AN OPERATOR THEORY APPROACH TO NONLINEAR CONTROLLER DESIGN

Thesis by
Constantin George Economou

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
(Submitted June 20, 1985)

στην ομορφη Stella
(οπως μαγευτικα προβαλλει σε κιτρινες φωτογραφιες του '50...)

Acknowledging...

It is with my deepest appreciation that I acknowledge my advisor, Professor Manfred Morari, for his continuous care, support and encouragement during my graduate studies. His insightful criticism, comments and suggestions guided my initiation and progress in the fascinating world of academic research. I am obliged for the ample freedom he allowed me in conducting my work.

The financial support from the National Science Foundation through the University of Wisconsin and Caltech is gratefully acknowledged.

The company of modern (and ancient) Hellenes was both exciting and helpful: Makis Caracotsios had off-the-shelf solutions to all my computational problems. Evanghelos Zafiriou meant trouble and heated discussions, sparking new ideas. Alex Seressiotis was around and that meant a lot to me. Leonidas Paparizos amply offered his math skills and intuition.

And then there was John Doyle, providing the inspiration and motivation that only a free spirit can.

It is unfortunate that I cannot mention everyone who contributed to my personal growth and enjoyment of this period of my life. Friendly faces abound in my memories:

The Economou family in Athens, Stella and George and Filia and Marianthe, my *oh! so dear* family.

My teachers at Madison and Caltech for nurturing a high quality environment of learning, intelligence, challenge and freedom of thought, in particular Professors W.H. Ray and Peter M. Thompson.

My friends and co-workers at UW and Caltech for their warmhearted support and fruitful interactions: from the "seniors" Carlos Garcia, Alok Saboo and Brad Holt, to the "youngsters" Chris Webb and Roland Heersink.

Donald Knuth (1984) for TEX.

David Thomas, T.S. Elliot, Iannis Minis, Ian Curtis and the bay area Residents for their company in isolation...

It is inappropriate to thank my wife Vaso "B.B." Leventakou. After all these years of love and togetherness it would be like thanking myself.

ABSTRACT

Strong similarities between control theory and the theory on the solution of operator equations have been observed and basic results in control theory have been derived from operator theory arguments. The purpose of this work is to investigate the theory of controller design as an application of basic operator theory principles and to establish a unified framework in which control theory can benefit from a "rich" operator theory. The major impact is anticipated in nonlinear feedback control theory: controller design can be formulated as selection of an iterative algorithm to solve a nonlinear operator equation corresponding to the control objective. As an example, controllers induced by the method of successive substitution and the Newton method are introduced and the corresponding analysis and synthesis issues are studied. Applied to linear systems, the proposed concepts have a straightforward interpretation in terms of familiar notions in linear controller design theory. Applications are presented and extensions of the current results are suggested to conclude the thesis.

<u>CONTENTS</u>

LIST OF EXAMPLES

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Control theory has had positive interactions with operator theory so far. A number of control researchers have either noticed (Åström and Wittenmark 1984) or alluded to (Goodwin and Sin 1984) and some have used an underlying duality between control theory and the theory on the solution of operator equations, to establish strong quantitative results.

It suffices to mention a few : Kalman (1960) was first to use Contraction Principle arguments to study the stability of autonomous discrete nonlinear systems. Using his original notation, the autonomous dynamic system described by the discrete evolution equation

$$x(t_{k+1}) = H(x(t_k)) \cdot x(t_k)$$

is globally asymptotically stable if the norm of the operator $H(x)$ is less than 1 for all x.

Zames (1966) used an input-output formalism, abstracting the system to an operator mapping $\mathcal{L}_2$ input functions to extended $\mathcal{L}_2$ output functions. He established closed loop stability on the condition that the induced system operator norm is less than 1 (Small Gain Theorem), in effect employing the same Contraction Mapping Principle in a different framework. The well known conicity and circle conditions surfaced by applying the principle to the special case of static nonlinearities. In the same work it is shown, how the sufficient conditions of the contraction mapping theorem can be strengthened to necessary and sufficient conditions by manipulating input-output relations. Safonov (1980) extended Zames' results to a more general setting based on the theory of topological separation in function spaces.

Rosenbrock (1974) introduced the Gerschgorin theorem, used in the analysis of iterative linear equation solution algorithms (Jacobi, Gauss-Seidel etc., Ortega and Rheinboldt 1970), to study the stability of decentralized control structures.

Perhaps the most profound demonstration of the duality for the case of linear systems, is the work of Doyle and Stein (Doyle and Stein 1979 and 1981, Stein 1981). The authors noticed that loop shaping methods produced compensators that inherently contained the inverse of the linear system operator. They showed how to construct inverses of linear dynamical systems by adjusting the noise parameters of Linear Quadratic Regulator compensators. To analyse the robust stability of the resulting control structures the method of Singular Value Decomposition, commonly used in studies of sensitivity of linear operator inversion, naturally surfaced(Lehtomaki 1981, Doyle and Stein 1981).

On the othe hand, basic control theory results appear in operator theory. For example, the von Neumann convergence analysis (Richtmeyer and Morton 1969) for linear partial differential equation solution procedures is a basic form of the Nyquist stability criterion. In its context, recursive solution schemes are Fourier transformed from time to frequency domain and the stability of the scheme is established as a standard application of the Nyquist stability theorem.

To the above, Chapter IV adds a number of quantitative results in support of the duality argument. Practically all the related results are confined to analysis issues, such as stability and robustness. The implications to synthesis and design have yet to be studied. The focus of this work is feedback controller design: if the design problem were to be formulated as an operator equation, it could benefit in both the analysis and synthesis aspects from a relatively well developed theory on the solution of operator equations.

For linear systems no major gains are to be expected, since the implied operator inversion has been either explicitly (Garcia and Morari 1982, Zames 1981) or implicitly (Stein 1981) used in control studies; still some insight in the issue of inverting control might be gained.

Compared to linear systems however, there are very few results in nonlinear controller design, mainly on stability analysis, while only limited attempts to a general synthesis theory have been reported: For autonomous systems, methods emanating from stability analysis of differential equations (Lyapunov 1892, LaSalle and Lefschetz 1961) have been employed, Kalman (1960) being first in providing a formal adaptation to systems analysis. For closed loop systems, Popov's (1973) and Zames' (1966) respective approaches are prominent, although their impact is practically confined to systems that can be represented by interconnections of linear dynamic operators with static nonlinear elements. Later, a number of applications of the Lyapunov method to stability (and instability) of systems with nonstationary nonlinear elements were reported by Eastern researchers (Skorodinskii 1981 and 1982, Barabanov 1982, Molchanov and Pyatinskii 1982).

An ad hoc approach characterizes controller synthesis methods for nonlinear systems, combined with extrapolations of linear controller design techniques like adaptive control (Goodwin and Sin 1984) and robust control (Doyle 1984).

Nonlinear Optimal Control (Athans and Falb 1966, Bryson and Ho 1975), is historically the first direct effort towards nonlinear control synthesis that proved useful in a number of aerospace and other non-continuous (batch) applications. Although nonlinear optimal control has been criticized from different viewpoints, namely that weighting does not provide insight to the final design properties, on line solution of the two point boundary value problem is in general prohibitive to applications etc., it appears that the real shortcoming is that even for linear systems stability is not assured in the face of modelling error (Doyle 1978).

Nonlinear Internal Model Control was introduced by Economou and Morari (1985) as an extension of a linear controller design technique. The controller structure is based on the inverse operator of the nonlinear system, obtained by analytical or numerical inversion.

The theory of Hunt, Su and Meyer (1983), provides a theoretically rigorous alternative to controller design for nonlinear systems. The basic idea is to derive analytic state, input and output transformations that reduce a nonlinear system to a linear dynamic element. Subsequently basic linear control techniques can be employed. The first successful implementation of the method was recently reported (Meyer 1985). It has been pointed out however that the method is complicated in general, involving the need for analytical solution of a set of recursive partial differential equations. The conditions for the existence of solutions are non-trivial to establish. Another point of criticism is that the linear controller is designed for the transformed inputs and outputs, which in general have no physical interpretation at all.

The purpose of this thesis is to establish the duality between controller design and algorithm development for the solution of operator equations. A number of meaningful control objectives can be formulated as operator inversion and/or optimization problems, which in turn have a good practical as well as theoretical support. This framework allows us to address nonlinear controller design in a general and intuitively clear manner, which at the same time naturally extends familiar notions from linear systems control. At the present stage, no hope is expressed to exhaust the subject, but rather to expose a concept and illustrate its applications.

In Chapter II the notation and necessary computational tools are introduced. Chapter III summarizes operator equation theory fundamentals, while a hybrid algorithm motivated by systems control requirements is developed. In Chapter IV the basic duality features are presented and the equivalence of the control problem to an operator equation is established for a class of control objectives. A general analysis theory is detailed in Chapter V. Control law synthesis is investigated in the next Chapters, where Contraction Principle (Chapter VI) and Newton (Chapter VII) controllers are introduced and analyzed in the light of iterative operator equation solution algorithms. Chapter VIII summarizes and concludes the work.

CHAPTER II

PRELIMINARIES

1. ASSUMPTIONS

The systems considered are governed by the vector ordinary differential equations:

$$\frac{\partial x}{\partial t} = f(x, u(t)) \tag{II.1}$$

where $x \in \mathbf{R}^n$ is the state of the system and for every $t \in [0, \infty)$ $u(t) \in \mathbf{R}^m$ is the input, with the corresponding output map ($y \in \mathbf{R}^m$):

$$y = g(x) \tag{II.2}$$

Example II.1 : The reversible exothermic reaction

$$A \underset{k_R}{\overset{k_A}{\rightleftharpoons}} R$$

is carried out in the ideal stirred tank reactor of fig. II.1. The following differential-algebraic equations model the dynamics of the reactor. They are derived from differential mass and energy balances:

$$\begin{aligned}
\frac{\partial x_1}{\partial t} &= \frac{1}{\tau}(A_i - x_1) - k_A e^{-Q_A/Rx_3} x_1 + k_R e^{-Q_R/Rx_3} x_2 \\
\frac{\partial x_2}{\partial t} &= \frac{1}{\tau}(R_i - x_2) + k_A e^{-Q_A/Rx_3} x_1 - k_R e^{-Q_R/Rx_3} x_2 \\
\frac{\partial x_3}{\partial t} &= \frac{1}{\tau}(T_i - x_3) + \frac{-\Delta H}{\rho c_p}(k_A e^{-Q_A/Rx_3} x_1 - k_R e^{-Q_R/Rx_3} x_2) \\
y &= \frac{x_2}{x_1 + x_2}
\end{aligned}$$

where $x_1 = \mathrm{A}_o$ (concentration of A in the tank), $x_2 = \mathrm{R}_o$ (concentration of R in the tank) and $x_3 = T_o$ (tank temperature).

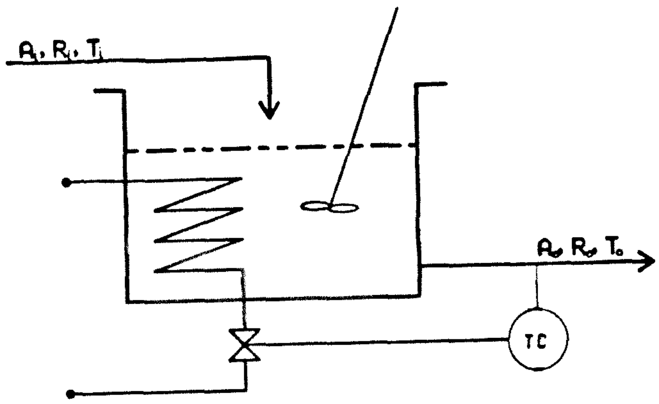

FIGURE II.1 : *Continuous Stirred Tank Reactor with reversible reaction. A: concentration of A, R: concentration of R, T: temperature. Subscript i denotes feed conditions, subscript o denotes tank (and outlet) conditions.*

A system of this form is treated in more detail in Chapters V, VI and VII.

The systems considered are assumed to have one and only one solution $x(t)$ for given data t^0, $x(t^0)$ and $u(t)$. Conditions for existence and uniqueness of solutions of (II.1) can be found in standard texts (Holtzman 1970, Vidyasagar 1978) and will not be discussed here. Instead, following Kalman (1960), the dynamic systems considered are defined axiomatically through the following set of axioms:

1.i. (Existence) There is a function $\chi(t; t^0, x^0, u(t))$, called the transition (or, state evolution) function, satisfying (II.1)): $\dfrac{\partial \chi(t; t^0, x^0, u(t))}{\partial t} = f(\chi(t; t^0, x^0, u(t)), u(t)), \quad t \geq t^0$

1.ii. $\chi(t; t^0, x^0, u(t))$ is defined for all x^0, t^0, $t \geq 0$

1.iii. $\chi(t^0; t^0, x^0, u(t)) = x^0$ for all t^0, x^0

1.iv. (Uniqueness) $\chi(t^2; t^1, \chi(t^1; t^0, x^0, u(t)), u(t)) = \chi(t^2; t^0, x^0, u(t))$ for all x^0, t^0, t^1, t^2

1.v. $\chi(t; t^0, x^0, u(t))$ is continuous with respect to all arguments

1.vi. $\chi(t; t^0, x^0, u(t))$ is differentiable with respect to all arguments

Example II.2 : Consider the system (Vidyasagar 1978) described by

$$\frac{\partial x_1}{\partial t} = \alpha x_1(\beta^2 - x_1^2)$$

$$\frac{\partial x_2}{\partial t} = -1 + x_1 u$$

with $t^0 = 0$, $x_1^0 = x_{10}$, $x_2^0 = x_{20}$ *and* $u(t) = 0$. *It can be verified that the system has a unique solution*

$$\chi_1(t; t^0, x^0, u) = \left[1 + (\frac{\beta^2}{x_{10}^2} - 1)e^{-2\beta^2\alpha t}\right]^{-\frac{1}{2}}$$

$$\chi_2(t;t^0,x^0,u) = x_{20} - t$$

which satisfies axioms 1.i - 1.vi .

Example II.3 : *Consider the linear system described by*

$$\frac{\partial x}{\partial t} = Ax + Bu$$

with $t^0 = 0$, $x^0 = x_0$ and $u = u_f$ (constant). The differential equations can be integrated analytically (Kailath 1980), yielding

$$\chi(t;x^0,u_f) = e^{At}x_0 + (e^{At} - I)A^{-1}Bu_f$$

2. NOTATION

The system inputs are assumed to be piecewise constant functions to reduce the problem at hand to a finite dimensional space. The letter s is used as a superscript to mark the discrete time. The s^{th} sampling interval extends from t^s to t^{s+1}; $T = t^{s+1} - t^s$ is the (constant) sampling time; x^s is the state at t^s; u^s is the system input, held constant over $(t^s, t^{s+1}]$.

In the discrete setting of the study, $\chi(t_2;t_1,x,u)$ is the solution of (1) at time t_2, for $u(t) = u$ $(t_1 < t \leq t_2)$, and initial condition $\chi(t_1;t_1,x,u) = x$; χ^s will denote the state of the system at $t = t^{s+1}$, i.e. x^{s+1}:

$$\chi^s \stackrel{\text{def}}{=} x^{s+1} = \chi(t^s + T; t^s, x^s, u^s)$$

Since (II.1) is stationary: $\chi(t_1 + \Delta t; t_1, x, u) = \chi(t_2 + \Delta t; t_2, x, u)$, time will be dropped from the parameter list and the following convention will be used:

$$\chi^s = \chi(T;x^s,u^s) = \chi(t^s + T; t^s, x^s, u^s)$$

Example II.4 : *Referring to the system of the Example II.2,*

$$\chi_1^s = \left[1 + (\frac{\beta^2}{x_{10}^2} - 1)e^{-2\beta^2\alpha T}\right]^{-\frac{1}{2}}$$
$$\chi_2^s = x_{20} - T$$

The derivatives of χ^s with respect to x^s and u^s will be $\Phi^s \left(\stackrel{\text{def}}{=} \frac{\partial \chi^s}{\partial x^s}\right)$ and $\Gamma^s \left(\stackrel{\text{def}}{=} \frac{\partial \chi^s}{\partial u^s}\right)$ respectively. $y^s \left(\stackrel{\text{def}}{=} g(x^s)\right)$ is the system output at t^s.

Example II.5 : For the system in example II.2 straightforward calculations show:

$$\Phi^s = \begin{pmatrix} \frac{\partial \chi_1^s}{\partial x_{10}} & \frac{\partial \chi_1^s}{\partial x_{20}} \\ \frac{\partial \chi_2^s}{\partial x_{10}} & \frac{\partial \chi_2^s}{\partial x_{20}} \end{pmatrix} = \begin{pmatrix} \frac{\beta^2}{x_{10}^3} e^{-2\beta^2\alpha T} \left[1 + (\frac{\beta^2}{x_{10}^2} - 1) e^{-2\beta^2\alpha T}\right]^{-\frac{3}{2}} & 0 \\ 0 & 1 \end{pmatrix}$$

$C^s \stackrel{\text{def}}{=} \left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right|_{\varsigma = x^s}$ will denote the derivative of the output map (II.2) at $x = x^s$.

Capital letters F, P, Q etc., denote operators and script letters $\mathcal{Y}$, $\mathcal{X}$, $\mathcal{U}$ are used for metric spaces.

3. STATE DERIVATIVES

The state derivatives with respect to initial conditions (Φ^s) and inputs (Γ^s) frequently appear throughout the paper. Except for simple cases, as in example II.2, analytical evaluation of the derivatives is rarely possible. In the following a computational theory for related quantities is presented. The statements are proved in Appendix II.

a. Φ^s is the solution at $t = t^{s+1}$ of the initial value problem

$$\frac{\partial \Phi(t)}{\partial t} = \left.\frac{\partial f(\varsigma, \xi)}{\partial \varsigma}\right|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}} \Phi(t) \tag{II.3}$$

with initial conditions

$$\Phi(t^s) = I \tag{II.4}$$

It is implied that $\chi(t; x^s, u^s)$ has been already computed by solving the initial value problem (II.1) and subsequently (II.3) is integrated along the trajectory $\chi(t; x^s, u^s)$. Caracotsios and Stewart (1984) and Leis and Kramer (1984), using the linearity of (II.3) with respect to $\Phi(t)$, showed that no additional integration is necessary: (II.1) is integrated by a standard iterative predictor-corrector implicit integration formula. When convergence has been attained, the derivative term on the right hand side of (II.3) becomes available. Then (II.3) is a linear system of equations which is solvable in one forward step, with no additional iterations necessary. In this procedure, effectively a system of n ordinary differential equations is solved, instead of the complete system of $n + n^2$ equations of (II.1) and (II.3) . General purpose software implementing the procedure is currently available (Caracotsios and Stewart 1984).

Example II.6 : For a linear system (III.3) and (III.4) can be integrated analytically (see Appendix II), yielding:

$$\Phi^s = e^{AT} \tag{II.5}$$

i.e., the state transition matrix of discrete state space representations of linear systems (see also example II.3). $C\Phi^s$ then is the autonomous system response $(u(t) = 0)$ at time $t = T$, to a unit strength impulse in every state at $t = 0$ when the system is at equilibrium.

b. Γ^s is the solution at $t = t^{s+1}$ of the initial value problem

$$\frac{\partial \Gamma(t)}{\partial t} = \left.\frac{\partial f(\varsigma, \xi)}{\partial \varsigma}\right|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}} \Gamma(t) + \left.\frac{\partial f(\varsigma, \xi)}{\partial \xi}\right|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}} \tag{II.6}$$

with initial conditions

$$\Gamma(t^s) = 0 \tag{II.7}$$

As in the Φ^s case, no additional integration is necessary for the computation of Γ^s.

Example II.7 : For a linear system (II.6) *and* (II.7) *can be integrated analytically (Appendix II), yielding:*

$$\Gamma^s = (e^{AT} - I)A^{-1}B \tag{II.8}$$

Γ^s has a familiar interpretation for linear systems: It is the input matrix of discrete state space representations (see also example II.3). $C\Gamma^s$ is the system response at $t = T$ to a unit step change in all the inputs at $t = 0$ when the system is at equilibrium.

c. In the remainder of the section, a computational theory for the second order state derivatives is detailed. Second order derivatives appear in the stability analysis of the Newton controllers in Chapter VII. The reader may skip this material at a first reading.

In the beginning of the section, first order derivatives were shown to be matrices resulting from the differentiation of a vector function with respect to a vector variable. Second order differentiation involves the derivative of a matrix function with respect to a vector variable and some new quantities have to be introduced (†).

Definition II.1 : A bilinear matrix H of dimension $m \times l \times n$ is an ordered collection of real numbers $h_{\alpha\beta\gamma}$, $\alpha = 1,2\ldots m$, $\beta = 1,2\ldots l$, $\gamma = 1,2\ldots n$. It is highlighted by inclusion in brackets: $\{H\}$, or $\{h_{\alpha\beta\gamma}\}$.

Definition II.2 : The derivative of an $m \times n$ matrix function $F(u) = [f_{\alpha\gamma}(u)]$ with respect to the l-dimensional vector $u = [u_\beta]$ is the $m \times l \times n$ bilinear matrix $\{H\}$ with elements

$$h_{\alpha\beta\gamma} \overset{\text{def}}{=} \frac{\partial f_{\alpha\gamma}(u)}{\partial u_\beta}$$

(†) A complete treatment of the bilinear operators appearing in matrix differentiation is given by Rall (1979). Here only the necessary notions are discussed.

Definition II.3 : The right dot product of an $m \times l \times n$ bilinear matrix $\{F\}$ with a regular $n \times k$ matrix G is the $m \times l \times k$ bilinear matrix $\{H\} = \{F\} \bullet G$ with elements

$$h_{\alpha\beta\gamma} = \sum_{\delta=1}^{\delta=n} f_{\alpha\beta\delta}\, g_{\delta\gamma}$$

Definition II.4 : The left dot product of an $m \times l \times n$ bilinear matrix $\{F\}$ with a regular $k \times m$ matrix G is the $k \times l \times n$ bilinear matrix $\{H\} = G \bullet \{F\}$ with elements

$$h_{\alpha\beta\gamma} = \sum_{\delta=1}^{\delta=m} g_{\alpha\delta}\, f_{\delta\beta\gamma}$$

Definition II.5 : The circle product of an $m \times l \times n$ bilinear matrix $\{F\}$ and a regular $l \times k$ matrix G is the $m \times k \times n$ bilinear matrix $\{H\} = \{F\} \circ G$ with elements

$$h_{\alpha\beta\gamma} = \sum_{\delta=1}^{\delta=l} f_{\alpha\delta\gamma} g_{\delta\gamma}$$

Two useful differentiation properties are stated. The proofs are in Appendix II.

Differentiation of a product

$$\{\frac{\partial A(u)B(u)}{\partial u}\} = \{\frac{\partial A(u)}{\partial u}\} \bullet B(u) + A(u) \bullet \{\frac{\partial B(u)}{\partial u}\}$$

Differentiation of composition

$$\{\frac{\partial A(x(u))}{\partial u}\} = \{\frac{\partial A(x)}{\partial x}\} \circ \frac{\partial x}{\partial u}$$

In this context the second and higher order derivatives can be computed. In Appendix II the following statements are proven.

$\{\Phi_x^s\} \stackrel{\text{def}}{=} \{\frac{\partial \Phi^s}{\partial x^s}\} = \{\frac{\partial^2 \chi(t; x^s, u^s)}{\partial (x^s)^2}\}$ is the solution at $t = t^{s+1}$ of the initial value problem

$$\frac{\partial\{\Phi_x(t)\}}{\partial t} = \frac{\partial f(\varsigma, \xi)}{\partial \varsigma}\bigg|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}} \bullet \{\Phi_x(t)\} + \{\frac{\partial^2 f(\varsigma, \xi)}{\partial \varsigma^2}\bigg|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}}\} \circ \Phi(t) \bullet \Phi(t) \quad \text{(II.9)}$$

where $\Phi(t)$ as in (II.3) , and initial condition

$$\{\Phi_x(t^s)\} = \{0\}$$

Example II.8 : For a linear system :

$$\{\Phi^s_x\} = \{0\} \tag{II.10}$$

(Proof in Appendix II.)

Then, $\{\Phi^s_u\} \stackrel{\text{def}}{=} \{\frac{\partial \Phi^s}{\partial u^s}\} = \{\frac{\partial^2 \chi(T; x^s, u^s)}{\partial x^s \partial u^s}\}$ is the solution at $t = t^{s+1}$ of the initial value problem

$$\{\frac{\partial \Phi_u(t)}{\partial t}\} = \frac{\partial f(\varsigma,\xi)}{\partial \varsigma}\Big|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}} \bullet \{\Phi_u(t)\} + \{\frac{\partial^2 f(\varsigma,\xi)}{\partial \varsigma^2}\Big|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}}\} \circ \Gamma(t) \bullet \Phi(t)$$
$$+ \{\frac{\partial^2 f(\varsigma,\xi)}{\partial \varsigma \partial \xi}\Big|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}}\} \bullet \Phi(t) \tag{II.11}$$

where $\Gamma(t)$ as in (II.6) , and initial condition

$$\{\Phi_u(t^s)\} = \{0\}$$

Example II.9 : For a linear system :

$$\{\Phi^s_u\} = \{0\} \tag{II.12}$$

(Proof in Appendix II.)

Also, $\{\Gamma^s_x\} \stackrel{\text{def}}{=} \{\frac{\partial \Gamma^s}{\partial x^s}\} = \{\frac{\partial^2 \chi(T; x^s, u^s)}{\partial u^s \partial x^s}\}$ is equal to $\{\Phi^s_u\}$ and formulæ (II.11) and (II.12) can be used directly.

Finally, $\{\Gamma^s_u\} \stackrel{\text{def}}{=} \{\frac{\partial \Gamma^s}{\partial u^s}\} = \{\frac{\partial^2 \chi(T; x^s, u^s)}{\partial (u^s)^2}\}$ is the solution at $t = t^s + 1$ of the initial value problem

$$\frac{\partial \{\Gamma_u(t)\}}{\partial t} = \frac{\partial f(\varsigma,\xi)}{\partial \varsigma}\Big|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}} \bullet \{\Gamma_u(t)\} + \{\frac{\partial^2 f(\varsigma,\xi)}{\partial \varsigma^2}\Big|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}}\} \circ \Gamma(t) \bullet \Gamma(t)$$
$$+ \{\frac{\partial^2 f(\varsigma,\xi)}{\partial \varsigma \partial \xi}\Big|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}}\} \bullet \Gamma(t) + \{\frac{\partial^2 f(\varsigma,\xi)}{\partial \xi \partial \varsigma}\Big|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}}\} \circ \Gamma(t)$$
$$+ \{\frac{\partial^2 f(\varsigma,\xi)}{\partial \xi^2}\Big|_{\substack{\varsigma = \chi(t; x^s, u^s) \\ \xi = u^s}}\} \tag{II.13}$$

with initial condition

$$\{\Gamma_u(t^s)\} = \{0\}$$

Example II.10 : For linear systems:

$$\{\Gamma^s_u\} = \{0\} \tag{II.14}$$

(Proof in Appendix II.)

It should be noted that equations (II.9) , (II.11) and (II.13) are ordinary differential equations of dimension $n \times n \times n$, $n \times n \times m$, $n \times m \times n$ and $n \times m \times m$ respectively, which are solved by any standard integrator by rearrangement to one dimensional vector differential equations.

Example II.11 : The following example of a linear system is illustrative of the different notions introduced so far. Consider the linear, continuous system

$$\begin{aligned}
\frac{\partial x_1}{\partial t} &= 4x_1 - 3x_2 + u \\
\frac{\partial x_2}{\partial t} &= x_1 \\
y &= x_2
\end{aligned}$$

with a sampling time of $T = 0.1$. Then, irrespective of initial conditions:

$$\Phi^s = \begin{pmatrix} 1.470 & -0.367 \\ 0.122 & 0.983 \end{pmatrix} \qquad \Gamma^s = \begin{pmatrix} 0.122 \\ 5.7 \times 10^{-3} \end{pmatrix}$$

$$\Phi^s_x = 0 \qquad \Phi^s_u = \Gamma^s_x = 0 \qquad \Gamma^s_u = 0$$

From example II.8, the discrete state space description of the system for the given sampling time is

$$\begin{aligned}
x_1^{s+1} &= 1.470x_1^s - 0.367x_2^s + 0.122u^s \\
x_2^{s+1} &= 0.1223x_1^s + 0.983x_2^s + 5.7 \times 10^{-3}u^s \\
y^{s+1} &= x_2^{s+1}
\end{aligned}$$

It is observed that the state transition matrix is equal to Φ^s and that the input matrix is equal to Γ^s. This is by no means coincidence, but rather a unique characteristic of linear systems. It demonstrates that for linear systems the derivative of the system operator is itself, a fact discussed in Chapter III.

4. REMARKS

Every continuous system, when sampled at a period T, gives rise to a discrete system of the form

$$\chi^s = x^{s+1} = F(x^s, u^s) \tag{II.1'}$$

Explicit functional relationships of this form, are possible only in the (rare) occasion when (II.1) can be integrated analytically, as is for example the case of linear systems. On the other hand, description (II.1′) is more general than (II.1), because not every discrete system arises by sampling a continuous system, as is the case of linear discrete systems with an odd number of negative real poles (Kalman 1960).

In any case, although only continuous systems of the form (II.1) are treated in the present context, the theory applies equally well to discrete systems of the form (II.1′). Then $\Phi^s = \partial F(x^s, u^s)/\partial x^s$, $\Gamma^s = \partial F(x^s, u^s)/\partial u^s$ and every result in the chapters to follow holds for systems of this form.

Appendix II

Computation of the state derivatives w.r.t. initial conditions

Define the function $\Phi(t) \stackrel{\text{def}}{=} \dfrac{\partial\chi(t;x^s,u^s)}{\partial x^s}$, $t \in [t^s,\infty)$. By definition $\Phi^s = \dfrac{\partial\chi(T;x^s,u^s)}{\partial x^s} = \Phi(t^{s+1})$. Differentiate $\Phi(t)$ with respect to t to obtain:

$$\begin{aligned}\frac{\partial\Phi(t)}{\partial t} &= \frac{\partial}{\partial t}\frac{\partial\chi(t;x^s,u^s)}{\partial x^s} = \frac{\partial}{\partial x^s}\frac{\partial\chi(t;x^s,u^s)}{\partial t}\\ &= \frac{\partial f(\chi(t;x^s,u^s),u^s)}{\partial x^s} = \left.\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\right|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}} \frac{\partial\chi(t;x^s,u^s)}{\partial x^s}\end{aligned}$$

Now substitute $\Phi(t)$ from its definition:

$$\frac{\partial\Phi(t)}{\partial t} = \left.\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\right|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}} \Phi(t) \tag{II.3}$$

At $t = t^s$, $\chi(t^s;x^s,u^s) = x^s$ so that:

$$\Phi(t^s) = \frac{\partial x^s}{\partial x^s} = I \tag{II.4}$$

Summing up, Φ^s is the solution at $t = t^{s+1}$ of the initial value problem (II.3) with initial conditions (II.4) .

For a linear system, using the notation of Section II, (II.3) and (II.4) become:

$$\frac{\partial\Phi(t)}{\partial t} = A\Phi$$

$$\Phi(t^s) = I$$

The solution is found in any standard textbook (Kailath 1980).

$$\begin{aligned}\Phi^s = \Phi(t^{s+1}) &= e^{A(t^{s+1}-t^s)}\Phi(t^s)\\ &= e^{AT}\cdot I\\ &= e^{AT}\end{aligned} \tag{II.5}$$ ■

Computation of state derivatives w.r.t. inputs:

Define the function $\Gamma(t) \stackrel{\text{def}}{=} \dfrac{\partial\chi(t;x^s,u^s)}{\partial u^s}$, $t \in [t^s,\infty)$. By definition $\Gamma^s = \dfrac{\partial\chi(T;x^s,u^s)}{\partial u^s} = \Gamma(t^{s+1})$. Differentiate $\Gamma(t)$ with respect to t to obtain:

$$\frac{\partial\Gamma(t)}{\partial t} = \frac{\partial}{\partial t}\frac{\partial\chi(t;x^s,u^s)}{\partial u^s} = \frac{\partial}{\partial u^s}\frac{\partial\chi(t;x^s,u^s)}{\partial t}$$

$\chi(t;x^s,u^s)$ is the solution of the state evolution equation (II.1), therefore:

$$\begin{aligned}\frac{\partial\Gamma(t)}{\partial t} &= \frac{\partial f(\chi(t;x^s,u^s)}{\partial u^s} \qquad \text{(chain differentiation)}\\ &= \left.\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\right|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}} \frac{\partial\chi(t;x^s,u^s)}{\partial u^s} + \left.\frac{\partial f(\varsigma,\xi)}{\partial\xi}\right|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\end{aligned}$$

Substitute $\Gamma(t)$ from its definition in the above expression:

$$\frac{\partial\Gamma(t)}{\partial t} = \left.\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\right|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}} \Gamma(t) + \left.\frac{\partial f(\varsigma,\xi)}{\partial\xi}\right|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}} \tag{II.6}$$

At $t=t^s$, $\chi(t^s;x^s,u^s)=x^s$ which is independent of u^s. It follows:

$$\Gamma(t^s) = \frac{\partial\chi(t^s;x^s,u^s)}{\partial u^s} = 0 \tag{II.7}$$

Summing up, Γ^s is the solution at $t=t^{s+1}$ of the initial value problem (II.6) with initial conditions (II.7) .

For a linear system, using the notation of Section II, (II.6) and (II.7) become:

$$\frac{\partial\Gamma(t)}{\partial t} = A\Gamma(t)+B$$

$$\Gamma(t^s)=0$$

The solution is found in any standard textbook (Kailath 1980).

$$\begin{aligned}\Gamma^s=\Gamma(t^{s+1}) &= e^{A(t^{s+1}-t^s)}\Gamma(t^s) - e^{A(t^{s+1}-t^s)}(e^{-A(t^{s+1}-t^s)}-I)A^{-1}B\cdot 1\\ &= -e^{AT}(e^{-AT}-I)A^{-1}B\\ &= (e^{AT}-I)A^{-1}B\end{aligned} \tag{II.8}$$ ■

Proof of the bilinear matrix properties

a. Product differentiation : Consider the $(\alpha\beta\gamma)^{th}$ element of $\{\frac{\partial AB}{\partial u}\}$. It is

$$\begin{aligned}\{\frac{\partial AB}{\partial u}\}_{\alpha\beta\gamma} \stackrel{\text{def}}{=} \frac{\partial(AB)_{\alpha\gamma}}{\partial u_\beta} &= \frac{\partial\sum_\delta(A_{\alpha\delta}B_{\delta\gamma})}{\partial u_\beta}\\ &= \sum_\delta\frac{\partial A_{\alpha\delta}}{\partial u_\beta}B_{\delta\gamma} + \sum_\delta A_{\alpha\delta}\frac{\partial B_{\delta\gamma}}{\partial u_\beta}\end{aligned}$$

which dy definitions II.3 and 4 is the $(\alpha\beta\gamma)^{th}$ element of $\frac{\partial A}{\partial u}\bullet B + A\bullet\frac{\partial B}{\partial u}$. ■

b. Composition Differentiation : Consider the $(\alpha\beta\gamma)^{th}$ element of $\{\frac{\partial A(x(u))}{\partial u}\}$. It is

$$\{\frac{\partial A(x(u))}{\partial u}\}_{\alpha\beta\gamma} = \frac{\partial A_{\alpha\gamma}(x(u))}{\partial u_\beta} = \sum_\delta \frac{\partial A_{\alpha\gamma}(x)}{\partial x}\frac{\partial x_\delta}{\partial u_\beta}$$

By Def II.5 the term on the right is the circle product $\{\frac{\partial A}{\partial x}\} \circ \frac{\partial x}{\partial u}$ ∎

Computation of second order state derivatives

Φ_x: For the second order state derivatives w.r.t. initial conditions, consider the function $\{\Phi_x(t)\} \stackrel{\text{def}}{=} \{\frac{\partial\Phi(t)}{\partial x^s}\} = \{\frac{\partial^2\chi(t;x^s,u^s)}{\partial(x^s)^2}\}$. By definition, $\{\Phi_x^s\} = \{\Phi_x(t^{s+1})\}$. Differentiate $\{\Phi_x(t)\}$ with respect to t to obtain:

$$\begin{aligned}
\frac{\partial\{\Phi_x(t)\}}{\partial t} &= \frac{\partial}{\partial t}\{\frac{\partial\Phi(t)}{\partial x^s}\} = \{\frac{\partial}{\partial x^s}\frac{\partial\Phi(t)}{\partial t}\} \\
&= \{\frac{\partial}{\partial x^s}\left[\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\Phi(t)\right]\} \\
&= \{\frac{\partial}{\partial x^s}\left[\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right]\}\bullet\Phi(t) + \frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\bullet\{\frac{\partial\Phi(t)}{\partial x^s}\} \\
&= \{\frac{\partial^2 f(\varsigma,\xi)}{\partial\varsigma^2}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\}\circ\Phi(t)\bullet\Phi(t) \\
&\quad + \frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\bullet\{\Phi_x(t)\}
\end{aligned} \tag{II.9}$$

At $t = t^s$,

$$\{\Phi_x(t^s)\} \stackrel{\text{def}}{=} \{\frac{\partial\Phi(t^s)}{\partial x^s}\} = \{\frac{\partial I}{\partial x^s}\} = \{0\}$$ ∎

For a linear system $\{\frac{\partial^2 f(\varsigma,\xi)}{\partial\varsigma^2}\} = \{\frac{\partial^2(A\varsigma+B\xi)}{\partial\varsigma^2}\} = \{0\}$. Then (II.9) becomes

$$\frac{\partial\{\Phi_x(t)\}}{\partial t} = A\bullet\{\Phi_x(t)\}$$

with $\{\Phi_x(t^s)\} = \{0\}$, i.e. an unforced linear system with zero initial condition. Readily then $\{\Phi_x(t)\} = \{0\}$ and more specifically

$$\{\Phi_x^s\} = \{\Phi_x(t^{s+1})\} = \{0\} \tag{II.10}$$

Φ_u: The second order derivative of the states with respect to inputs is taken up next. For this purpose define the function $\{\Phi_u(t)\} \stackrel{\text{def}}{=} \{\frac{\partial\Phi(t)}{\partial u^s}\}$. By definition $\{\Phi_u^s\} = \{\Phi_u(t^{s+1})\}$. Differentiate

$\{\Phi_u(t)\}$ with respect to t to obtain:

$$\frac{\partial\{\Phi_u(t)\}}{\partial t} = \frac{\partial}{\partial t}\{\frac{\partial\Phi(t)}{\partial u^s}\} =$$

$$= \{\frac{\partial}{\partial u^s}\left[\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\Phi(t)\right]\}$$

$$= \{\frac{\partial}{\partial u^s}\left[\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right]\}\bullet\Phi(t) + \frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\bullet\{\frac{\partial\Phi(t)}{\partial u^s}\}$$

$$= \left[\{\frac{\partial^2 f(\varsigma,\xi)}{\partial\varsigma^2}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\}\circ\Gamma(t) + \{\frac{\partial^2 f(\varsigma,\xi)}{\partial\varsigma\partial\xi}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\}\right]\bullet\Phi(t)$$

$$+ \frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\bullet\{\Phi_u(t)\} \qquad \text{(II.11)}$$

At $t = t^s$, $\Phi(t) = I$ so that:

$$\{\Phi_u(t^s)\} = \{\frac{\partial I}{\partial u^s}\} = \{0\} \qquad \blacksquare$$

For a linear system the second order derivatives of $f(x,u)$ with respect to both arguments are zero and (II.11) becomes

$$\frac{\partial\{\Phi_u(t)\}}{\partial t} = A\bullet\{\Phi_u(t)\}$$

Again, it is a linear unforced system, with zero initial conditions, therefore

$$\{\Phi_u^s\} = \{\Phi_u(t^{s+1})\} = \{0\} \qquad \text{(II.12)} \ \blacksquare$$

$\underline{\Gamma_u}$: The second order derivative of the state with respect to the inputs is computed similarly. Define the function $\{\Gamma_u(t)\} \stackrel{\text{def}}{=} \{\frac{\partial\Gamma(t)}{\partial t}\} = \{\frac{\partial^2\chi(t;x^s,u^s)}{\partial u^s}\}$. By definition $\{\Gamma_u^s\} = \{\Gamma_u(t^{s+1})\}$. Differentiate $\{\Gamma_u(t)\}$ with respect to time to obtain:

$$\frac{\partial\{\Gamma_u(t)\}}{\partial t} = \frac{\partial}{\partial t}\{\frac{\partial\Gamma(t)}{\partial u^s}\} = \{\frac{\partial}{\partial u^s}\frac{\partial\Gamma(t)}{\partial t}\}$$

$$= \{\frac{\partial}{\partial u^s}\left[\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\Gamma(t) + \frac{\partial f(\varsigma,\xi)}{\partial\xi}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right]\}$$

$$= \{\frac{\partial}{\partial u^s}\left[\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right]\}\bullet\Gamma(t) + \frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\Big|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\bullet\{\frac{\partial\Gamma(t)}{\partial u^s}\}$$

$$+\left\{\frac{\partial}{\partial u^s}\frac{\partial f(\varsigma,\xi)}{\partial \xi}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right\}$$

$$=\left[\left\{\frac{\partial^2 f(\varsigma,\xi)}{\partial \varsigma^2}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right\}\circ\frac{\partial\chi(t;x^s,u^s)}{\partial u^s}+\left\{\frac{\partial^2 f(\varsigma,\xi)}{\partial\varsigma\partial\xi}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right\}\right]\bullet\Gamma(t)$$

$$+\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\bullet\{\Gamma_u(t)\}$$

$$+\left\{\frac{\partial^2 f(\varsigma,\xi)}{\partial\xi\partial\varsigma}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right\}\circ\frac{\partial\chi(t;x^s,u^s)}{\partial u^s}+\left\{\frac{\partial^2 f(\varsigma,\xi)}{\partial\xi^2}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right\}$$

$$=\left[\left\{\frac{\partial^2 f(\varsigma,\xi)}{\partial \varsigma^2}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right\}\circ\Gamma(t)+\left\{\frac{\partial^2 f(\varsigma,\xi)}{\partial\varsigma\partial\xi}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right\}\right]\bullet\Gamma(t)$$

$$+\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\bullet\{\Gamma_u(t)\}$$

$$+\left\{\frac{\partial^2 f(\varsigma,\xi)}{\partial\xi\partial\varsigma}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right\}\circ\Gamma(t)+\left\{\frac{\partial^2 f(\varsigma,\xi)}{\partial\xi^2}\bigg|_{\substack{\varsigma=\chi(t;x^s,u^s)\\ \xi=u^s}}\right\}\qquad\text{(II.13)}$$

At $t=t^s$, $\{\Gamma_u(t^s)\}=\{\frac{\partial\Gamma(t^s)}{\partial u^s}\}$, with $\{\Gamma(t^s)\}=\{0\}$. It follows

$$\{\Gamma_u(t^s)\}=\{0\}$$ ■

For a linear system the second order derivatives of $f(x,u)$ with respect to its arguments are zero and (II.13) is

$$\{\frac{\partial\Gamma_u(t)}{\partial t}\}=A\bullet\{\Gamma_u(t)\}$$

i.e. an unforced nonlinear system with zero initial conditions. It follows

$$\{\Gamma_u^s\}=\{\Gamma_u(t^{s+1})\}=\{0\}\qquad\text{(II.14)}$$ ■

CHAPTER III

OPERATOR EQUATION SOLUTION

Functional Analysis suggests useful algorithms for obtaining solutions to a wide class of operator equations. A well developed analysis theory supports the properties of the algorithms, namely existence of solutions, convergence, rate of convergence and sensitivity to approximation error. A few results, providing the basic analysis framework are summarized in the following. They can be found in any standard text on the subject (Kolmogorov and Fomin 1957, Kantorović and Akilow 1964, Ortega and Rheinboldt 1970, Rall 1979). Then a hybrid algorithm, arising in control applications, is developed and its convergence properties are investigated.

1. THE METHOD OF SUCCESSIVE SUBSTITUTIONS

Let the equation to be solved be:

$$P(z) = 0 \tag{III.1}$$

where P is an operator on a normed space Z. (III.1) is equivalent to the operator equation:

$$z = z + Q(P(z)) \stackrel{\text{def}}{=} F(z) \tag{III.2}$$

where Q is any continuous operator on the range of P with the property

$$Q(\xi) = 0 \quad \Leftrightarrow \quad \xi = 0 \tag{III.3}$$

(III.2) generates the recursive sequence:

$$z^{m+1} = F(z^m), \quad m = 0, 1, 2, \ldots \tag{III.4}$$

If F is continuous and if the sequence (III.4) converges, it converges to a solution of (III.1). The Contraction Mapping Theorem (Banach 1922, Cacciopoli 1931) and related results establish a computational theory for (III.4). Some definitions are in order:

Definition III.1 : If F is an operator on a Banach space Z, then any $z^* \in Z$ with the property $z^* = F(z^*)$ is called a fixed point of the operator.

Definition III.2 : An operator F on a Banach space Z into itself is called a contraction mapping of the closed ball (†) $\bar{U}(z^0, r)$ *if there exists a real number* θ, $0 \leq \theta < 1$ *such that:*

$$\|F(z^1) - F(z^2)\| \leq \theta \, \|z^1 - z^2\| \tag{III.5}$$

for all $z^1, z^2 \in \bar{U}(z^0, r)$.

Definition III.3 : The quantity θ *in Def. III.2 is called the contraction constant of F in* $\bar{U}(z^0, r)$.

Establishing the validity of (III.5) is a cumbersome undertaking. When F is differentiable the situation is significantly simplified: an exact characterization of the contraction property can be developed.

Lemma III.1 : (Curtain and Pritchard 1977) If the operator F on a Banach space Z is twice differentiable in a closed convex subset Ω *of Z then*

$$\|F(z^1) - F(z^2)\| \leq \sup_{0 \leq \lambda \leq 1} \|F'(\lambda z^1 + (1-\lambda) z^2)\| \, \|z^1 - z^2\|, \quad \forall z^1, z^2 \in \Omega \tag{III.6}$$

$$\|F(z^1) - F(z^2) - F'(z^2)(z^1 - z^2)\| \leq \tfrac{1}{2} \sup_{0 \leq \lambda \leq 1} \|F''(\lambda z^1 + (1-\lambda) z^2\| \, \|z^1 - z^2\|^2, \quad \forall z^1, z^2 \in \Omega \tag{III.7}$$

Lemma III.2 : Let the operator F on a Banach space Z be differentiable in $\bar{U}(z^0, r)$. *F is a contraction of* $\bar{U}(z^0, r)$ *if and only if*

$$\|F'(z)\| \leq \theta < 1, \quad \forall z \in \bar{U}(z^0, r) \tag{III.8}$$

where $\|\cdot\|$ *is any induced operator norm*(‡).

Proof : In Appendix III.

In the following, use of condition (III.8) will be made to characterize contraction properties of operators. When the operator is not differentiable the theorems remain unchanged except that (III.5) has to be used instead of (III.8) .

The basic Contraction Mapping Theorem is stated next:

Theorem III.1 : (Holtzman, 1970) If F maps a set $\bar{U}$ *into itself and F is a contraction mapping of the set with contraction constant* θ, *then:*

1. *F has a fixed point* z^* *in* $\bar{U}$.
2. z^* *is unique in* $\bar{U}$.
3. *The sequence* $\{z^m\}$ *defined by (III.4) converges to* z^* *with*

$$\|z^m - z^*\| \leq \theta^m \|z^0 - z^*\| \tag{III.9}$$

There is an inherent difficulty in applying Theorem III.1, namely that a set has to be found that maps into itself, which is not always easy. Rall (1979) replaces this condition with another one, more suitable for computation.

(†) *A closed ball* $\bar{U}(z^0, r)$ *is defined by:* $\bar{U}(z^0, r) = \{z \in Z : \|z - z^0\| \leq r\}$

(‡) *An induced operator norm is defined for every vector norm by:* $\|F\| = \sup_{z \neq 0} \frac{\|F(z)\|}{\|z\|}$

Theorem III.2 : (Rall 1979) If F is a contraction mapping of $\bar{U}(z^0, r)$ with contraction constant θ for

$$r \geq \frac{1}{1-\theta}\|F(z^0) - z^0\| \stackrel{\text{def}}{=} r^0 \quad , \quad \text{then :}$$

1. *F has a fixed point z^* in $\bar{U}(z^0, r^0)$.*
2. *z^* is unique in $\bar{U}(z^0, r)$.*
3. *The sequence $\{z^m\}$ generated by (III.4) converges to z^* with*

$$\|z^m - z^*\| \leq \theta^m \|z^0 - z^*\| \tag{III.10}$$

4. *Even more, the sequence $\{\tilde{z}^m\}$ generated by*

$$\tilde{z}^{m+1} = F(\tilde{z}^m), \quad \tilde{z}^0 \in \bar{U}(z^0, r^0), \quad m = 0, 1, 2, \ldots$$

converges to z^* with

$$\|\tilde{z}^m - z^*\| \leq \theta^m \|\tilde{z}^0 - z^*\| \tag{III.11}$$

Claim 4 of the theorem is not very useful if the objective is to solve a particular operator equation. Once it has been established that the equation can be solved by an iteration starting at z^0, there is little interest to know whether it would be solvable if the iteration started at any other point. On the other hand this claim will be shown to be very important in stability analysis of nonlinear systems (Chapter V).

Theorem III.2 has an instructive graphical interpretation, which at the same time serves as a constructive procedure for establishing the contraction conditions. To this end, let z^0 be an arbitrary point, $F(z^0)$ the value of the operator at z^0 and $\bar{U}(z^0, l)$ a ball of radius l centered at z^0. Then define the function of l

$$\theta(l) = \sup_{z \in \bar{U}(z^0, l)} \|F'(z)\| \tag{III.12}$$

By definition $\theta(l)$ is a non-negative, continuous, non-decreasing function of l. It depends on the operator F and on the particular operator norm used in its definition. Fig. III.1 displays possible shapes for $\theta(l)$ for three different operators when the same norm is used. Fig. III.2 shows possible shapes for $\theta(l)$ for the same operator when three different norms are used.

Let l_1 be the ball radius when $\theta(l)$ becomes 1 (if for every l, $\theta(l) > 1$ as in fig. III.1 curve a, set $l_1 = 0$; if $\theta(l) < 1$ for all l as in fig. III.1 curve b, set $l_1 = \infty$). Then:

Case 1 If $l_1 = 0$ a contraction condition cannot be established because θ is always greater than 1; another norm or another point z^0 should be considered.

Case 2 If $l_1 = \infty$, F is a contraction of the space as a whole (global contraction). Theorem III.2 readily applies to show that the equation $z = F(z)$ has a unique solution in the space to which (III.4) converges for every initial point in the space.

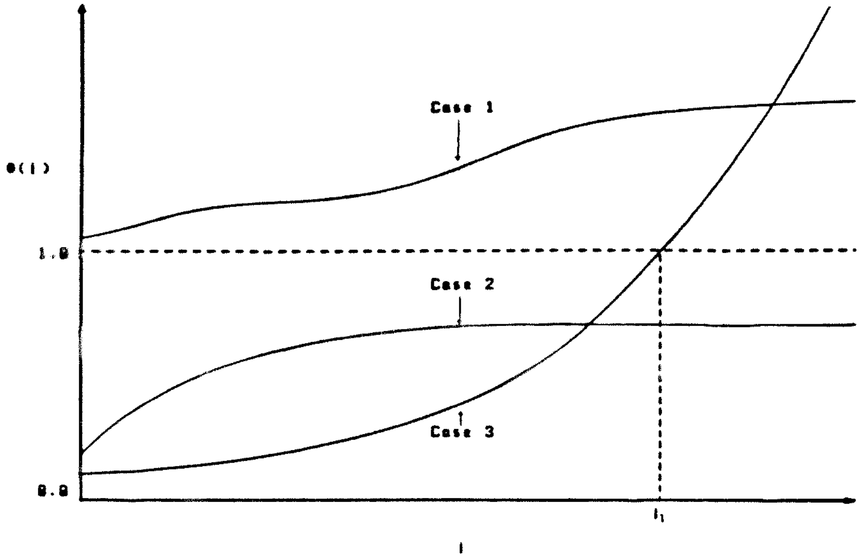

FIGURE III.1 : *Typical $\theta(l)$ shapes for different operators, same norm.*

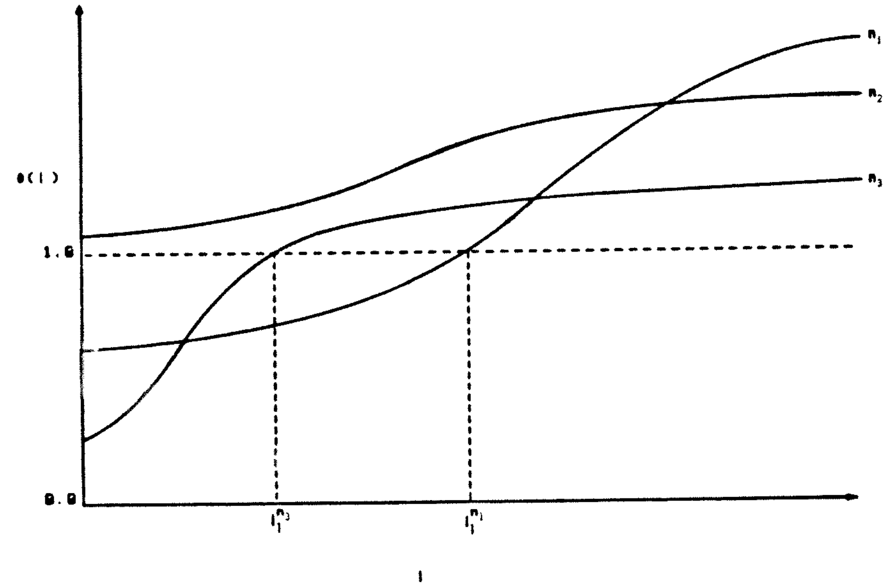

FIGURE III.2 : *Typical $\theta(l)$ shapes for different norms, same operator.*

Case 3 If $0 < l_1 < \infty$ (fig. III.1, curve c), for l in $[0, l_1)$ the function $r^0(l)$ is defined

$$r^0(l) = \|F(z^0) - z^0\|/(1 - \theta(l)) \tag{III.13}$$

By definition $r^0(l)$ is non-negative, continuous, non-decreasing function of l. Typical $r^0(l)$ shapes are shown in figures III.3 and III.4 for different operators and different norms in relation to the curve $r^0 = l$.

Theorem III.2 can be interpreted in terms of the quantities l, $\theta(l)$ and $r^0(l)$. Consider some $\theta^* \in [0, 1)$. To this implicitly corresponds an $l^* \in [0, l_1)$ from (III.12) and explicitly an r^0 from (III.13) . Given θ^*, the conditions of theorem III.2 are satisfied if $\|F'(z)\| < \theta^*, \quad \forall z \in \bar{U}(z^0, r^0)$.

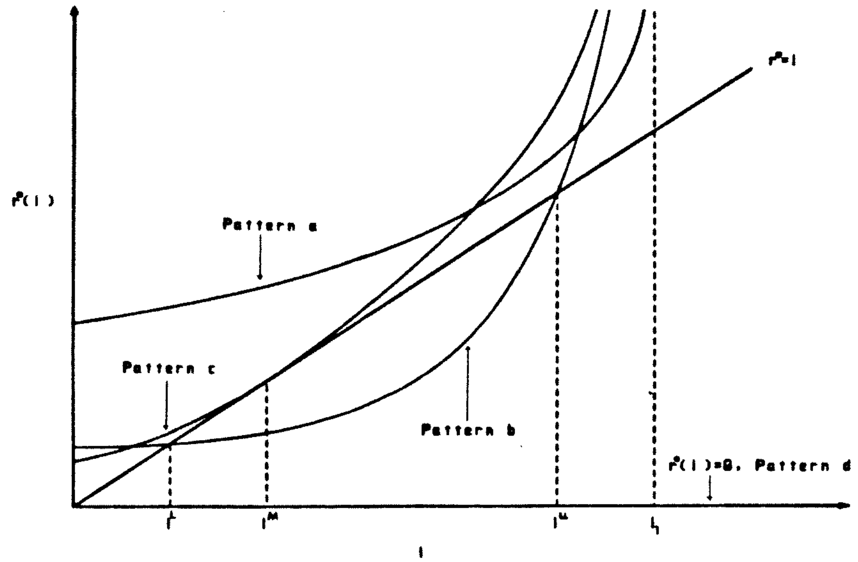

FIGURE III.3 : *Typical $r^0(l)$ shapes for different operators, same norm.*

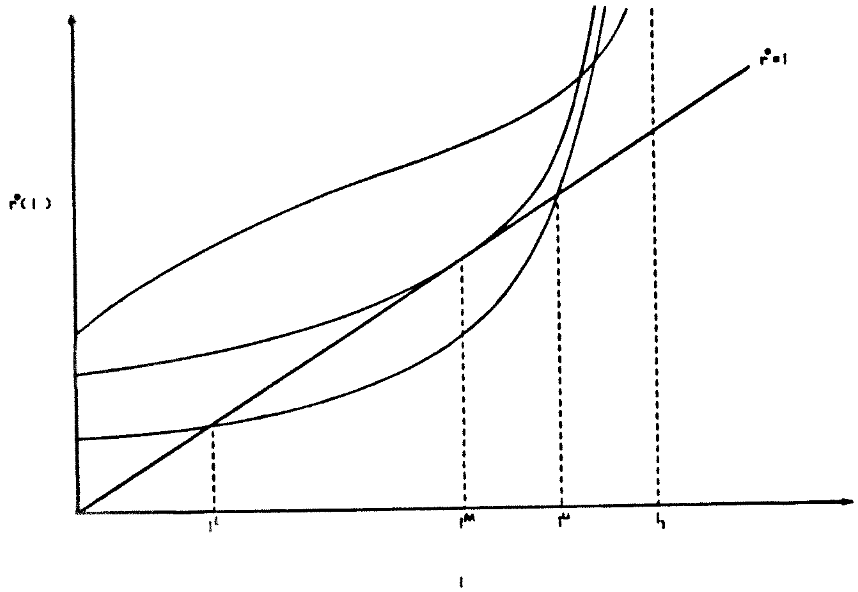

FIGURE III.4 : *Typical $r^0(l)$ shapes for different norms, same operator.*

If $r^0 \leq l^*$, then this is true by (III.12) . If on the other hand $r^0 > l^*$ then there are $z \in \bar{U}(z^0, r^0)$ such that $||F'(z)|| > \theta^*$ and the conditions of Theorem III.2 are violated.

In the present case 3, four basic patterns can occur:

Pattern a: $r^0(l)$ does not intersect line $r^0 = l$ (fig. III.3, curve a). Then the conditions of Theorem III.2 cannot be satisfied. Another norm or initial point should be considered. To understand why, select any l^*, $0 < l^* < l_1$. The corresponding $r^0(l^*)$ is always greater than l^*. It follows from the arguments of the previous paragraph that the contraction conditions cannot be met. This will be true for any l^*, QED.

Pattern b: $r^0(l)$ intersects line $r^0 = l$ at two points, l^L and l^U (fig. III.3, curve b). Then there is a solution in $\bar{U}(z^0, l^L)$, it is unique in $\bar{U}(z^0, l^U)$ and (III.4) converges to it for every initial point $z \in \bar{U}(z^0, l^U)$. To understand why, consider any $l^*, l^L \leq l^* \leq l^U$. The corresponding $r^0(l^*)$ is then smaller or equal ot l^*. It follows that the conditions are satisfied and the claims derive from Theorem III.2.

Pattern c: $r^0(l)$ intersects line $r^0 = l$ at only one point l^M (fig. III.3, curve c). Then there is a solution in $\bar{U}(z^0, l^M)$, which is unique in this ball and (III.4) converges to it for every initial point $z \in \bar{U}(z^0, l^M)$. The same argumentation used in pattern 3.b establishes the claims.

Pattern d: $r^0(0) = 0$ (fig. III.3, curve d). Then z^0 is a solution of $z = F(z)$ and $r^0(l) = 0$ for all l. This case is treated in detail in the following (Theorem III.3) (†).

The issues and application aspects related to Theorem III.2 become significantly simpler when z^0 is the solution of the operator equation (III.1) . This is a preposterous situation from an operator theory point of view: if the solution of an operator equation is known, there is little incentive to develope solution methods and study their properties. Nevertheless it will become evident in later Chapters, that treatment of this situation is of crucial importance when stability of equilibrium states of nonlinear systems is considered.

Theorem III.3 : *Let z^* be a solution of the equation $z = F(z)$. If*

$$||F'(z)|| \leq \theta < 1, \qquad \forall z \in \bar{U}(z^*, r) \tag{III.14}$$

then:

1. *z^* is the unique fixed point of F in $\bar{U}(z^*, r)$.*
2. *The sequence $\{z^m\}$ generated by (III.4) converges to z^* with*

$$||z^m - z^*|| \leq \theta^m ||z^0 - z^*|| \tag{III.15}$$

for every $z^0 \in \bar{U}(z^*, r)$.

Proof: In Appendix III.

Theorem III.3 suggests a particularly simple procedure to characterize a set of points that generate sequences $\{z^m\}$ converging to z^* at least as fast as $\theta_0^m ||z^0 - z^*||$: Define the function $\theta(l)$ as in (III.12) ; then find any l such that $\theta(l) \leq \theta_0$; $\bar{U}(z^*, l)$ is a set of points with the desired property.

Linear operators

For linear finite dimensional operators the contraction conditions can be strengthened further. Related results are summarized and discussed in the following.

Definition III.4 *: Let A be a linear operator on a finite dimensional space. The maximum modulus eigenvalue of A is called the spectral radius of A and is denoted by $\rho(A)$.*

Lemma III.3 *: (Rall 1979) The (Frechet) derivative of a linear operator is the operator itself.*

(†) *Patterns with more than two intersections can occur, but then the problem can always be referred to either pattern b or c*

Lemma III.4 : (Rall 1979) The (Frechet) derivative of an operator at a point z^0 is a linear operator.

Lemma III.5 : (Stakgold 1979) Every linear operator on a finite dimensional space has a matrix representation which is unique with respect to a basis of the space.

Lemma III.6 : (Ortega and Rheinboldt 1970, Desoer and Vidyasagar 1975) Let $\mathcal{N}$ be the set of all induced operator norms on a finite dimensional space. Then, for any linear operator A

$$\inf_{\|\cdot\|_i \in \mathcal{N}} \|A\|_i = \rho(A)$$

In addition, given any $\epsilon > 0$, a norm can be constructed with the property: $\|A\|_i \leq \rho(A) + \epsilon$. This norm is defined by

$$\|A\|_i = \|(PD)^{-1}APD\|_1$$

where P is the similarity transformation of A to its Jordan form, $D = diag(1, \epsilon, \ldots, \epsilon^{n-1})$ and $\|\cdot\|_1$ is the 1-matrix norm:

$$\|A\|_1 = \| [a_{ij}] \| = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|$$

Corollary III.1 : For a linear operator A on a finite dimensional space, the successive substitution sequence

$$z^{m+1} = Az^m + B, \quad m = 0, 1, 2, \ldots$$

1. *When $\rho(A) < 1$, converges to $(I - A)^{-1}B$, the unique solution of $z = Az + B$*
2. *When $\rho(A) > 1$ and $(I - A)$ is invertible, does not converge in general.*
3. *When $\rho(A) = 1$ no conclusions can be drawn.*

Proof: In Appendix III.

Example III.1 : Consider the linear operator with matrix representation

$$A = \begin{pmatrix} 1.470 & -0.367 \\ 0.122 & 0.983 \end{pmatrix}$$

The eigenvalues of A are 1.105 and 1.35, therefore $\rho(A) = 1.35$. The sequence $z^{m+1} = A(z^m)$ does not converge for any non-zero initial point z^0.

Remarks

a. The conditions for convergence of successive substitution algorithms are only sufficient, i.e. the sequence $\{z^m\}$ of Theorem III.1 might converge even though the contraction constant θ is greater than 1. For the special case of linear operators, Corollary III.1 establishes necessary and sufficient conditions for convergence when $\rho(A) \neq 1$.

b. Because the spectral radius is smaller than any induced operator norm, it is a promising candidate for the least conservative convergence criteria, as is the case of Corollary III.1 for linear operators. On the other hand by Lemma III.4 $F'(z)$ is a linear operator, therefore for any induced norm

$$\rho(F'(z)) \leq \|F'(z)\|_i$$

However a condition like

$$\rho(F'(z)) \leq \theta < 1$$

cannot replace (III.8) when F is nonlinear. The reason is that although for any fixed z, an induced norm can be defined that is arbitrarily close to $\rho(F'(z))$ (Lemma III.6), this norm is "tailored" to the specific z and can be significantly larger than the spectral norm at another point in $\bar{U}(z^0, r)$. Tailoring is only useful when F is linear because then $F'(z)$ is independent of z.

c. The spectral radius can be useful as a necessary convergence condition, i.e. when $\rho(F'(z)) > 1$ for some $z \in \bar{U}(z^0, r)$ there is no induced operator norm that satisfies the conditions of Theorem III.1 in $\bar{U}(z^0, r)$.

The method of successive substitution is simple and general. It does not suggest however how to select the operator Q in (III.2) so that the convergence properties can be affected in a desirable manner. This is done in the context of the Newton method.

2. THE NEWTON METHOD

Assuming differentiability and smoothness of the derivatives, the operator P in (III.1) is expanded in its Taylor series around a point z^0:

$$P(z) = P(z^0) + P'(z^0)(z - z^0) + O[\|z - z^0\|^2] \tag{III.16}$$

The standard procedure to derive the Newton method is to assume that z is a solution of $P(z) = 0$ and then truncate the higher order terms. As a result (III.16) becomes a linear operator equation in z

$$0 = P(z^0) + [P'(z^0)](z - z^0)$$

Solving for z yields

$$z = z^0 - [P'(z^0)]^{-1} P(z^0) \tag{III.17}$$

which in turn generates the recursive Newton sequence:

$$z^{m+1} = z^m - [P'(z^m)]^{-1} P(z^m) \tag{III.18}$$

Note that for $Q = -[P'(z^m)]^{-1}$, (III.2) and (III.18) are identical. There are a number of modifications of the Newton method. Two of the most commonly used are defined next:

Definition III.5 : The simplified Newton method is generated when the derivative in (III.18) *is always computed at the same point z_{ref}:*

$$z^{m+1} = z^m - [P'(z_{ref})]^{-1} P(z^m) \tag{III.19}$$

Definition III.6 : The relaxed Newton method is generated when the the z update in (III.18) *is relaxed by some factor λ:*

$$z^{m+1} = z^m - \lambda [P'(z^m)]^{-1} P(z^m) \tag{III.20}$$

where the value of λ depends on some monotonicity criterion (Stoer 1972).

The Kantorović theorem and related results establish the computational theory of the basic Newton method (III.18) :

Theorem III.4 : (Kantorović 1964) Consider an operator P on a Banach space Z, such that P is twice differentiable and the following conditions hold:

i. *There is a $z^0 \in Z$ such that $[P'(z^0)]^{-1}$ exists with*

$$\| [P'(z^0)]^{-1} \| = \beta_0, \quad \| [P'(z^0)]^{-1} P(z^0) \| \leq \eta_0$$

ii. $\|P''(z)\| \leq K$ *in a closed ball* $\bar{U}(z^0, 2\eta_0)$

iii. $h_0 \stackrel{\text{def}}{=} \beta_0 \eta_0 K < \frac{1}{2}$

Then the sequence (III.18) *exists for all $m \geq 0$ and converges to a solution of* (III.1) *which exists and is unique in $\bar{U}(z^0, 2\eta_0)$.*

A few remarks are illustrative of the "mechanics" of the Newton method. Using Lemma III.1

$$\begin{aligned} \|P'(z) - P'(z^0)\| &\leq \sup_{0 \leq \lambda \leq 1} \|P''(\lambda z + (1-\lambda) z^0)\| \, \|z - z^0\| \\ &\leq K \|z - z^0\| \end{aligned}$$

Multiplication by the norm of linear operator $[P'(z^0)]^{-1}$ (which exists and is bounded by condition i of Theorem III.4) and rearranging, yields

$$\| [P'(z^0)]^{-1} \| \, \|P'(z) - P'(z^0)\| \leq \| [P'(z^0)]^{-1} \| K \|z - z^0\|$$

When the conditions of Theorem III.4 are satisfied

$$\| [P'(z^0)]^{-1} \| \, \|P'(z^0) - P'(z)\| \leq \beta K \, 2\eta_0 < 1 \tag{III.21}$$

(III.21) has a straightforward interpretation: the term on the left is an expression for the magnitude of the relative change of the first operator derivative in the ball $\bar{U}(z^0, 2\eta_0)$, i.e. Theorem III.4 asserts that the Newton method will converge if the relative change of the operator derivative in the ball $\bar{U}(z^0, 2\eta_0)$ is less than 100%.

If a solution z^* of the operator equation (III.1) is known, stronger conclusions can be deduced:

Theorem III.5 : Consider the operator P on a Banach space Z. Let z^* be such that $P(z^*) = 0$. Assume P' has a bounded inverse and P'' is bounded in $\bar{U}(z^*, r)$ with $B = \sup_{z \in \bar{U}(z^*,r)} \| [P'(z)]^{-1} \|$ *and* $K = \sup_{z \in \bar{U}(z^*,r)} \|P''(z)\|$. *Consider also the conditions:*

i. $\theta_1 \stackrel{\text{def}}{=} \frac{1}{2} BKr < 1$

ii. $\theta_2 \stackrel{\text{def}}{=} \frac{1}{2} \| [P'(z^*)]^{-1} \| Kr < 1$

Then:

1. *Assuming i., the sequence*

$$z^{m+1} = z^m - [P'(z^m)]^{-1} P(z^m) \tag{III.22}$$

converges to z^* for any $z^0 \in \bar{U}(z^*, r)$ with

$$\|z^{m+1} - z^*\| \le \tfrac{1}{2} BK \|z^m - z^*\|^2$$

2. *Assuming ii., the sequence*

$$z^{m+1} = z^m - [P'(z^*)]^{-1} P(z^m) \tag{III.23}$$

converges to z^* for any $z^0 \in \bar{U}(z^*, r)$ with

$$\|z^{m+1} - z^*\| \le \tfrac{1}{2} BK \|z^m - z^*\|^2$$

3. *If i. and/or ii. hold, z^* is the unique solution of $P(z) = 0$ in $\bar{U}(z^*, r)$.*

Proof: In Appendix III.

The implications of Theorem III.5 is that quadratic convergence can be guaranteed under the assumption of boundedness of the second derivative in some neighbor of the solution. By definition $\theta_2 < \theta_1$ and consequently condition i. is stronger than ii.

Note also that algorithm (III.23) is a simplified Newton method (the derivative is always computed at the same point z^*) with quadratic convergence to the solution. The consequence is that algorithm (III.23) is not only more computationally efficient, but also guarantees a larger radius of convergence than (III.22) .

3. A HYBRID ALGORITHM

Application of operator equation solution methods to dynamic control systems motivates development of hybrid algorithms, where some subset of the variables are under successive substitution, while the rest are updated by another algorithm. Perhaps the best way to describe related algorithms is by a representing example.

Consider the operator

$$\mathcal{X} \times \mathcal{U} \longmapsto \mathcal{X} \times \mathcal{U}$$

$$(x,u) \overset{P}{\longmapsto} P(x,u)$$

the associated operator equation

$$P(x,u) = 0 \tag{III.24}$$

and the equivalent successive substitution form

$$(x,u) = F(x,u) \tag{III.25}$$

with

$$F(x,u) \overset{\text{def}}{=} (x,u) + Q(P(x,u)), \quad Q(\varsigma,\xi) = 0 \quad \Longleftrightarrow (\varsigma,\xi) = 0$$

The following hybrid Newton algorithm (HN1) is defined:

Step 0: Select an initial point (x^0, u^0).

Step 1: Compute a new point using the successive substitution algorithm (III.25) :

$$(x_c^{m+1}, u_c^{m+1}) = F(x^m, u^m) \tag{III.26}$$

Step 2: Compute a second new point using the Newton method

$$(x_N^{m+1}, u_N^{m+1}) = (x^m, u^m) - [P'(x^m, u^m)]^{-1} P(x^m, u^m) \tag{III.27}$$

Step 3: Set $(x^{m+1}, u^{m+1}) = (x_c^{m+1}, u_N^{m+1})$. Go to step 1.

A similar (and simpler) algorithm (HN2) is obtained if the derivative at (III.27) is always computed at the solution (x^*, u^*):

$$(x_N^{m+1}, u_N^{m+1}) = (x^m, u^m) - [P'(x^*, u^*)]^{-1} P(x^m, u^m) \tag{III.28}$$

Theorems III.6 and III.7 establish convergence conditions for the respective hybrid algorithms. The ∞–norm in the product space $\mathcal{X} \times \mathcal{U}$(†) is used.

Theorem III.6 : Consider the operator P and let (x^*, u^*) be such that $P(x^*, u^*) = 0$. Assume P' has a bounded inverse and P'' is bounded in $\bar{U}((x^*, u^*), r)$ with

$$B = \sup_{(x,u)\in\bar{U}((x^*,u^*),r)} \| [P'(x,u)]^{-1} \|$$

$$K = \sup_{(x,u)\in\bar{U}((x^*,u^*),r)} \|P''(x,u)\|$$

If

i. *The operator F of (III.25) is a contraction of $\bar{U}((x^*, u^*), r)$ with contraction constant θ_c*

(†) *The ∞-norm in a product space $\mathcal{X} \times \mathcal{U}$ is defined by $\|(x,u)\|_\infty = \max\{\|x\|, \|u\|\}$, where $\|\cdot\|$ is any vector norm in either space*

ii. $\theta_N \stackrel{\text{def}}{=} \frac{1}{2}BKr < 1$

Then:

1. *Algorithm (HN1) converges to* (x^*, u^*) *for any* $(x^0, u^0) \in \bar{U}((x^*, u^*), r)$, *with*

$$||(x^{m+1}, u^{m+1}) - (x^*, u^*)|| \leq max\{\theta_c, \theta_N\}||(x^m, u^m) - (x^*, u^*)|| \tag{III.29}$$

and, there exists some m^* *such that for* $m > m^*$

$$||(x^{m+1}, u^{m+1}) - (x^*, u^*)|| \leq \theta_c||(x^m, u^m) - (x^*, u^*)|| \tag{III.30}$$

2. (x^*, u^*) *is the unique solution of* $P(x, u) = 0$ *in* $\bar{U}((x^*, u^*), r)$

<u>Proof</u>: In Appendix III.

Theorem III.7 : Consider the operator P *and let* (x^*, u^*) *be such that* $P(x^*, u^*) = 0$. *Assume* P' *has a bounded inverse at* (x^*, u^*) *and* P'' *is bounded in* $\bar{U}((x^*, u^*), r)$ *with*

$$B = ||\,[P'(x^*, u^*)]^{-1}\,||$$
$$K = \max_{(x,u)\in\bar{U}((x^*,u^*),r)} ||P''(x, u)||$$

If

i. *The operator* F *of* (III.25) *is a contraction of* $\bar{U}((x^*, u^*), r)$ *with contraction constant* θ_c

ii. $\theta_N \stackrel{\text{def}}{=} \frac{1}{2}BKr < 1$

Then:

1. *Algorithm (HN2) converges to* (x^*, u^*) *for any* $(x^0, u^0) \in \bar{U}((x^*, u^*), r)$, *with*

$$||(x^{m+1}, u^{m+1}) - (x^*, u^*)|| \leq max\{\theta_c, \theta_N\}||(x^m, u^m) - (x^*, u^*)|| \tag{III.31}$$

and, there exists some m^* *such that for* $m > m^*$

$$||(x^{m+1}, u^{m+1}) - (x^*, u^*)|| \leq \theta_c||(x^m, u^m) - (x^*, u^*)|| \tag{III.32}$$

2. (x^*, u^*) *is the unique solution of* $P(x, u) = 0$ *in* $\bar{U}((x^*, u^*), r)$

<u>Proof</u>: In Appendix III.

Appendix III

Proof of Lemma III.2

a) (If) Let $z^1, z^2 \in \bar{U}(z^0, r)$. Then by Lemma III.1

$$||F(z^1) - F(z^2)|| \le ||F'(\lambda z^1 + (1-\lambda)z^2)||\ ||z^1 - z^2||, \quad \forall\lambda \in [0,1] \tag{III.33}$$

Since $\bar{U}(z^0, r)$ is convex, $z_\lambda \stackrel{\text{def}}{=} \lambda z^1 + (1-\lambda)z^2 \in \bar{U}(z^0, r)$. By assumption then

$$||F'(z_\lambda)|| \le \theta$$

and (III.33) shows

$$||F(z^1) - F(z^2)|| \le \theta||z^1 - z^2|| \tag{III.5}$$

b) (Only if) Suppose there were $z^1, z^2 \in \bar{U}(z^0, r)$ such that $||F(z^1) - F(z^2)|| > \theta||z^1 - z^2||$. By Lemma III.1

$$\theta||z^1 - z^2|| < ||F(z^1) - F(z^2)|| \le ||F'(z_\lambda)||\ ||z^1 - z^2||$$

This implies that

$$||F'(z_\lambda)|| > \theta,\ z_\lambda \in \bar{U}(z^0, r)$$

which contradicts the assumption. ∎

Proof of Theorem III.3

1. Lemma III.2 establishes that F is a contraction of $\bar{U}(z^*, r)$. By Theorem III.1 the result will be established if F maps $\bar{U}(z^*, r)$ into itself, i.e if $z \in \bar{U}(z^*, r)$ then $F(z) \in \bar{U}(z^*, r)$:

$$||F(z) - z^*|| = ||F(z) - F(z^*)|| \stackrel{\text{Lem.1}}{\le} ||F'(\lambda z + (1-\lambda)z^*)||\ ||z - z^*|| \le ||z - z^*|| \le r$$

which shows that $F(z) \in \bar{U}(z^*, r)$. 2.

$$||z^{m+1} - z^*|| = ||F(z^m) - F(z^*)|| \stackrel{\text{Lem.1}}{\le} ||F'(\lambda z^m + (1-\lambda)z^*)||\ ||z^m - z^*|| \le \theta||z^m - z^*||$$

And by induction

$$||z^{m+1} - z^*|| \le \theta^{m+1}||z^0 - z^*||$$ ∎

Proof of Corollary III.1

i. The derivative of the affine linear operator $Az + B$ is the operator A. From Lemma III.5, for any $\epsilon > 0$ there exists an induced norm of A such that

$$\rho(A) < ||A||_i < \rho(A) + \epsilon$$

Choose $\epsilon = (1 - \rho(A))/2$. Then

$$\|A\|_i < \rho(A) + (1 - \rho(A))/2 = (1 + \rho(A))/2 < 1$$

and from Lemma III.3 it follows

$$\|(Az + B)'\|_i = \|A\|_i < 1, \quad \forall z \in Z$$

From Theorem III.3 A is a contraction of the total space and the sequence $\{z^m\}$ converges for all z.

ii. $(I - A)^{-1}$ is invertible by assumption, therefore we can shift the origin by $\bar{z} = (I - A)^{-1}b$: $\varsigma = z - (I - A)^{-1}b$ and the successive substitution algorithm becomes:

$$\varsigma^{m+1} = A(\varsigma^m)$$

Let z_{max} be the eigenvector of A corresponding to the (possibly non-unique) eigenvalue of A of maximum magnitude λ_{max}. Set $\varsigma^0 = z_{max}$. Then $\varsigma^1 = A\varsigma^0 = \lambda_{max} z_{max}$ and iterating $\varsigma^m = \lambda_{max}^m z_{max}$. It follows: $\|\varsigma^m\| = |\lambda_{max}|^m \|z_{max}\|$ and $\lim_{m\to\infty} \|\varsigma^m\| = \infty$, since by assumption $|\lambda_{max}| > 1$. This implies that $\lim_{m\to\infty} \|z^m\| = \lim_{n\to\infty} \|\varsigma^m + (I - A)^{-1}b\| = \infty$. Now, for any ς^0 with a non-zero projection on z_{max}, the ς^0 component along z_{max} will be amplified to ∞. In fact it is trivial (but not very interesting in this context) to extend the result to any ς^0 which has a non-zero projection to the span of the eigenvectors of A corresponding to eigenvalues of magnitude greater than 1.

Proof of Theorem III.5

1) Since by assumption $P(z^*) = 0$ and $P'(z)^{-1}$ exists and is bounded in $\bar{U}(z^*, r)$, it is true that

$$z^* = z^* - [P'(z^m)]^{-1} P(z^*) \tag{III.34}$$

Subtract (III.34) from (III.18) to obtain

$$\begin{aligned}
& z^{m+1} - z^* = z^m - z^* - [P'(z^m)]^{-1}(P(z^m) - P(z^*)) \\
\Longleftrightarrow\; & z^{m+1} - z^m = [P'(z^m)]^{-1}[P'(z^m)(z^m - z^*) - (P(z^m) - P(z^*))] \\
\overset{\text{Lem1}}{\Longrightarrow}\; & \|z^{m+1} - z^m\| \le \|[P'(z^m)]^{-1}\| \tfrac{1}{2} \sup_{0\le\lambda\le 1} \|P''(\lambda z^m + (1-\lambda)z^*)\| \, \|z^m - z^*\|^2 \\
\Longrightarrow\; & \|z^{m+1} - z^m\| \le \tfrac{1}{2} B\, K \, \|z^m - z^*\|^2
\end{aligned} \tag{III.35}$$

Let $\sigma_m \overset{\text{def}}{=} \frac{1}{2} BK\|z^m - z^*\|$. It will be shown in the following that $\sigma_m \le \theta_1^m \sigma_0$.

FirstStep : Show that $\sigma_m \le \theta_1$ by induction.

For $k = 0$:

$$\sigma_0 = \tfrac{1}{2} BK\|z^0 - z^*\| \le \tfrac{1}{2} BKr = \theta_1$$

because $z^0 \in \bar{U}(z^*, r)$.
For $k = m$ assume $\sigma_m \le \theta_1$. Then

$$\begin{aligned}
\sigma_{m+1} &= \tfrac{1}{2}BK\|z^{m+1} - z^*\| \\
&\le \tfrac{1}{2}BK\left[\tfrac{1}{2}BK\|z^m - z^*\|^2\right] && \text{From (III.35)} \\
&= \left[\tfrac{1}{2}BK\|z^m - z^*\|\right]^2 \\
&= (\sigma_m)^2 \\
&\le \theta_1^2 && \text{Induction assumption} \\
&\le \theta_1 && \text{Theorem assumption i.}
\end{aligned}$$

<u>*SecondStep*</u> : Show that $\sigma_{m+1} \le \theta_1 \sigma_m$:

$$\begin{aligned}
\sigma_{m+1} &\le (\sigma_m)^2 && \text{From above} \\
&\le \theta_1 \sigma_m && \text{From above } \sigma_m \le \theta_1
\end{aligned}$$

It follows trivially that $\sigma_m \le \theta_1^m \sigma_0$. Then

$$\lim_{m \to \infty} \sigma_m = \lim_{m \to \infty} \theta_1^m \sigma_0 = 0$$

This implies that

$$\lim_{m \to \infty} \frac{1}{2} BK\|z^m - z^*\| = 0$$

which in turn shows that

$$\lim_{m \to \infty} z^m = z^*$$

2) The proof is identical to the above. The only change is that now (III.34) will be

$$z^* = z^* - [P'(z^*)]^{-1} P(z^*) \tag{III.36}$$

3)Assume that there is another solution z_* in $\bar{U}(z^*, r)$. Then

$$z_* = z_* - [P'(z_*)]^{-1} P(z^*) \tag{III.37}$$

It is also

$$z^* = z^* - [P'(z_*)]^{-1} P(z^*) \tag{III.38}$$

Following the same steps as in obtaining (III.35) :

$$\|z_* - z^*\| \le \tfrac{1}{2}BK\|z_* - z^*\|^2 \le \theta_1\|z_* - z^*\| < \|z_* - z^*\| \tag{III.39}$$

(III.39) is a contradiction.

<u>**Proof of Theorem III.6**</u>

1) Consider the m^{th} step of HN1. Since F is a contraction:

$$\begin{aligned}
&\|(x_c^{m+1}, u_c^{m+1}) - (x^*, u^*)\| \le \theta_c \|(x^m, u^m) - (x^*, u^*)\| \\
\Longleftrightarrow\; &\max\{\|x_c^{m+1} - x^*\|, \|u_c^{m+1} - u^*\|\} \le \theta_c \|(x^m, u^m) - (x^*, u^*)\| \quad \text{Norm definition} \\
\Longrightarrow\; &\|x_c^{m+1} - x^*\| \le \theta_c \|(x^m, u^m) - (x^*, u^*)\| \\
\Longrightarrow\; &\|x^{m+1} - x^*\| \le \theta_c \|(x^m, u^m) - (x^*, u^*)\| \qquad \text{(HN1 step 3)} \qquad \text{(III.40)}
\end{aligned}$$

Since P has the properties required by Theorem III.5, using the notation of the Theorem III.5 proof, it is:

$$\begin{aligned}
&\|(x_N^{m+1}, u_N^{m+1}) - (x^*, u^*)\| \le \sigma_m \|(x^m, u^m) - (x^*, u^*)\| \qquad \text{From (III.35)} \\
\Longleftrightarrow\; &\max\{\|x_N^{m+1} - x^*\|, \|u_N^{m+1} - u^*\|\} \le \sigma_m \|(x^m, u^m) - (x^*, u^*)\| \\
\Longrightarrow\; &\|u_N^{m+1} - u^*\| \le \sigma_m \|(x^m, u^m) - (x^*, u^*)\| \\
\Longrightarrow\; &\|u^{m+1} - u^*\| \le \sigma_m \|(x^m, u^m) - (x^*, u^*)\| \qquad \text{(HN1 step 3)} \qquad \text{(III.41)}
\end{aligned}$$

From (III.40) and (III.41) it follows

$$\max\{\|x^{m+1} - x^*\|, \|u^{m+1} - u^*\|\} \le \max\{\theta_c, \sigma_m\} \|(x^m, u^m) - (x^*, u^*)\| \qquad \text{(III.42)}$$

From the proof of theorem III.5 $\sigma_m \le \theta_N$ and (III.42) shows

$$\|(x^{m+1}, u^{m+1}) - (x^*, u^*)\| \le \max\{\theta_c, \theta_N\} \|(x^m, u^m) - (x^*, u^*)\| \qquad \text{(III.43)}$$

In the proof of Theorem III.5 it is shown that $\sigma_m \to 0$ for $m \to \infty$. This implies that for $\epsilon = \theta_c > 0$, there exists some $m^* > 0$ such that $\sigma_m < \theta_c$ for all $m > m^*$. Then (III.42) implies

$$\|(x^{m+1}, u^{m+1}) - (x^*, u^*)\| \le \theta_c \|(x^m, u^m) - (x^*, u^*)\| \qquad \text{(III.44)}$$

2) The uniqueness of (x^*, u^*) is established by Theorem III.3, because by assumption F is a contraction, therefore it has a unique solution.■

<u>Proof of Theorem III.7</u>

The proof is identical to that given for Theorem III.6.

CHAPTER IV

CONTROL THEORY vs. OPERATOR EQUATION THEORY

The noted duality between control theory and the theory on the solution of operator equations can be sought in the feedback mechanism that underlines both. Consider for example the solution of the operator equation

$$Pu^* = y^* \tag{IV.1}$$

where for given P and y^*, u^* is to be computed.

Trial inputs u^s are injected into the operator P. The outcome of the operation Pu^s, i.e. $y^s = Pu^s$, is compared to y^* and an algorithm A, which in general depends on P, produces a new trial input u^{s+1}

$$u^{s+1} = A(u^s, y^s, y^*; P)$$

with the objective of u^{s+1} being in some sense "closer" to u^* than u^s. A generic block diagram representation of the procedure is shown in fig. IV.1; at the same time it is the basic block structure of feedback control where the algorithm block takes the place of the more familiar controller block.

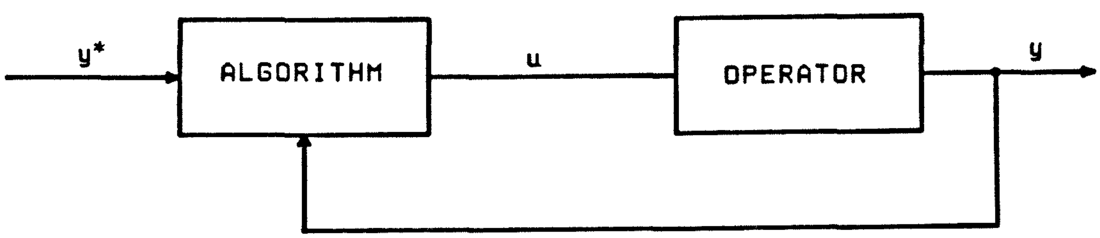

FIGURE IV.1 : *Block diagram representation of an iterative operator equation solution procedure.*

Reversing the argument, a number of important control problems could be formulated as operator equation problems and controller design could be viewed as selection of an algorithm for the solution of an operator equation.

It was emphasized in the introduction that the underlying duality has been explicitly used to address a number of analysis issues. At the same time however, methods for the solution of operator equations are inherently used in well known synthesis approaches.

The case of proportional control is the simplest to examine. The associated block structure appears in fig. IV.2, where the dashed box encloses the equivalent algorithm for the solution of an operator equation:

$$u^{s+1} = k \cdot (y^* - Pu^s) \tag{IV.2}$$

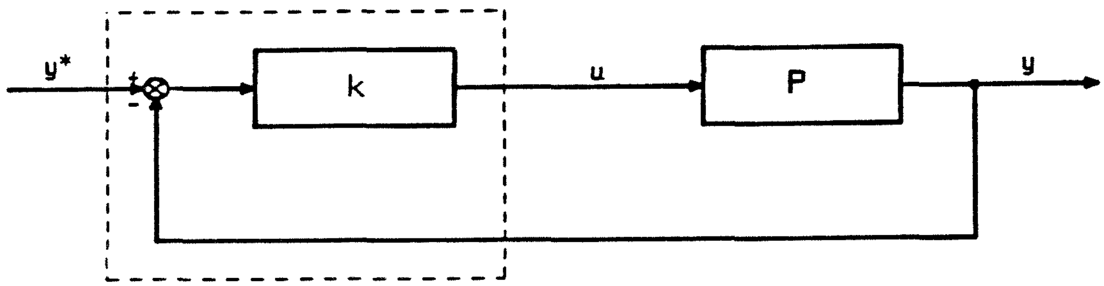

FIGURE IV.2 : *Block diagram of a proportional control structure*

From an operator theory point of view, the method is successive substitution with the sufficient convergence condition

$$|k| \cdot ||P|| < 1$$

i.e. the familiar small gain condition on the loop gain. To uncover the equation that algorithm (IV.2) solves, the limit of (IV.2) is taken as $s \to \infty$. If (IV.2) converges the operator equation is shown to be:

$$u^* - k \cdot (y^* - Pu^*) = 0$$

This is <u>not</u> the desired objective expressed by (IV.1) and the associated offset is not surprising.

Offsets are eliminated by stabilizing controllers with integral action. Fig. IV.3 shows the block structure of an integral action controller, where Q is its non-integral part.
The algorithm is

$$u^{s+1} = u^s + Q(y^* - Pu^s) \tag{IV.3}$$

Again a basic successive substitution method surfaces, this time however the <u>correct</u> equation is solved, as can be easily verified by assuming convergence and taking the limit of (IV.3) on both sides as $s \to \infty$. It is evident that integral action controllers correspond to algorithms for the solution of the <u>correct</u> operator equation.

Next a representing scheme of the family of inverting controllers is examined (Internal Model Control, Garcia and Morari 1982). Fig. IV.4 is the block diagram representation of the basic algorithm, where P is the plant, M is the (linear) model and M^{-1} is the inverse of the model:

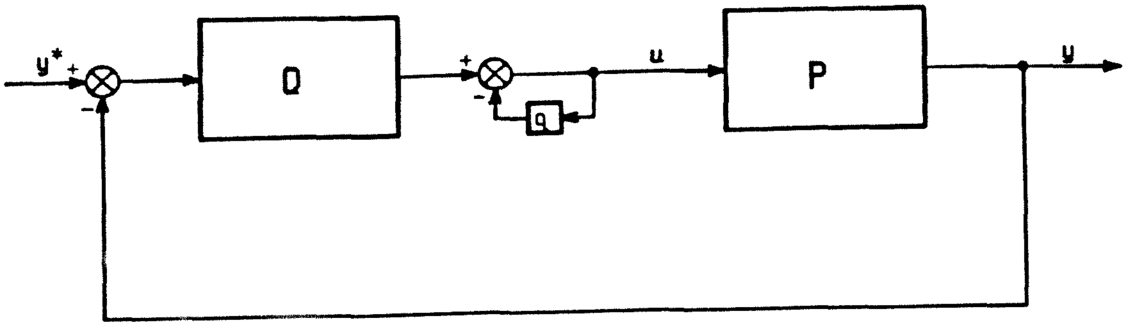

FIGURE IV.3 : *Block diagram of the integral action controller* $C(z) = (z-1)^{-1}Q(z)$ *(q: backward shift - delay - operator).*

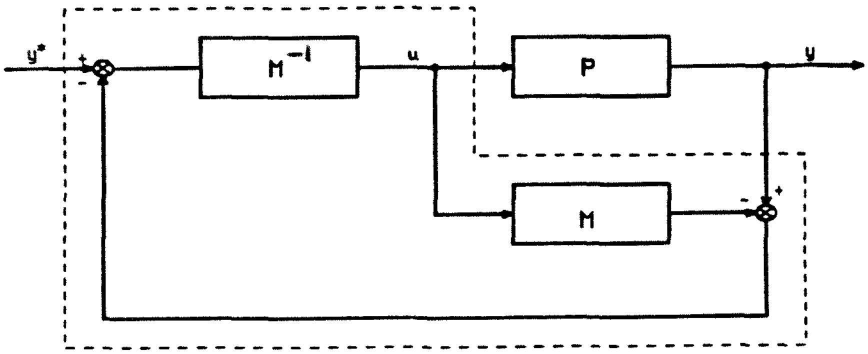

FIGURE IV.4 : *Block diagram of an inverting parametrization controller*

$$\begin{aligned} u^{s+1} &= M^{-1}(y^* - Pu^s + Mu^s) \\ &= u^s - M^{-1}(Pu^s - y^*) \end{aligned} \qquad \text{(IV.4)}$$

Taking into account that the derivative of the linear operator M is itself, (IV.4) can be rewritten:

$$u^{s+1} = u^s - [M']^{-1}(Pu^s - y^*)$$

For $M = P$ it corresponds to the exact Newton method for the solution of $Pu = y^*$, while for $M \neq P$ it corresponds to some simplified Newton method.

To make the point concrete, control problems are posed through various objectives. These include following a specified trajectory, minimizing some performance index, disturbance rejection, dead-beat action, state dead-beat action etc. For a number of meaningful objectives the control problem is equivalent to the solution of an operator equation.

As an example, consider a certain objective, that of following a given trajectory, together with a simple operator interpretation of a dynamic system: under the assumptions 1.i-vi of Chapter II, the system defines an operator, which maps states x^s at the beginning of a sampling interval t^s and inputs u^s constant over that sampling interval, to states $x^{s+1} = \chi(T; x^s, u^s)$ and outputs

$y^{s+1} = g(x^{s+1})$ at t^{s+1}. The objective is to drive the system to a steady state $(x^{s+1} = x^s)$ with its output at $y^{s+1} = y^*$. If the system operator is S

$$\mathcal{X} \times \mathcal{U} \longmapsto \mathcal{X} \times \mathcal{Y} (\dagger)$$

$$(x, u) \overset{S}{\longmapsto} (\chi, y)$$

then the control objective is expressed by the operator equation

$$S(x, u) = (x, y^*) \tag{IV.5}$$

Control law computations to achieve the objective can be based on iterative algorithms for the solution of (IV.5) . Potential gains of this approach stem from a well developed theory on algorithms for the solution of operator equations, especially in areas where control theory has not progressed as much, as is the case of nonlinear systems.

A number of important issues in controller design, such as stability, performance, robustness etc., have their well studied counterparts in the theory of operator equations: convergence, speed of convergence, sensitivity to approximation error etc. Some notable differences do exist however and have to be reconciled :

I. From the definition of the system operator S, at the current state x^s and input u^s an iterative algorithm for the solution of (IV.5) produces a new pair (x^{s+1}, u^{s+1}) for the next iteration. In the operator theory interpretation a new iteration can start at any point in the domain of P. In the control interpretation, although u^{s+1} can be assigned as a system input arbitrarily (assuming no input constraint), x^{s+1} is the state of the system at $t = t^{s+1}$ and, disregarding unrealistic impulses in the states, cannot be arbitrary: it has to equal the system state at $t = t^{s+1}$, as it has evolved during the s^{th} sampling interval i.e.,

$$x^{s+1} = \chi(T; x^s, u^s) \tag{IV.6}$$

Not every algorithm has this property. Basic operator algorithms have to be modified so that (IV.6) holds. The following definition discriminates between algorithms that do have this property and may therefore be used in deriving control laws, from those that cannot be used.

Definition IV.1 : An algorithm for the solution of (IV.5) *is said to be consistent in the derivation of control laws, if the new state x^{s+1} generated by the algorithm is such that* (IV.6) *holds.*

Consistency Requirement
$x^{s+1} = \chi(T; x^s, u^s)$

($\dagger$) *Note that in the discretized problem considered,* $\mathcal{X}$, $\mathcal{U}$ *and* $\mathcal{Y}$ *become* $\mathbf{R}^n$, $\mathbf{R}^m$ *and* $\mathbf{R}^m$ *respectively.*

This requirement on algorithms necessitates use of hybrid algorithms similar to (HN1) and (HN2) of Chapter III.

II. The solutions of operator equations are rarely, if ever, known a priori.

In the framework presented earlier in the chapter, desired steady states of dynamic nonlinear systems are solutions to appropriate operator equations. Most of the time these solutions are known a priori, either by experimentation, or off-line simulation, or the solution of steady-state algebraic equations.

Therefore, the interest is focused in studying the stability properties of known steady states, as well as driving systems to them from perturbed states, rather than using operator equation methods to compute steady states of a system. Theorems like Theorem III.3, which are rather meaningless from an operator theory point of view, become very important as analysis tools in the following Chapter.

III. An algorithm for the solution of a particular operator equation is used only once. On the other hand, if the control problem is based on iterative algorithms, these will be continuously implemented for as long as system operation continues. Therefore extensive study to improve the convergence properties of control algorithms is justified, although it is not always profitable for operator equations.

For example, the convergence properties of algorithm (III.4) are adjusted by the "user-supplied" operator Q. If the algorithm is used to solve the operator equation (III.1), then any Q that happens to generate convergence to the solution for some initial guess is satisfactory in general. There is no need to find a Q that will guarantee convergence for a whole set of initial guesses. It is not crucial also to go to great lengths to adjust Q so as the fastest convergence rate is attained, since the equation will be solved only once. On the other hand, if (III.4) is the control algorithm for the system represented by (III.1), it is very important to search for a Q that will stabilize the system for all possible initial conditions, as well as speed-up settling time.

IV. An important issue in both operator equation theory and feedback control is how convergence properties are affected by errors in the computation of the algorithms.

In most cases, error sources are conceptually similar. For example errors arising from approximation or truncation can be treated in the same fashion as errors arising from measurement error or noise. It is also often that instead of solving the exact equation, an approximate yet simpler equation is solved. In the same sense, instead of basing the control calculations on the exact system description, often a simplified model is used.

However there is a marked difference: the operators involved in operator equations are for all practical purposes exactly known (or can be approximated to any desired degree of accuracy). This is seldom the case in control, where exact system descriptions are rarely available. This motivates the modification of some basic convergence theorems to account for the situation and is discussed in detail in the next Chapters.

CHAPTER V

NONLINEAR SYSTEMS ANALYSIS

In this Chapter analysis tools are developed to characterize the properties of dynamic nonlinear systems, with or without feedback. The criteria stem from convergence analysis of operator equation solution algorithms. Emphasis is placed on the issues of stability and robustness.

The Chapter is divided in five sections. The first section provides a framework for the analysis. Then the state-feedback stability case is studied, which is defined by the assumptions of complete knowledge of the state vector and exact modelling. In the following sections these assumptions are successively dropped: section 3 assumes only exact modelling and details criteria for model reference stability, while in section 4 robustness issues (i.e. stability in the face of modelling error) are investigated. Finally section 5 contains a few remarks on applying the analysis theorems.

The Chapter lays the foundation for the analysis of the control laws that are developed in later Chapters.

1. A FRAMEWORK FOR THE ANALYSIS

There are different notions of stability of nonlinear systems and, before we proceed any further, it is appropriate to define the type of stability relevant in the present context. To do so, a description of the underlying physical framework and some definitions are in order.

A system of the form (II.1)

$$\begin{aligned} \frac{\partial x}{\partial t} &= f(x,u) \\ y &= g(x) \end{aligned} \tag{II.1}$$

sampled at a constant sampling interval T, with its input vector u held constant at a certain level u_f, is the framework for open-loop stability. A number of questions naturally arise pertaining to the dynamic behavior of the system: knowing that at the beginning of the observations the system state vector x is at a level x^0, will the system tend to settle to some state x^* as we observe it at consecutive sampling instants? Is there a state x^* with the property that if the system is at x^* at any observation, it will remain at this state for all subsequent observations? On the other hand, if there is such a state, will the system return to it after some perturbation? And, finally, if this is the case, what is the extent of perturbations that the system at x^* can tolerate?

A set of definitions puts these questions in a quantitative format :

Definition V.1 : (Open-loop System Operator) Under the assumptions of existence and uniqueness of solutions to (II.1), constant input $u = u_f$ and sampling at a constant rate T, system (II.1) gives rise to a well defined nonlinear operator O_f that maps the state at sampling instant s to the state at the next sampling instant $s+1$:

$$\mathbf{R}^n \longmapsto \mathbf{R}^n$$

$$x^s \overset{O_f}{\longmapsto} x^{s+1} = O_f(x^s) \overset{\text{def}}{=} \chi(T; x^s, u_f)$$

where the subscript f underlines the dependence of the system operator on the particular system input u_f.

Example V.1 : In a well stirred tank, the reversible dissociation reaction of a chemical species A to species B is carried out.

$$\mathrm{A} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 2\mathrm{B}$$

The dissociation reaction rate r_1 is

$$r_1 = k_1 C_A \ mol \cdot l^{-1} \cdot s^{-1}, \quad k_1 > 0, \quad C_A \ concentration \ of \ A$$

with the association reaction rate r_2 being

$$r_2 = k_2 C_B^2 \ mol \cdot l^{-1} \cdot s^{-1}, \quad k_2 > 0, \quad C_B \ concentration \ of \ B$$

The tank is insulated, while the heat of the reaction is negligible. As a result the reaction conditions are isothermal and the reaction rate coefficients k_1 and k_2 are constant. The differential mass balance for the reactants is

$$\frac{\partial C_A}{\partial t} = F_A(t) - k_1 C_A + \frac{1}{2} k_2 C_B^2$$
$$\frac{\partial C_B}{\partial t} = F_B(t) + 2k_1 C_A - k_2 C_B^2$$

where $F_A(t)$ and $F_B(t)$ are the respective rates of addition of reactants into the tank. F_A and F_B are the inputs to the system. Assuming $F_A = u_{f1} = 0$, $F_B = u_{f2} = 0$ (no addition of reactants), then multiplying the first equation by 2 and adding to the second shows that

$$\frac{\partial (2C_A + C_B)}{\partial t} = 0 \implies C_0 \overset{\text{def}}{=} 2C_A + C_B = \text{constant}$$

Consequently, letting $x = C_B$ and $C_A = (C_0 - C_B)/2 = (C_0 - x)/2$, the reactor dynamics are described by

$$\frac{\partial x}{\partial t} = 2k_1 \frac{C_0 - x}{2} - k_2 x^2$$
$$= -k_2 x^2 - k_1 x + k_1 C_0$$
$$y_1 = C_B = x$$
$$y_2 = C_A = \frac{C_0 - x}{2}$$

It is shown in Appendix V that the system operator is given by

$$x^{s+1} = O_f(x^s) = \beta\Delta \frac{2x^s + \alpha - \Delta}{2(1-\beta)x^s + (1-\beta)\alpha + (1+\beta)\Delta} + \frac{\Delta - \alpha}{2}$$

where

$$\alpha = \frac{k_1}{k_2}, \ \Delta = \sqrt{\left(\frac{k_1}{k_2}\right)^2 + 4\frac{k_1}{k_2}C_0} \quad \text{and } \beta = \exp(-k_2 T\sqrt{(\frac{k_1}{k_2})^2 + 4\frac{k_1}{k_2}C_0}\ )$$

Definition V.2 : (Equilibrium State) A state vector x_f^* is an equilibrium state of the open loop system of Def. V.1 if it has the property

$$f(x_f^*, u_f) = 0 \qquad (V.1)$$

Def V.2 and (II.1) imply that

$$\frac{\partial x}{\partial t} = 0, \quad \text{for} \quad x(0) = x_f^*$$

therefore x_f^* has the additional property

$$x_f^* = \chi(T; x_f^*, u_f) \qquad (V.1')$$

Example V.2 : For the reaction system of Example V.1, possible equilibrium states are solutions of the steady-state equation

$$-k_2(x_f^*)^2 - k_1 x_f^* + k_1 C_0 = 0$$

This quadratic equation has two solutions, one negative and one positive. The negative solution corresponds to negative concentrations and therefore has no physical significance. The positive solution is

$$x_f^* = \frac{1}{2}\left[-\frac{k_1}{k_2} + \sqrt{\left(\frac{k_1}{k_2}\right)^2 + 4\frac{k_1}{k_2}C_0}\right]$$

and is the unique equilibrium state of the system. It can be verified, using the expressions of example V.1, that

$$x_f^* = O_f(x_f^*) = \beta\Delta \frac{2x_f^* + \alpha - \Delta}{2(1-\beta)x_f^* + (1-\beta)\alpha + (1+\beta)\Delta} + \frac{\Delta - \alpha}{2}$$

Definition V.3 : (Equilibrium State Stability) An equilibrium state x_f^* is stable if there is some real constant $r_f > 0$ and for any $\epsilon > 0$ there exists a sampling instant $s_0 = s_0(\epsilon)$, such that for any initial state in $\bar{U}(x_f^*, r_f)$ the system evolves with

$$||x^s - x_f^*|| < \epsilon, \qquad \forall s > s_0 \tag{V.2}$$

Stability in the sense of Def. V.3 is a local concept, as it basically asserts that arbitrarily small perturbations about a stable equilibrium state result in arbitrarily small perturbations of the trajectory $\{x^s\}$ at large enough times. A concept to characterize stability to finite perturbations is introduced next.

Definition V.4 : (Region of Attraction) A set $\Omega_f \subseteq \mathbf{R}^n$ is a region of (exponential) attraction for x_f^* if every trajectory starting at any initial state $x^0 \in \Omega_f$ converges to x_f^*, with

$$||x^s - x_f^*|| \leq \theta^s ||x^0 - x_f^*||, \quad 0 \leq theta < 1$$

A system can have none, one or many equilibrium states for every u_f. To each stable equilibrium state corresponds at least one non-empty region of attraction. If the region of attraction is $\mathbf{R}^n$ itself, a strong type of stability arises.

Definition V.5 : (Global Stability) An equilibrium state is globally stable, if $\mathbf{R}^n$ is a region of attraction.

The stability concepts introduced thus far pertain to what is usually referred to as uniform asymptotic stability. Uniformity stems from the system time-invariance assumption (equations (II.1) are time-independent). At the same time, since the scope of this work is nonlinear systems control rather than nonlinear system dynamics, asymptotic stability is the stability concept of interest.

Finally, for linear systems the situation is significantly simplified. In general, linear systems have a unique equilibrium state (†) which, if stable, has $\mathbf{R}^n$ as a region of attraction (globally stable).

(†) Linear systems can have whole subspaces of $\mathbf{R}^n$ as equilibrium states when unity is an eigenvalue of the state transition matrix, or when zero is an eigenvalue of the A matrix of the state space representation of the continuous system.

The same concepts of stability carry over to the closed-loop case. Then, the input vector is no longer constant. It changes according to some control law of the form

$$u^{s+1} = \psi(x^s, u^s, y_f^*)(\dagger) \tag{V.3}$$

where y_f^* is any external input (for example a set-point command).

Definition V.6 : (Closed-loop System Operator) The open-loop system of Def. V.1 augmented by control law (V.3) generates a well defined nonlinear operator C_f that maps the state and input at sampling instant s to the state and input at sampling time $s+1$:

$$\mathbf{R}^{n+m} \longmapsto \mathbf{R}^{n+m}$$

$$\begin{pmatrix} x^s \\ u^s \end{pmatrix} \overset{C_f}{\longmapsto} \begin{pmatrix} x^{s+1} \\ u^{s+1} \end{pmatrix} = \begin{pmatrix} \chi(T; x^s, u^s) \\ \psi(x^s, u^s, y_f^*) \end{pmatrix}$$

Vector $\begin{pmatrix} x^s & u^s \end{pmatrix}^T$ is called the (augmented) closed-loop state.

For analysis purposes C_f can be treated in the same fashion O_f is, where now the augmented closed-loop state replaces the open-loop state in the definitions given and the theorems to follow.

2. THE STATE-FEEDBACK CASE

Assumptions : At any sampling instant, the state vector is completely known.
The model of the system is exact

2.1. Open-loop stability

The Contraction Mapping Theorems (CMT) of Chapter III provide a natural framework for addressing all the issues raised in section 1 in a unified manner. Using CMT arguments the following stability theorems are stated and then proved.

Theorem V.1 : Consider the open-loop system of Def. V.1 and an equilibrium state x_f^*. If

$$\left\| \frac{\partial \chi(T; x, u_f)}{\partial x} \right\| \leq \theta < 1, \qquad \forall x \in \bar{U}(x_f^*, r) \tag{V.4}$$

Then :

1. *The equilibrium state x_f^* is unique in $\bar{U}(x_f^*, r)$*

(†) This is a general type of control law. For example, any linear or nonlinear state and/or output control law can be expressed in this form:

$$u^{s+1} = K(x^{s+1}, y_f^* - y^{s+1}) = K(\chi(T; x^s, u^s), y_f^* - g(\chi(T; x^s, u^s))) \stackrel{\text{def}}{=} \psi(x^s, u^s, y_f^*)$$

2. *The equilibrium state x_f^* is stable*

3. *$\bar{U}(x_f^*, r)$ is a region of attraction for x_f^*.*

Proof : In Appendix V.

In Theorem V.1 knowledge of x_f^* is assumed, which is usually the case (through experimentation, simulation or solution of the algebraic steady state equations). This will be also assumed in the remainder of the Chapter. In case x_f^* is not known (as for example when unknown disturbances are present), Theorem V.1 can be modified to establish the existence of an equilibrium state and characterize its stability at the same time :

Theorem V.1′ : Consider the open-loop system of Def. V.1 and an initial state x^0. If

$$\left\| \frac{\partial \chi(T; x, u_f)}{\partial x} \right\| \leq \theta < 1, \qquad \forall x \in \bar{U}(x^0, r) \tag{V.5}$$

where $r \geq r^0 \stackrel{\text{def}}{=} \|\chi(T; x^0, u_f) - x^0\|/(1-\theta)$, then

1. *There is a unique equilibrium state x_f^* in $\bar{U}(x^0, r)$*

2. *x_f^* is stable*

3. *$\bar{U}(x^0, r^0)$ is a region of attraction for x_f^*.*

Proof : In Appendix V.

Theorems V.1 and V.1′ relate to each other the way Theorems III.2 and III.3 do.

Example V.3 : The derivative of the system operator of example V.1 is

$$\frac{\partial O_f}{\partial x} = \frac{4\beta\Delta^2}{[2(1-\beta)x + (1-\beta)\alpha + (1+\beta)\Delta]^2}$$

where by definition

$$\alpha > 0,\ \Delta > 0,\ 0 < \beta < 1$$

The derivative operator is a monotone decreasing function of x. For $x \in [0, \infty)$ it attains its maximum value at $x = 0$

$$\max_{x \in [0,\infty)} \left\| \frac{\partial O_f}{\partial x} \right\| = \left\| \left. \frac{\partial O_f}{\partial x} \right|_{x=0} \right\| = \frac{4\beta\Delta^2}{[(1-\beta)\alpha + (1+\beta)\Delta]^2}$$

It can be easily shown that for $\beta < 1$:

$$4\beta < (1+\beta)^2$$

Then

$$\left\| \frac{\partial O_f}{\partial x} \right\| \leq \frac{4\beta\Delta^2}{[(1-\beta)\alpha + (1+\beta)\Delta]^2}$$
$$< \frac{4\beta\Delta^2}{[(1+\beta)\Delta]^2}$$
$$= \frac{4\beta}{(1+\beta)^2}$$
$$< 1$$

This shows that the conditions of Theorem V.1 are satisfied in any closed ball centered at x_f^* *(see example V.2) of radius* $r \leq x_f^*$. *As a result* x_f^*, *the only equilibrium state of the system, is stable and has* $\bar{U}(x_f^*, r)$, $r \leq x_f^*$ *as a region of convergence. The contraction constant is (using the monotonicity of the derivative operator):*

$$\theta = \max_{x \in \bar{U}(x_f^*, r)} \left\| \frac{\partial O_f}{\partial x} \right\| = \frac{4\beta\Delta^2}{[2(1-\beta)(x_f^* - r) + (1-\beta)\alpha + (1+\beta)\Delta]^2}$$

(The result can be strengthened by noticing that O_f *is a contraction of* $[0, \infty)$ *that maps* $[0, \infty)$ *to itself. Then, using Theorem III.1, convergence to* x_f^* *can be established for any* $x^0 \in [0, \infty)$, *therefore the positive real axis is a region of attraction for* x_f^*.*)*

Example V.4 : *Analytic expressions for the system operator and its derivative are rare in practice and the associated quantities have to be computed numerically. This example examines the open loop stability of the ideal continuous stirred tank reactor of example II.1. To facilitate graphical illustration of the stability conditions, it will be assumed that the feed stream concentration is kept constant, with* $A_i = 1.0$ *and* $R_i = 0$. *This assumption will effectively reduce the reactor to a two state system and consequently will allow for graphical interpretation of the stability condition of Theorem V.1.*

The following two differential equations describe the reactor state evolution, after appropriate values for the coefficients are introduced:

$$x_1 = 1 - x_2$$
$$\frac{\partial x_2}{\partial t} = -x_2 + 3 \times 10^5 \exp(-5000/x_3)x_1 - 6 \times 10^7 \exp(-7000/x_3)x_2$$
$$\frac{\partial x_3}{\partial t} = T_i - x_3 + 0.05[3 \times 10^5 \exp(-5000/x_3)x_1 - 6 \times 10^7 \exp(-7000/x_3)x_2]$$

where t *in min,* x_1, x_2 *in* $mol \cdot l^{-1}$, x_3 *in* $K/100$ *and* T_i, *the inlet temperature, in* $K/100$. *Equilibrium states are obtained by solving the algebraic steady state equations for different values* *of the system input* T_i*:*

$$x_{1eq} = 1 - x_{2eq}$$
$$0 = -x_{2eq} + 3 \times 10^5 \exp(-5000/x_{3eq})x_{1eq} - 6 \times 10^7 \exp(-7000/x_{3eq})x_{2eq}$$
$$0 = T_i - x_{3eq} + 0.05[3 \times 10^5 \exp(-5000/x_{3eq})x_{1eq} - 6 \times 10^7 \exp(-7000/x_{3eq})x_{2eq}]$$

Fig. V.1 is then constructed after elimination of T_i*. It is the temperature - conversion equilibrium diagram for the reactor. A well defined maximum conversion point is at* $x_{1eq} = 0.492$, $x_{2eq} = 0.508$ *and* $x_{3eq} = 4.35$*. At this point* $T_i = 4.33$*.*

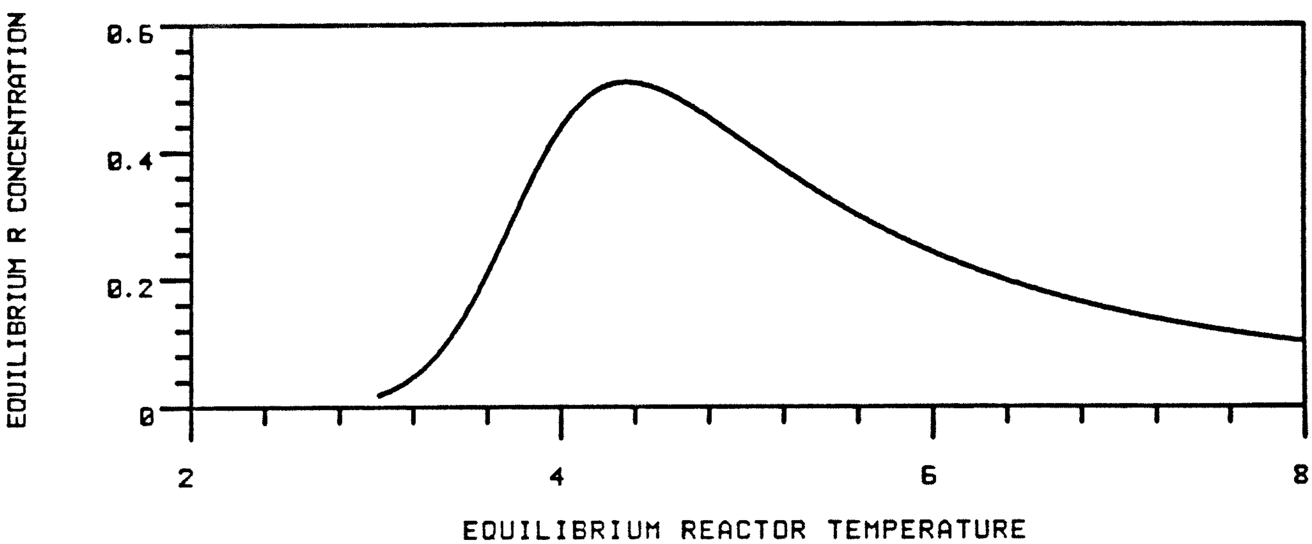

FIGURE V.1 : *Temperature - conversion equilibrium diagram for the Continuous Stirred Tank Reactor.*

For all practical purposes, the operating region of the reactor is confined in the two-dimensional interval

$$0.0 \le x_2 \le 1.0$$
$$2.0 \le x_3 \le 6.0$$

For two different values of the system input T_i*, the 2-norm of the system derivative operator:* $\|\frac{\partial O_f}{\partial x}\|_2 = \|\frac{\partial \chi}{\partial x}\|_2$ *is computed by*

a. selecting a sampling time of 1.0 min,

b. for (x_2, x_3) *in the operating region, solving the initial value problem (II.3) and (II.4)* (†)*. It is then plotted as a function of* x_2 *and* x_3 *in fig. V.2* $(u_f = T_i = 3.5)$ *and V.3* $(u_f = T_i = 4.33)$.

(†) Program DDASAC (Caracotsios and Stewart 1984) was used, that produced the derivative operator by integrating the differential equations over a sampling interval.

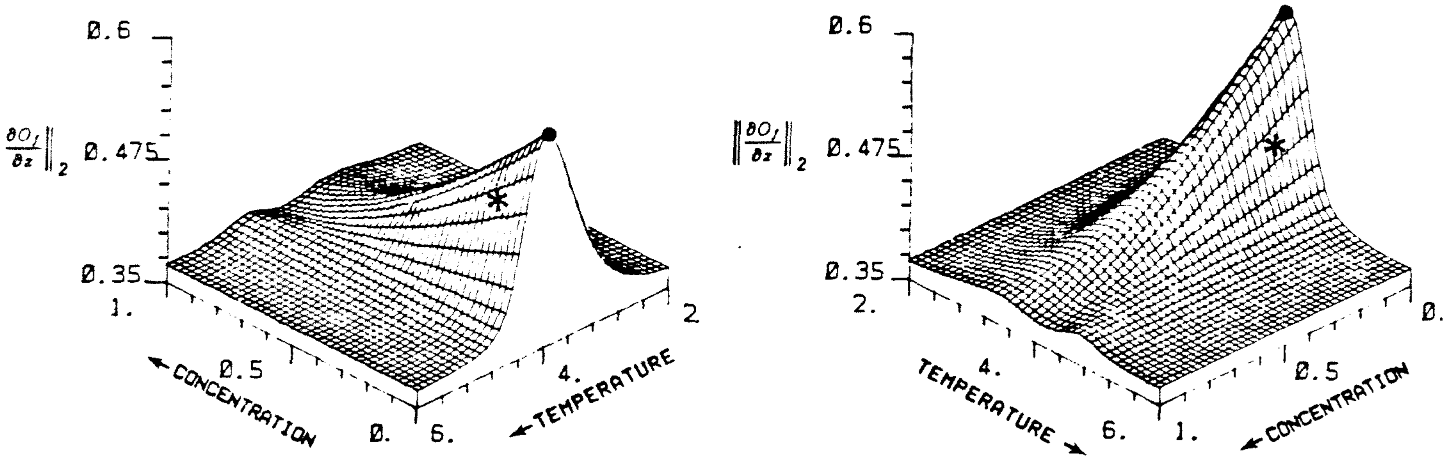

FIGURE V.2 : *2-norm of the derivative operator as a function of x_2 (R concentration) and x_3 (reactor temperature), for $T_i = 3.5$. Two different viewing angles of the surface are shown. The maximum value is 0.57 for $x_2 = 0.0$ and $x_3 = 4.1$ (* : equilibrium state, • : state of maximum norm).*

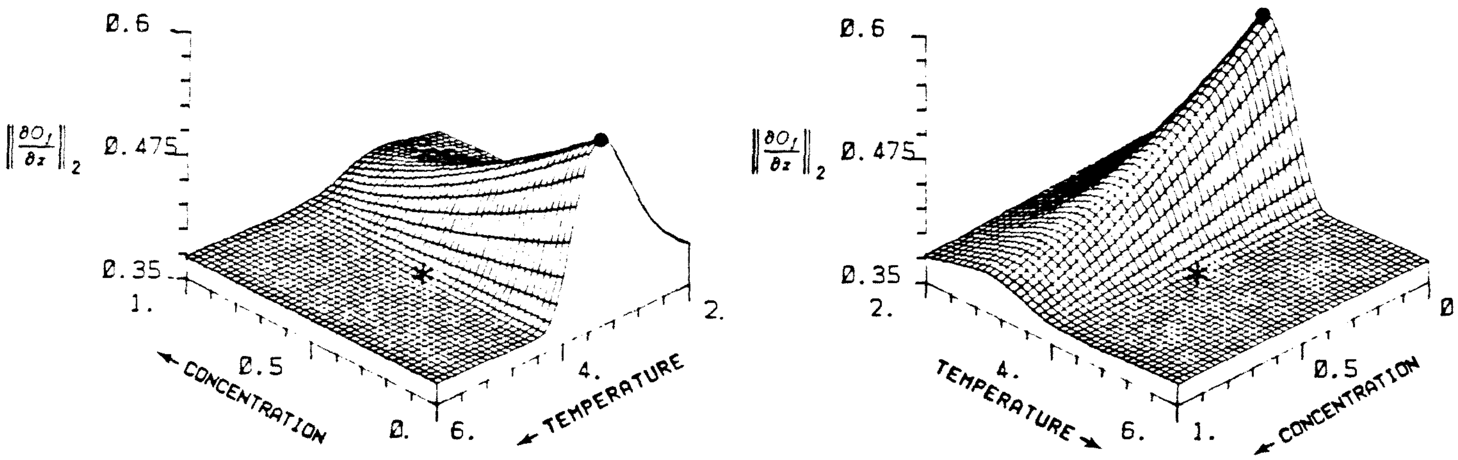

FIGURE V.3 : *2-norm of the derivative operator as a function of x_2 (R concentration) and x_3 (reactor temperature), for $T_i = 4.33$. Two different viewing angles of the surface are shown. The maximum value is 0.54 for $x_2 = 0.0$ and $x_3 = 3.6$ (* : equilibrium state, • : state of maximum norm).*

The following table can be constructed:

$Input, T_i$	x_{2eq}	x_{3eq}	$\max \lVert \frac{\partial O_f}{\partial x} \rVert$
3.50	0.157	3.51	0.57
4.33	0.508	4.33	0.54

Using theorem V.1 we deduce that the respective equilibrium states are stable. Any circle centered at the equilibrium state, lying inside the operating region is a region of attraction. The contraction constants (that characterize the speed with which the reactor returns to equilibrium after any perturbation within the operating region) are 0.57 and 0.54 respectively.

When the contraction condition holds for all $x \in \mathbf{R}^n$ stability can be established in a global sense.

Corollary V.1 : Consider the open-loop system of Def. V.1. If

$$\left\| \frac{\partial \chi(T; x, u_f)}{\partial x} \right\| \leq \theta < 1, \qquad \forall x \in \mathbf{R}^n \tag{V.6}$$

then

1. The system has a unique equilibrium state x_f^*

2. x_f^* *is globally stable*

3. $\mathbf{R}^n$ *is a region of attraction for* x_f^*.

Proof : Follows trivially from Theorem V.1 by setting $r = \infty$.

Theorem V.1. has a familiar interpretation for linear systems.

Corollary V.2 : For a linear system of the form

$$\frac{\partial x}{\partial t} = Ax + Bu_f \tag{V.7}$$

1. $x_f^* = -A^{-1}Bu_f$ *is the unique globally stable equilibrium state if the eigenvalues of the state transition matrix* $\Phi = e^{AT}$ *are inside the unit circle.*

2. If some of the eigenvalues of the state transition matrix are outside the unit circle, the system does not have a stable equilibrium state.

Proof : In Appendix V.

Example V.5 : Consider the linear, continuous system

$$\begin{aligned} \frac{\partial x_1}{\partial t} &= 4x_1 - 3x_2 + u \\ \frac{\partial x_2}{\partial t} &= x_1 \\ y &= x_2 \end{aligned}$$

with a sampling time of $T = 0.1$. *From example II.8, the discrete state space description of the system for the given sampling time is*

$$\begin{aligned} x_1^{s+1} &= 1.470x_1^s - 0.367x_2^s + 0.122u_f \\ x_2^{s+1} &= 0.122x_1^s + 0.983x_2^s + 5.7 \times 10^{-3}u_f \\ y^{s+1} &= x_2^{s+1} \end{aligned}$$

The state transition matrix is then (see also example II.11)

$$\Phi = \begin{pmatrix} 1.470 & -0.367 \\ 0.122 & 0.938 \end{pmatrix}$$

The eigenvalues of Φ are 1.105 and 1.35, i.e. both are outside the unit circle. As a result the closed loop system is (globally) unstable.

2.2 Closed-loop stability

It was mentioned in the introduction that open-loop stability arguments carry over to the closed-loop case by augmenting the open-loop system operator O_f by the feedback control law (V.3) . In this context Theorem V.1 and Corollaries V.1 and V.2 have their closed-loop counterparts which are stated next. The associated proofs are identical and will not be repeated(†).

Theorem V.2 : Consider the closed-loop system of Def. V.6 and an equilibrium state $\begin{pmatrix} x^* & u^* \end{pmatrix}^T$. *If*

$$\left\| \begin{matrix} \dfrac{\partial \chi}{\partial x} & \dfrac{\partial \chi}{\partial u} \\ \dfrac{\partial \psi}{\partial x} & \dfrac{\partial \psi}{\partial u} \end{matrix} \right\| \leq \theta < 1, \qquad \forall \begin{pmatrix} x \\ u \end{pmatrix} \in \bar{U}((x^*, u^*), r) \tag{V.8}$$

Then

1. *The equilibrium state* $\begin{pmatrix} x^* & u^* \end{pmatrix}^T$ *is unique in* $\bar{U}((x^*, u^*), r)$
2. $\begin{pmatrix} x^* & u^* \end{pmatrix}^T$ *is a stable equilibrium state*
3. $\bar{U}((x^*, u^*), r)$ *is a region of attraction for* $\begin{pmatrix} x^* & u^* \end{pmatrix}^T$.

Example V.6 : Consider the closed-loop configuration of an open-loop linear system with a static nonlinear feedback block, shown in fig V.4.

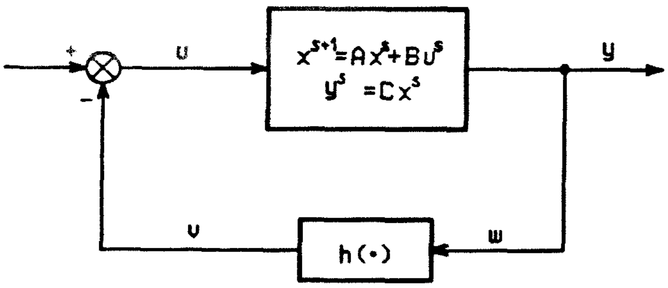

FIGURE V.4 : *Linear system with nonlinear feedback.*

(†) For the sake of brevity the dependence of χ and ψ on x, u, T and y_f^* is suppressed, as well as the subscripts f. It is implied throughout the remainder of the chapter that the assumptions of Def. V.6 hold.

The linear system is described by

$$x^{s+1} = \Phi x^s + \Gamma u^s$$
$$y^s = C x^s$$

The nonlinearity is assumed to be characterized by the nonlinear m*-variable differentiable vector function* $h(\cdot)$

$$v = h(w), \quad h(0) = 0$$

Setting $w^s = y^s = Cx^s$ *and* $u^s = -v^s = -h(w^s)$, *the closed-loop equations are obtained:*

$$x^{s+1} = \Phi x^s + \Gamma u^s$$
$$y^s = C x^s$$
$$u^s = -h(Cx^s)$$

It can be easily verified that $x^s = 0$, $u^s = 0$ *is an equilibrium state. Furthermore, for this particular case,* u^s *is a function of* x^s *alone and as a result the closed-loop stability condition is simplified: substituting* u^s *in the state evolution equation yields*

$$x^{s+1} = \Phi x^s - \Gamma h(Cx^s)$$

Theorem V.1 asserts that if

$$\|\Phi - \Gamma H(x) C\| < 1, \; \forall x \in \mathbf{R}^n, \; H(x) = \left. \frac{\partial h(\varsigma)}{\partial \varsigma} \right|_{\varsigma = Cx}$$

then 0 *is the unique globally stable equilibrium state of the system.*
In the following examples, a number of implications and applications of this stability condition will be investigated.

Example V.7 : *Let* $h(w) = 0, \forall w \in \mathbf{R}^m$, *i.e. the open-loop case. Then* $H(x) = 0$ *and the stability condition of example V.6 becomes* $\|\Phi\| < 1$. *Since* Φ *is a constant matrix, the spectral radius can replace the norm yielding*

$$\rho(\Phi) < 1$$

which is in agreement with Corollary V.2 (stability condition for open-loop linear systems).

Example V.8 : *Let* $h(w) = w$ *(linear unity feedback). Then* $H(x) = I$ *and the stability condition of example V.6 becomes:* $\rho(\Phi - \Gamma C) < 1$, *i.e. the eigenvalues of* $(\Phi - \Gamma C)$ *should be inside the unit circle.*

The standard stability condition for linear discrete systems can be derived in this fashion. The eigenvalues of $(\Phi - \Gamma C)$ are the roots of the closed-loop characteristic polynomial: $det(zI - \Phi + \Gamma C)$. Using Schur's formula for the determinant of block matrices (Kailath 1980), we obtain

$$
\begin{aligned}
det(zI - \Phi + \Gamma C) &= det \begin{pmatrix} I & C \\ -\Gamma & zI - \Phi \end{pmatrix} \\
&= det(zI - \Phi) det[I + C(zI - \Phi)^{-1}\Gamma] \\
&= det(zI - \Phi) det[I + G(z)]
\end{aligned}
$$

where $G(z)$ is the transfer function of the linear block. This shows that the closed-loop system will be stable if the roots of the product of the open-loop characteristic polynomial $(det(zI - \Phi))$ with the determinant of the return difference operator $(I + G(z))$, are inside the unit circle.

Example V.9 : *Consider the case where the linear system of example V.6 is single input-single output. Assume also that the derivative of the nonlinearity lies in some interval $[a, b] \subset \mathbf{R}$:*

$$a \leq \eta \stackrel{\text{def}}{=} \frac{\partial h(w)}{\partial w} \leq b, \quad \forall w \in \mathbf{R}$$

The stability condition of example V.6 becomes

$$||\Phi - \eta \Gamma C|| < 1, \ \forall \eta \in [a, b] \qquad (R1)$$

The result bears some close relationship with the well known describing functions and Kalman conjectures (Munro 1979) and some discussion is in order.

The describing function conjecture is that the closed-loop system will be stable if all the linear systems obtained by replacing the nonlinearity with all possible "instantaneous" gains $\frac{h(w)}{w}$ are stable, i.e. if

$$\rho(\Phi - \xi \Gamma C) < 1, \quad \forall \xi \stackrel{\text{def}}{=} \frac{h(w)}{w}, \ w \in \mathbf{R} \qquad (C1)$$

Two are the main differences between result (R1) and conjecture (C1). First, from basic calculus, the values ξ assumes are a subset of the values that η assumes:

$$a \leq \inf_w \frac{\partial h(w)}{\partial w} \leq \inf_w \frac{h(w)}{w} \leq \xi \leq \sup_w \frac{h(w)}{w} \leq \sup_w \frac{\partial h(w)}{\partial w} \leq b$$

and as a result (R1) imposes a much stronger condition. Second, the spectral radius is used in (C1) instead of the norm in (R1).

The conjecture according to Kalman, proclaims that the closed-loop system will be stable if all the linear systems obtained by replacing the nonlinearity with all possible slopes $\eta = \frac{\partial h(w)}{\partial w}$ it can attain are stable, i.e. if

$$\rho(\Phi - \eta \Gamma C) < 1, \ \forall \eta \in [a, b] \qquad (C2)$$

The only difference between (R1) and (C2) is the use of norm in the first, while the second employs a spectral radius condition. Consequently (R1) is a stronger condition than (C2).
The depicted differences partially explain why both (C1) and (C2) have been shown to be false.

Example V.10 : The special case of a constraint nonlinearity is treated in this example. The nonlinearity is characterized by

$$h(w) = \begin{cases} w, & \text{if } |w| \le 1; \\ 1, & \text{otherwise.} \end{cases}$$

and is shown in fig. V.5.

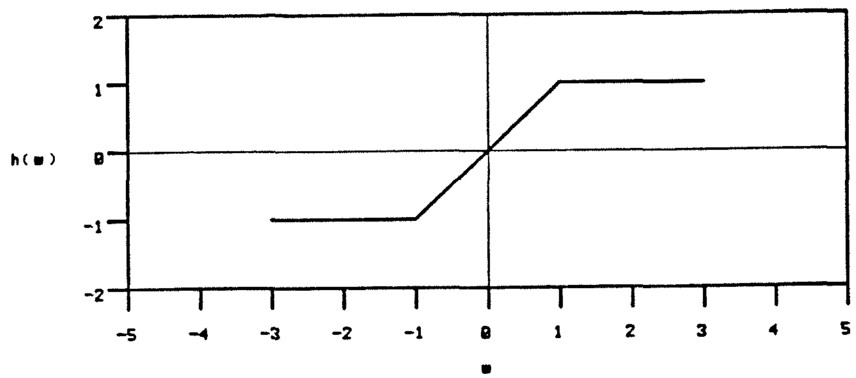

FIGURE V.5 : *A constraint nonlinearity.*

This type of nonlinearity is not differentiable and as a result condition (R1) does not formally apply. However $h(w)$ can be approximated to any desired degree of accuracy by a differentiable function $\bar{h}(w)$ whose derivative ranges from 0 to 1. Then condition (R1) applies to assert stability when

$$\|\Phi - \eta\Gamma C\| < 1, \ \forall \eta \in [0, 1]$$

Corollary V.3 : Consider the closed-loop system of Def. V.6. If

$$\left\| \begin{matrix} \frac{\partial \chi}{\partial x} & \frac{\partial \chi}{\partial u} \\ \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial u} \end{matrix} \right\| \le \theta < 1, \qquad \forall \begin{pmatrix} x \\ u \end{pmatrix} \in \mathbf{R}^{n+m} \tag{V.9}$$

then

1. The system has a unique equilibrium state $\begin{pmatrix} x^* & u^* \end{pmatrix}^T$
2. $\begin{pmatrix} x^* & u^* \end{pmatrix}^T$ is globally stable
3. $\mathbf{R}^{n+m}$ is a region of attraction for $\begin{pmatrix} x^* & u^* \end{pmatrix}^T$.

Corollary V.4 *: Consider the closed-loop system generated by augmenting an open-loop linear system of the form*

$$x^{s+1} = \Phi x^s + \Gamma u^s (\dagger) \tag{V.10}$$

with a linear feedback control law

$$u^{s+1} = \Psi x^s + \Upsilon u^s + \Theta y^* \tag{V.11}$$

Then

1. If the eigenvalues of the matrix $\begin{pmatrix} \Phi & \Gamma \\ \Psi & \Upsilon \end{pmatrix}$ *are inside the unit circle, the system has a unique globally stable equilibrium state*

2. If some eigenvalues of this matrix are outside the unit circle, the system does not have a stable equilibrium state.

> ***Example V.11*** *: Consider the closed-loop system generated by the linear system of example V.5 and control law*
>
> $$u^{s+1} = -33.75x_1^s - 71.3x_2^s - 2.46u^s + 81.7y^*$$
>
> *Then the closed-loop stability matrix becomes*
>
> $$\begin{pmatrix} 1.470 & -0.367 & 0.122 \\ 0.122 & 0.983 & 5.7 \times 10^{-3} \\ -33.75 & -71.3 & -2.46 \end{pmatrix}$$
>
> *which has a spectral radius of approximately* 0.01. *It follows that all eigenvalues are inside the unit circle and the closed loop system is globally stable.*

Corollaries V.2 and 4 have some more information content compared to Theorems V.1 and 2, because they provide both stability and instability conditions. In the following it is shown that instability conditions parallel to the stability conditions of Theorems V.1 and 2 can be derived for nonlinear systems. In order to do so, the relation of the stability conditions developed so far to traditional stability concepts, has to be brought up first.

2.3 Relation to Lyapunov stability

The indirect (first) method of Lyapunov is a powerful tool to characterize local stability of nonlinear systems. The method allows to draw conclusions about a nonlinear system by linearizing the system at some equilibrium state and studying the behavior of the resulting linear system.

Lemma V.1 *: (Vidyasagar 1978) Consider the autonomous system*

$$\frac{\partial x}{\partial t} = f(x) \tag{V.12}$$

($\dagger$) In Chapter II it was shown how this expression is obtained by sampling a continuous linear system of the form : $\dot{x} = Ax + Bu$.

and let

$$A = \left.\frac{\partial f(x)}{\partial x}\right|_{x_f^*} \tag{V.13}$$

with x_f^* an equilibrium state. Then

1. *x_f^* is stable if all eigenvalues of A have negative real parts.*

2. *x_f^* is unstable if at least one eigenvalue of A has a positive real part.*

The following Theorem characterizes the relation of the stability Theorems V.1 and 2 to the indirect Lyapunov method stability analysis.

Theorem V.3 :

1. *If an equilibrium state is stable by Theorems V.1 or 2, then it is stable in the sense of Lyapunov.*
2. *If an equilibrium state is stable in the sense of Lyapunov then it is stable by Theorem V.1 or V.2.*
3. *If an equilibrium state is unstable in the sense of Lyapunov, then the conditions of Theorems V.1 or 2 are not satisfied by any induced operator norm.*
4. *If the norm conditions of Theorem V.1 or V.2 are not satisfied by any induced operator norm in a neighborhood of an equilibrium state, then the equilibrium state is not Lyapunov stable.*

Proof : In Appendix V.

Theorem V.3 asserts that the stability results of the section are in good agreement with established results in nonlinear systems. The advantage of these theorems over the indirect Lyapunov method is that they give criteria for stability to finite perturbations, contrary to the local nature of the indirect Lyapunov method.

Example V.12 : For the system of example V.1

$$A = \left.\frac{\partial\left(-k_2x^2 - k_1x + k_1C_0\right)}{\partial x}\right|_{x_f^*} = -\left(2k_2x_f^* + k_1\right) < 0$$

because every term in the parenthesis is positive. It follows that x_f^* is Lyapunov stable. It was shown in Example V.8 that x_f^* is stable by Theorem V.1 also.

The direct (second) method of Lyapunov gives criteria for global stability of equilibrium states. The following theorem is the discrete version of the basic Krassovskii (1959) theorem which establishes a connection between Lyapunov functions and global stability of discrete systems of the form

$$x^{s+1} = \chi(T; x^s, u_f)$$

It is assumed that the unique equilibrium state of the system has been shifted to the origin.

Lemma V.2 : (Grujić and Šiljak 1973) The equilibrium state of the system is globally stable if and only if there exists a scalar function $V : \mathbf{R}^n \longmapsto \mathbf{R}^1$ (called a "Lyapunov" function) with the properties

i. $V(x)$ *is continuous in* $\mathbf{R}^n$

ii. $\eta_1 \|x\|_2 \leq V(x) \leq \eta_2 \|x\|_2$, (†) $\forall x \in \mathbf{R}^n$

iii. $V(x^{s+1}) \leq \eta_3 V(x^s)$, $\forall s > 0$,

where $\eta_1,\ \eta_2 > 0,\ 0 < \eta_3 < 1$ *are real numbers.*

The following theorem characterizes the relation of the stability Theorem V.1 (and with appropriate modifications of Theorem V.2) to the indirect Lyapunov method stability analysis.

Theorem V.4 : If

$$\|\frac{\partial \chi(T; x, u_f)}{\partial x}\|_{\alpha \mapsto \alpha} \leq \theta < 1 (\ddagger), \quad \forall x \in \mathbf{R}^n$$

and $\chi(T; 0, u_f) = 0$, *then*

1. $x = 0$ *is the unique equilibrium state of system* (V.13)
2. *There exist Lyapunov functions for the system*
3. $\|\chi(T; x, u_f) - x\|_\alpha$ *is a Lyapunov function for the system, where* $\|\cdot\|_\alpha$ *is any vector norm.*

Proof: In Appendix V.

Although Lemma V.2 gives necessary and sufficient conditions for global stability in terms of some Lyapunov function, it does not provide guidelines for constructing such functions. Corollary V.1 on the other hand serves a number of purposes at the same time: it establishes the existence, uniqueness and global stability of an equilibrium state, while as a side product (through Theorem V.4) it establishes the existence and gives explicit formulas for the construction of Lyapunov functions.

2.4 Instability Conditions

Instability conditions, analogous to the stability conditions of Theorems V.1 and 2, can be established at this point. They are summarized in Theorem V.5:

Theorem V.5 : Consider the open (closed) loop system of Def. V.1 (V.6) and a region $\Omega \subseteq \mathbf{R}^n(\mathbf{R}^{n+m})$. *If*

$$\rho(\frac{\partial \chi}{\partial x}) > 1, \qquad \forall x \in \Omega \tag{V.14}$$

$$\left(\text{or,} \quad \rho \begin{pmatrix} \frac{\partial \chi}{\partial x} & \frac{\partial \chi}{\partial u} \\ \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial u} \end{pmatrix} > 1 \quad \forall \begin{pmatrix} x \\ u \end{pmatrix} \in \Omega \quad \right) \tag{V.14'}$$

then Ω *does not contain any stable equilibrium states. Furthermore if* $\Omega = \mathbf{R}^n(\mathbf{R}^{n+m})$ *the system is globally unstable.*

(†) $\|\cdot\|_2$ *denotes the Euclidean vector norm:* $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

(‡) $\|\cdot\|_{\alpha \mapsto \alpha}$ *denotes the operator norm induced by the vector norm* $\|\cdot\|_\alpha$: $\|F\|_{\alpha \mapsto \alpha} = \sup_{\|x\| \neq 0} \frac{\|F(x)\|_\alpha}{\|x\|_\alpha}$

Proof : In Appendix V

Theorem V.5 is somewhat more convenient to apply than its counterparts V.1 and 2, because the instability condition depends only on a unique "measure", the spectral radius, whereas stability conditions depend on a non-uniquely defined norm. But more on that later in section V.5.

3. MODEL REFERENCE STABILITY

In this section the assumption of complete knowledge of the state vector is dropped. The importance of doing so is investigated first and then the state-feedback case conditions are modified to account for the new situation.

Assumption The model of the system is exact

In order to use control laws of the form (V.3) and subsequently apply the results of section 2, the state vector must be accessible at every sampling instant. The exact model of the system is available and, if the state of the system was known at some sampling instant (eg. t=0), the state vector could be inferred at any subsequent time by simulating the model of the system

$$\frac{\partial z}{\partial t} = f(z, u), \qquad at \quad t = 0: \quad z^0 = x^0 \tag{V.15}$$

and observing the model state z. In essence then the situation treated in section 2 would reappear. However the exact system state is rarely available at any time and a value for z^0 has to be assumed for control law computations and stability analysis, which is in general different from the system state : $z^0 \neq x^0$.

It turns out that this is a justifiable practice when the system has a unique globally stable equilibrium state. However, when the system is unstable or has more than one equilibrium states it is likely that the system and the model will display marked differences in behavior, depending on the initial estimate of the state. For example they might settle to different equilibrium states or either might settle while the other becomes unstable.

The depicted situation is very well understood in linear systems theory: simulation of the model to infer the system state corresponds to employing an open-loop observer and is a valid approach for open-loop stable systems. When the system is unstable, a closed-loop observer has to be realized by feeding back information about the discrepancies of model and system outputs to the model equations (V.15) .

The implications to the analysis of systems under state feedback are most important and are discussed next. It is appropriate to clarify the underlying physical situation:

The system evolution over a sampling interval depends explicitly on the state of the system at the beginning of the interval and on the system input. There is no explicit dependence on the model state. The dependence is given by

$$x^{s+1} = \chi(x^s, u^s) \tag{V.16}$$

where the T argument has been suppressed.

The model evolution on the other hand depends on the model state z^s, the input u^s and, assuming feedback of the system output $y^s = g(x^s)$, on the system state, through some relation

$$z^{s+1} = \phi(x^s, z^s, u^s) \tag{V.17}$$

To conform with Chapter VI, a particular expression for the functional relationship $\phi(\cdot,\cdot,\cdot)$ will be adopted :

$$z^{s+1} = \phi(x^s, z^s, u^s) = \chi(z^s, u^s) + Q(g(x^s) - g(z^s)) \tag{V.18}$$

Note that for $x^s = z^s$ it is $\phi(x^s, z^s, u^s) = \chi(z^s, u^s)$ and as a result the model evolution matches the evolution of the system states.

Finally, the control law depends on previous inputs u^s, model states z^s, possibly on system states through output feedback and external inputs y^* :

$$u^{s+1} = \psi(x^s, z^s, u^s, y^*) \tag{V.19}$$

The system operator for this case, as well as the notion of model reference are defined next:

Definition V.7 : *(Model Reference Operator) The open-loop system of Def. V.1 augmented by model (V.18) and control law (V.19) generates a well defined nonlinear operator S that maps system states, model states and inputs at a sampling instant s to states and inputs at sampling time $s+1$:*

$$\mathbf{R}^{2n+m} \longmapsto \mathbf{R}^{2n+m}$$

$$v^s \stackrel{\text{def}}{=} \begin{pmatrix} x^s \\ z^s \\ u^s \end{pmatrix} \stackrel{S}{\longmapsto} \sigma(v^s) \stackrel{\text{def}}{=} \begin{pmatrix} \chi(x^s, u^s) \\ \chi(z^s, u^s) + Q(g(x^s) - g(z^s)) \\ \psi(x^s, z^s, u^s, y^*) \end{pmatrix}$$

In this context, Theorem V.2 can be restated. The proof, being identical in this case, is omitted.

Theorem V.6 : *Consider the system of Def. V.7 and an equilibrium state $v^* = \begin{pmatrix} x^* & z^* & u^* \end{pmatrix}^T$. If*

$$\left\| \begin{matrix} \frac{\partial\chi}{\partial x} & 0 & \frac{\partial\chi}{\partial u} \\ Q\frac{\partial g}{\partial x} & \frac{\partial\chi}{\partial z} - Q\frac{\partial g}{\partial z} & \frac{\partial\chi}{\partial u} \\ \frac{\partial\psi}{\partial x} & \frac{\partial\psi}{\partial z} & \frac{\partial\psi}{\partial u} \end{matrix} \right\| \leq \theta < 1, \qquad \forall v \in \bar{U}(v^*, r) \tag{V.20}$$

Then the equilibrium state v^* is unique in $\bar{U}(v^*, r)$, it is stable and has $\bar{U}(v^*, r)$ as a region of attraction.

Compared to Theorem V.2, Theorem V.6 is more useful and meaningful, because it asserts stability in the face of incomplete knowledge of the system state. In effect the theorem establishes the stability of the real system for all perturbations inside the respective region of attraction, instead of establishing the stability of the model as Theorem V.2 essentially does. Results parallel to Corollaries V.3 and V.4 and Theorem V.5 carry over to the model reference case. Again the proofs are identical and are not included.

Corollary V.5 : Consider the closed-loop system of Def. V.7. If

$$\left\| \begin{matrix} \frac{\partial\chi}{\partial x} & 0 & \frac{\partial\chi}{\partial u} \\ Q\frac{\partial g}{\partial x} & \frac{\partial\chi}{\partial z} - Q\frac{\partial g}{\partial z} & \frac{\partial\chi}{\partial u} \\ \frac{\partial\psi}{\partial x} & \frac{\partial\psi}{\partial z} & \frac{\partial\psi}{\partial u} \end{matrix} \right\| \le \theta < 1, \qquad \forall v \in \mathbf{R}^{2n+m} \tag{V.21}$$

then the system has a unique equilibrium state which is globally stable and $\mathbf{R}^{2n+m}$ *is a region of attraction.*

Corollary V.6 : Consider the closed-loop system generated by augmenting a linear open-loop system of the form (V.10)

$$x^{s+1} = \Phi x^s + \Gamma u^s \tag{V.10}$$

with a linear model

$$z^{s+1} = \Phi z^s + \Gamma u^s + Q(Cx^s - Cz^s) \tag{V.22}$$

and a linear control law

$$u^{s+1} = \Psi z^s + \Upsilon u^s + \Lambda x^s + \Theta y^* \tag{V.23}$$

as well as the constant matrix

$$L = \begin{pmatrix} \Phi & 0 & \Gamma \\ QC & \Phi - QC & \Gamma \\ \Lambda & \Psi & \Upsilon \end{pmatrix} \tag{V.24}$$

1. If the eigenvalues of L are inside the unit circle, the closed-loop system has a unique globally stable equilibrium state
2. If some eigenvalues of L are outside the unit circle, the system has a unique unstable equilibrium state.

Example V.13 : Consider the closed-loop system generated by the linear system of example V.5 augmented by model (V.22) *, with* $Q = \begin{pmatrix} 17.35 & 2.46 \end{pmatrix}^T$, *i.e.*

$$z_1^{s+1} = 1.470z_1^s - 17.7z_2^s + 0.122u^s + 17.35x_2^s$$
$$z_2^{s+1} = 0.122z_1^s - 1.47z_2^s + 5.7 \times 10^{-3}u^s + 2.46x_2^s$$

and the control law

$$u^{s+1} = -906.7x_2^s - 33.75z_1^s + 835.3z_2^s - 2.46u^s + 81.7y^*$$

Then

$$L = \begin{pmatrix} 1.47 & -0.367 & 0 & 0 & 0.122 \\ .122 & 0.983 & 0 & 0 & 5.7\times10^{-3} \\ 0 & 17.35 & 1.47 & -17.7 & 0.122 \\ 0 & 2.46 & 0.122 & -1.47 & 5.7\times10^{-3} \\ 0 & -906.7 & -33.75 & 835.3 & -2.46 \end{pmatrix}$$

The spectral radius of L is approximately 0.01, *therefore all eigenvalues of L are inside the unit circle and the closed-loop system is globally stable.*

Theorem V.7 : Consider the closed-loop system of Def. V.7 and a region $\Omega \subseteq \mathbf{R}^{2n+m}$. If

$$\rho \begin{pmatrix} \frac{\partial\chi}{\partial x} & 0 & \frac{\partial\chi}{\partial u} \\ Q\frac{\partial g}{\partial x} & \frac{\partial\chi}{\partial z} - Q\frac{\partial g}{\partial z} & \frac{\partial\chi}{\partial u} \\ \frac{\partial\psi}{\partial x} & \frac{\partial\psi}{\partial z} & \frac{\partial\psi}{\partial u} \end{pmatrix} > 1, \qquad \forall v \in \Omega \tag{V.25}$$

then Ω does not contain any stable equilibrium states. Furthermore if $\Omega = \mathbf{R}^{2n+m}$ the system is globally unstable.

4. ROBUSTNESS ANALYSIS

In this section the assumption of exact modelling of the system is dropped. The stability conditions of section 3 carry over with minor mathematical, yet, from a practical viewpoint, crucial modifications.

For example the only modification necessary in Theorem V.6 is to change functional relationship (V.18) to reflect the modelling error :

$$z^{s+1} = \varsigma(z^s, u^s) + Q(g(x^s) - h(z^s)) \tag{V.26}$$

Definition V.8 : (Closed-loop Operator with General Modelling Error) The open-loop system of Def. V.1 augmented by model (V.26) and control law (V.19) generates a well defined nonlinear operator G that maps the n system states, l model states and m inputs at a sampling instant s to states and inputs at sampling time $s+1$:

$$\mathbf{R}^{n+l+m} \longmapsto \mathbf{R}^{n+l+m}$$

$$v^s \stackrel{\text{def}}{=} \begin{pmatrix} x^s \\ z^s \\ u^s \end{pmatrix} \stackrel{G}{\longmapsto} \gamma(v^s) \stackrel{\text{def}}{=} \begin{pmatrix} \chi(x^s, u^s) \\ \varsigma(z^s, u^s) + Q(g(x^s) - h(z^s)) \\ \psi(x^s, z^s, u^s, y^*) \end{pmatrix}$$

Theorem V.8 : Consider the system of Def. V.8 and an equilibrium state $v^* = \begin{pmatrix} x^* & z^* & u^* \end{pmatrix}^T$. *If*

$$\left\| \begin{matrix} \frac{\partial \chi}{\partial x} & 0 & \frac{\partial \chi}{\partial u} \\ Q\frac{\partial g}{\partial x} & \frac{\partial \zeta}{\partial z} - Q\frac{\partial h}{\partial z} & \frac{\partial \zeta}{\partial u} \\ \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial z} & \frac{\partial \psi}{\partial u} \end{matrix} \right\| \leq \theta < 1, \qquad \forall v \in \bar{U}(v^*, r) \tag{V.27}$$

Then the equilibrium state v^* *is unique in* $\bar{U}(v^*, r)$, *it is stable and has* $\bar{U}(v^*, r)$ *as a region of attraction.*

Theorem V.8 is of practical use only in the case where, although the system is exactly known, some other model is used for the – possible simplification of – control calculations. If on the contrary, the functional relationship $x^{s+1} = \chi(x^s, u^s)$ is not explicitly available, the usefulness of the theorem is minimal. Treatment of general structural uncertainty in nonlinear systems is a formidable task at present and will not be pursued any further.

However mathematical rigor and – to a certain extent – applicability, can be maintained in the case of parametric uncertainty, that is when the system is known up to a set of parameters.

In this case the following physical situation arises. The system is described by the evolution equation

$$x^{s+1} = \chi(x^s, u^s; p) \tag{V.28}$$

with p a constant – yet unknown – parameter vector, that takes values in an interval $I_p \subset \mathbf{R}^p$. In order to stress the dependence on the parameter vector the notation $\frac{\partial \chi}{\partial x}(p)$ and $\frac{\partial \chi}{\partial x}(p)$ is used for the state and input derivatives respectively, although in general $\frac{\partial \chi}{\partial x}$ and $\frac{\partial \chi}{\partial x}$ also depend on x, u and T. The model is described by the equation

$$z^{s+1} = \chi(z^s, u^s; p_0) + Q(g(x^s; p) - g(z^s; p_0)) \tag{V.29}$$

where p_0 is the nominal value of the parameter vector. Finally the control law is as in (V.19) :

$$u^{s+1} = \psi(x^s, z^s, u^s, y^*) \tag{V.30}$$

It should be noted that the equilibrium state depends on p. To underline the dependence, the notation v_p^* will be employed, $v_{p_0}^*$ denoting an equilibrium state of the nominal system. In this case, the equilibrium state is not explicitly known and consequently a theorem analogous to Theorem V.1′ will be used for stability analysis, with the addition that the contraction condition holds for all $p \in I_p$.

Definition V.9 : (Closed-loop Parametric Uncertainty Operator) The open-loop system (V.28) *augmented by model* (V.29) *and control law* (V.30) *generates a well defined nonlinear operator*

P that maps system states, model states and inputs at a sampling instant s to states and inputs at sampling time $s+1$:

$$\mathbf{R}^{2n+m} \longmapsto \mathbf{R}^{2n+m}$$

$$v^s \stackrel{\text{def}}{=} \begin{pmatrix} x^s \\ z^s \\ u^s \end{pmatrix} \stackrel{\mathrm{P}}{\longmapsto} \pi(v^s) \stackrel{\text{def}}{=} \begin{pmatrix} \chi(x^s, u^s; p) \\ \chi(z^s, u^s; p_0) + Q(g(x^s; p) - g(z^s; p_0)) \\ \psi(x^s, z^s, u^s, y^*) \end{pmatrix}$$

Theorem V.9 *: Consider the system of Def. V.9 and an equilibrium state $v^*_{p_0} = \begin{pmatrix} x^*_{p_0} & z^*_{p_0} & u^*_{p_0} \end{pmatrix}^T$ of the nominal system. If*

$$\left\| \begin{matrix} \frac{\partial \chi}{\partial x}(p) & 0 & \frac{\partial \chi}{\partial u}(p) \\ Q\frac{\partial g}{\partial x}(p) & \frac{\partial \chi}{\partial z}(p_0) - Q\frac{\partial g}{\partial z}(p_0) & \frac{\partial \chi}{\partial u}(p_0) \\ \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial z} & \frac{\partial \psi}{\partial u} \end{matrix} \right\| \leq \theta < 1, \qquad \forall v \in \bar{U}(v^*_{p_0}, r), \forall p \in I_p \tag{V.31}$$

where

$$r \geq r^0 \stackrel{\text{def}}{=} \sup_{p \in I_p} \left\| \begin{matrix} \chi(x^*_{p_0}, u^*_{p_0}; p) - x^*_{p_0} \\ Qg(x^*_{p_0}; p) - Qg(x^*_{p_0}) \\ 0 \end{matrix} \right\|$$

Then

1. *The system has a unique, stable equilibrium state v^*_p in $\bar{U}(v^*_{p_0}, r)$.*
2. *v^*_p has a region of attraction $\bar{U}(v^*_p, r_p)$ of radius $r_p \geq r - r^0$.*

<u>Proof</u>: In Appendix V.

Corollary V.7 *: Consider the closed-loop system of Def. V.9. If*

$$\left\| \begin{matrix} \frac{\partial \chi}{\partial x}(p) & 0 & \frac{\partial \chi}{\partial u}(p) \\ Q\frac{\partial g}{\partial x}(p) & \frac{\partial \chi}{\partial z}(p_0) - Q\frac{\partial g}{\partial z}(p_0) & \frac{\partial \chi}{\partial u}(p_0) \\ \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial z} & \frac{\partial \psi}{\partial u} \end{matrix} \right\| \leq \theta < 1, \qquad \forall v \in \mathbf{R}^{2n+m}, \quad \forall p \in I_p \tag{V.32}$$

then the system has a unique equilibrium state which is globally stable and $\mathbf{R}^{2n+m}$ is a region of attraction.

Corollary V.8 *: Consider the closed-loop system generated by augmenting an open-loop linear system of the form (V.10)*

$$x^{s+1} = \Phi(p)x^s + \Gamma(p)u^s \tag{V.10}$$

with a linear model

$$z^{s+1} = \Phi(p_0)z^s + \Gamma(p_0)u^s + Q(C(p)x^s - C(p_0)z^s) \tag{V.33}$$

and a linear control law

$$u^{s+1} = \Psi z^s + \Upsilon u^s + \Lambda x^s + \Theta y^* \tag{V.34}$$

as well as the matrix

$$L = \begin{pmatrix} \Phi(p) & 0 & \Gamma(p) \\ QC(p) & \Phi(p_0) - QC(p_0) & \Gamma(p_0) \\ \Lambda & \Psi & \Upsilon \end{pmatrix} \tag{V.35}$$

1. If the eigenvalues of L are inside the unit circle for all $p \in I_p$, the closed-loop system has a unique globally stable equilibrium state

2. If some eigenvalues of L are outside the unit circle for all $p \in I_p$, the system has a unique unstable equilibrium state.

Theorem V.10 : Consider the closed-loop system of Def. V.9 and a region $\Omega \subseteq \mathbf{R}^{2n+m}$. If

$$\rho \begin{pmatrix} \frac{\partial \chi}{\partial x}(p) & 0 & \frac{\partial \chi}{\partial u}(p) \\ Q\frac{\partial g}{\partial x}(p) & \frac{\partial \chi}{\partial z}(p_0) - Q\frac{\partial g}{\partial z}(p_0) & \frac{\partial \chi}{\partial u}(p_0) \\ \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial z} & \frac{\partial \psi}{\partial u} \end{pmatrix} > 1, \qquad \forall v \in \Omega \quad \forall p \in I_p \tag{V.36}$$

then Ω does not contain any stable equilibrium states. Furthermore if $\Omega = \mathbf{R}^{2n+m}$ the system is globally unstable.

5. CONCLUDING REMARKS

5.1 Disturbances

Up to this point no explicit mention to disturbances was given. Stability was studied in the face of perturbations from equilibrium states. However disturbances frequently appear in control systems and some additional discussion is in order. Both constant and time-varying disturbances will be treated. Only the open-loop case will be covered, extension to the other cases being straightforward. The notation

$$x^{s+1} = \chi(x^s, u_f; d^s) \tag{V.37}$$

is introduced to denote the effect of disturbances to the system evolution, where d^s is the disturbance vector at the s^{th} sampling interval. In this context

$$x^{s+1} = \chi(x^s, u_f; 0) \overset{\text{def}}{=} \chi(x^s, u_f) \tag{V.38}$$

will denote the nominal (undisturbed) system evolution.

5.1.1 Constant Disturbances :

5.1.1.1 Additive constant disturbances: The disturbance set is characterized by

$$d^s = d \in I_d \subset \mathbf{R}^n$$

and the system evolution equation (V.37) becomes

$$x^{s+1} = \chi(x^s, u_f) + d \tag{V.39}$$

The system equilibrium state depends on d and the notation x_d^* is used to underline this dependence:

$$x_d^* = \chi(x_d^*, u_f) + d \tag{V.40}$$

x_d^* is unknown and consequently Theorem V.1 cannot be used. The following theorem characterizes the stability properties of the disturbed system. Note that the contraction conditions do not depend on d. The only difference from Theorem V.1 is that there is a minimum radius for the contraction ball.

Theorem V.11 *: Consider the open-loop system described by*

$$x^{s+1} = \chi(x^s, u_f) + d \tag{V.39}$$

where $d \in I_d \subset \mathbf{R}^n$. Let δ be the maximum possible magnitude of d:

$$\delta = \sup_{d \in I_p} ||d|| \tag{V.41}$$

and let x_f^* be an equilibrium state of the undisturbed system (V.38) . If

$$||\frac{\partial \chi(x, u_f)}{\partial x}|| \leq \theta < 1, \quad \forall x \in \bar{U}(x_f^*, r) \tag{V.42}$$

with

$$r \geq R \stackrel{\text{def}}{=} \frac{\delta}{1-\theta} \tag{V.43}$$

Then

1. *There is an equilibrium state x_d^* of the disturbed system (V.39) in $\bar{U}(x_f^*, R)$.*
2. *x_d^* is unique in $\bar{U}(x_f^*, r)$.*
3. *x_d^* is stable.*
4. *$\bar{U}(x_f^*, r)$ is a region of attraction for x_d^*.*

<u>Proof</u>: In Appendix V.

5.1.1.2. Constant non-additive disturbances : Constant disturbances act the same way uncertain parameters do (i.e. modelling uncertainty), therefore the analysis of section V.4 (parametric uncertainty) directly applies.

<u>5.1.2.</u> Time varying disturbances :

5.1.2.1. Bounded disturbances: The set of disturbances is characterized by

$$||d^s|| \leq \delta, \quad d^s \in I_d \subset \mathbf{R}^d \tag{V.44}$$

The following theorem gives conditions for bounded system response to bounded time-varying disturbances.

Theorem V.12 : Consider the open-loop system described by

$$x^{s+1} = \chi(x^s, u_f; d^s) \tag{V.38}$$

where d^s is bounded:

$$\|d^s\| \le \delta, \quad d^s \in I_d \subset \mathbf{R}^d \tag{V.44}$$

Let x_f^* be an equilibrium solution of the undisturbed system

$$x_f^* = \chi(x_f^*, u_f; 0) \tag{V.45}$$

and define the quantity B by

$$B \stackrel{\text{def}}{=} \sup_{\substack{\|d\| \le \delta \\ x \in \bar{U}(x_f^*, r_d)}} \left\| \frac{\partial \chi(x, u_f; d)}{\partial d} \right\|$$

If the <u>undisturbed</u> *system is a contraction*

$$\left\| \frac{\partial \chi(x, u_f)}{\partial x} \right\| \le \theta < 1, \quad \forall x \in \bar{U}(x_f^*, r) \tag{V.46}$$

with r such that $R \stackrel{\text{def}}{=} B\delta/(1-\theta) \le r \le r_d$, then

1. Every trajectory that starts in a ball $\bar{U}(x_f^*, r_0)$, $r_0 \le R$, remains bounded in $\bar{U}(x_f^*, R)$.

2. Every trajectory that starts in a ball $\bar{U}(x_f^*, r_0)$, $R < r_0 \le r$, remains bounded in $\bar{U}(x_f^*, r_0)$. Furthermore, for any $\epsilon > 0$, there exists $s_0 = s_0(\epsilon)$ such that for $s > s_0$ the trajectory is bounded in $\bar{U}(x_f^*, R + \epsilon)$.

<u>Proof</u>: In Appendix V.

5.1.2.2. Asymptotically constant disturbances: The set of disturbances is characterized by

$$\|d^s\| \le \lambda^s \delta, \quad 0 \le \lambda < 1 \tag{V.47}$$

i.e. they tend to zero asymptotically. The following Theorem gives stability conditions for this case.

Theorem V.13 : Consider the open-loop system described by

$$x^{s+1} = \chi(x^s, u_f; d^s) \tag{V.39}$$

with

$$\|d^s\| \le \lambda^s \delta, \quad d^s \in I_d \subset \mathbf{R}^d \tag{V.47}$$

Let x_f^* be an equilibrium solution of the undisturbed system

$$x_f^* = \chi(x_f^*, u_f; 0) \tag{V.39}$$

and define the quantity B by

$$B \stackrel{\text{def}}{=} \sup_{\substack{\|d\| \in I_d \\ x \in \bar{U}(x_f^*, r_d)}} \left\| \frac{\partial \chi(x, u_f; d)}{\partial d} \right\|$$

If the <u>undisturbed</u> *system is a contraction*

$$\left\| \frac{\partial \chi(x, u_f)}{\partial x} \right\| \leq \theta < 1, \quad \forall x \in \bar{U}(x_f^*, r) \tag{V.48}$$

with r such that $R \stackrel{\text{def}}{=} B\delta/(1-\theta) \leq r \leq r_d$, *then*

1. x_f^* *is the unique equilibrium state of* (V.38) .
2. *Every trajectory that starts at* $x^0 \in \bar{U}(x_f^*, r)$ *converges to* x_f^* *with*

$$\|x^s - x_f^*\| \leq \theta^s \|x^0 - x_f^*\| + sB\delta \cdot max\{\lambda, \theta\}^{s-1} \tag{V.49}$$

<u>Proof</u>: In Appendix V.

<u>5.2. Selecting a Norm</u>

There is a marked difference between the stability and instability conditions of the previous sections: in both cases some measure of the derivative of the system operator is compared against 1; instability is conditional on a unique measure, the spectral radius of the derivative operator, whereas any norm out of an infinite multitude of induced norms can be used for stability analysis. Some of these norms produce more conservative results than others and the question of which norm is more appropriate to use, naturally arises.

It is instructive to review the stability analysis theorems for a moment. System stability was established on the condition that the derivative – which will be called $L(v)$ here to generalize notation – of the system operator is bounded from above by 1 for every state v in some region. This L had different forms depending on the context. For example in the case of Theorem V.1 it was called $\frac{\partial \chi}{\partial x}$ and was a function of the open-loop system state x; in Theorem V.2 it was $\begin{pmatrix} \frac{\partial \chi}{\partial x} & \frac{\partial \chi}{\partial u} \\ \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial u} \end{pmatrix}$ a function of x and u etc.

For any particular v, $L(v)$ is a linear operator (Lemma III.4). By Lemma III.6 every induced norm of $L(v)$ is greater than or equal to its spectral radius. Furthermore, a norm can be found that approaches arbitrarily close to the spectral radius. It deems promising then to compute $\rho(L(v))$ and then select a norm that approximates it for the subsequent analysis. It turns out that this may not be a good practice, because the particular norm employed is tailored to v and may be considerably larger (and therefore conservative) than the spectral radius $\rho(L(v'))$ at another state v' in the region of interest. However it is a justifiable practice when either the system operator is smoothly nonlinear (note that by Corollary V.2 the spectral norm is an exact characterization for linear systems), or when only a small region, over which $L(v)$ does not change significantly, is of interest.

5.3. Implementation aspects

It is evident that stability analysis of nonlinear systems involves a sizable amount of computation. It seems this is a price to be paid in a quest for a general treatment of nonlinear systems.

No systematic way to carry out the computations is available at present. It poses as a future research topic. At the current state of this theory, a 'brute-force' approach is adopted to implement the theoretical findings. It is described in the following for the case of open-loop stability (Theorem V.1). The same procedure with appropriate modifications is used to establish the conditions of the other Theorems.

Step 1: Fix the external input $(u = u_f)$. Compute an equilibrium state of interest (x_f^*) by solving equation (V.1) and selecting one of its solutions.

Step 2: Solve the initial value problem (II.3-4) at x_f^* to compute $\Phi(x_f^*) = \left.\frac{\partial \chi}{\partial x}\right|_{x_f^*}$.

i) If $\rho(\Phi(x_f^*)) > 1$, x_f^* is unstable and stability analysis concludes.
ii) If $\rho(\Phi(x_f^*)) = 1$, no conclusions can be drawn, analysis terminates.
iii) If $\rho(\Phi(x_f^*)) < 1$, continue to establish region of attraction.

Step 3: Select an induced norm.

i) If $\|\Phi(x_f^*)\| \geq 1$ the norm is not appropriate. Go to Step 3.
ii) If $\|\Phi(x_f^*)\| < 1$, select a sufficiently small radius r, continue.

Step 4: Divide the surface S of the ball $\bar{U}(x_f^*, r)$ into a finite grid of points (states). At each grid point solve problem (II.3-4).

Step 5: Compute $\theta = \max_{x \in S} \|\Phi(x)\|$.

i) If $\theta < 1$ increase r, go to Step 4.
ii) If $\theta \geq 1$ reduce r to its previous value, continue.

Step 6: x_f^* is stable, with $\bar{U}(x_f^*, r)$ a region of attraction.

i) If the region of attraction is satisfactory, analysis concludes.
ii) If not, alternative norms should be considered, go to Step 3.

Appendix V

Facts related to the reaction system of example V.1

1. Integration to obtain $x^{s+1} = O_f(x^s)$:

$$\frac{\partial x}{\partial t} = -k_2 x^2 - k_1 x + k_1 C_0 = -k_2\left(x^2 + \frac{k_1}{k_2}x - \frac{k_1}{k_2}C_0\right)$$

$$= -k_2(x^2 + \alpha x - \alpha C_0) \qquad \text{where } \alpha = \tfrac{k_1}{k_2}$$

$$\Longleftrightarrow \frac{\partial x}{x^2 + \alpha x - \alpha C_0} = -k_2 \partial t$$

$$\Longleftrightarrow \int_{x^s}^{x^{s+1}} \frac{\partial x}{x^2 + \alpha x - \alpha C_0} = -\int_{t^s}^{t^{s+1}} k_2 \partial t \qquad \text{at } t = t^s\ x = x^s$$

$$\Longleftrightarrow \frac{1}{\Delta}\log\left[\frac{2x^{s+1} + \alpha - \sqrt{(\frac{k_1}{k_2})^2 + 4\frac{k_1}{k_2}C_0}}{2x^{s+1} + \alpha + \sqrt{(\frac{k_1}{k_2})^2 + 4\frac{k_1}{k_2}C_0}} \, \frac{2x^s + \alpha + \sqrt{(\frac{k_1}{k_2})^2 + 4\frac{k_1}{k_2}C_0}}{2x^s + \alpha - \sqrt{(\frac{k_1}{k_2})^2 + 4\frac{k_1}{k_2}C_0}}\right] = -k_2 T$$

$$\Longleftrightarrow \frac{2x^{s+1} + \alpha - \Delta}{2x^{s+1} + \alpha + \Delta}\,\frac{2x^s + \alpha + \Delta}{2x^s + \alpha - \Delta} = \exp(-k_2 T\Delta) \qquad \text{where } \Delta = \sqrt{(\tfrac{k_1}{k_2})^2 + 4\tfrac{k_1}{k_2}C_0}$$

$$\Longleftrightarrow x^{s+1} = \beta\Delta \frac{2x^s + \alpha - \Delta}{2(1-\beta)x^s + (1-\beta)\alpha + (1+\beta)\Delta} + \frac{\Delta - \alpha}{2} \quad \text{where } \beta = \exp(-k_2 T\Delta) < 1$$

2. Differentiation:

$$\frac{\partial O_f(x^s)}{\partial x^s} = \frac{\partial x^{s+1}}{\partial x^s}$$

$$= \beta\Delta \frac{[4(1-\beta)x^s + 2(1-\beta)\alpha + 2(1+\beta)\Delta] - [4(1-\beta)x^s + 2(1-\beta)\alpha - 2(1+\beta)\Delta]}{[2(1-\beta)x + (1-\beta)\alpha + (1+\beta)\Delta]^2}$$

$$= \frac{4\beta\Delta^2}{[2(1-\beta)x + (1-\beta)\alpha + (1+\beta)\Delta]^2}$$

Proof of Theorem V.1

The proof is based on the fact that the properties of the open-loop system are characterized by the properties of the operator O_f of Def. V.1.

1. By (V.1′) an equilibrium state x_f^* is a fixed point of the operator O_f. By assumption

$$\left\|\frac{\partial O_f}{\partial x}\right\| \le \theta < 1, \quad \forall x \in \bar{U}(x_f^*, r)$$

Theorem III.3 then asserts that x_f^* is unique.

2. Let r_f in Def. V.3 be equal to r. By the assumptions of this theorem, Theorem III.3 implies that if $x^{s+1} = O_f(x^s)$ then

$$\|x^s - x_f^*\| \le \theta^s \|x^0 - x_f^*\|, \quad \forall x^0 \in \bar{U}(x_f^*, r_f)$$

Choose any $\epsilon > 0$ and let $s_0 > [ln(\epsilon/||x^0 - x_f^*||)/ln(\theta)]$. For $s > s_0$:

$$||x^s - x_f^*|| \leq \theta^s ||x^0 - x_f^*|| < \theta^{s_0} ||x^0 - x_f^*|| < \epsilon \quad \forall x^0 \in \bar{U}(x_f^*, r_f)$$

By Def. V.3, there exists $r_f = r > 0$ such that for any $\epsilon > 0$ there exists $s_0 = s_0(\epsilon)$ such that for every trajectory that starts in $\bar{U}(x_f^*, r_f)$, $||x^s - x_f^*|| < \epsilon \;\; \forall s > s_0$.

3. Part 2 of the proof established that $\lim_{s \to \infty} ||x^s - x_f^*|| = 0$, i.e.

$$\lim_{s \to \infty} x^s = x_f^*, \quad \forall x^0 \in \bar{U}(x_f^*, r)$$

i.e. every trajectory starting at an initial state $x \in \bar{U}(x_f^*, r)$ converges to x_f^*. By Def. V.4 it follows that $\bar{U}(x_f^*, r)$ is a region of attraction for x_f^*.■

Proof of Theorem V.1′

The proof is based on the fact that the properties of the open-loop system are characterized by the properties of the operator O_f of Def. V.1.

1. By (V.1′) an equilibrium state x_f^* is a fixed point of the operator O_f. By assumption

$$||\frac{\partial O_f}{\partial x}|| \leq \theta < 1, \quad \forall x \in \bar{U}(x^0, r)$$

Theorem III.2 then asserts that x_f^* exists and is unique.

2. Let r_f in Def. V.3 be equal to r^0. By the assumptions of this theorem, Theorem III.2 implies that if $x^{s+1} = O_f(x^s)$ then

$$||x^s - x_f^*|| \leq \theta^s ||x^0 - x_f^*||, \quad \forall x^0 \in \bar{U}(x_f^*, r_f)$$

Choose any $\epsilon > 0$ and let $s_0 > [ln(\epsilon/||x^0 - x_f^*||)/ln(\theta)]$. For $s > s_0$:

$$||x^s - x_f^*|| \leq \theta^s ||x^0 - x_f^*|| < \theta^{s_0} ||x^0 - x_f^*|| < \epsilon \quad \forall x^0 \in \bar{U}(x_f^*, r_f)$$

By Def. V.3, there exists $r_f = r^0 > 0$ such that for any $\epsilon > 0$ there exists $s_0 = s_0(\epsilon)$ such that for every trajectory that starts in $\bar{U}(x_f^*, r_f)$, $||x^s - x_f^*|| < \epsilon \; \forall s > s_0$.

3. Part 2 of the proof establishes that $\lim_{s \to \infty} ||x^s - x_f^*|| = 0$, i.e.

$$\lim_{s \to \infty} x^s = x_f^*, \quad \forall x^0 \in \bar{U}(x_f^*, r^0)$$

i.e. every trajectory starting at an initial state $x \in \bar{U}(x_f^*, r^0)$ converges to x_f^*. By Def. V.4 it follows that $\bar{U}(x_f^*, r^0)$ is a region of attraction for x_f^*.■

Proof of Corollary V.2

1. Example II.3 shows that the system operator O_f is

$$x^{s+1} = e^{AT} x^s + (e^{AT} - I) A^{-1} B u_f$$

To obtain the equilibrium state, the following equation is solved

$$x_f^* = e^{AT}x_f^* + (e^{AT} - I)A^{-1}Bu_f \tag{V.50}$$

yielding

$$x_f^* = -A^{-1}Bu_f(\dagger)$$

Corollary III.1 establishes that x_f^* is the unique fixed point of O_f, i.e. the unique equilibrium state of the system. It also establishes that every trajectory starting at any $x^0 \in \mathbf{R}^n$ converges to x_f^* and consequently x_f^* is globally stable.

2. Assume that $(e^{AT} - I)$ is invertible. Then again $x_f^* = -A^{-1}Bu_f$ is the unique equilibrium state because (V.50) has a unique solution. Select a ball $\bar{U}(x_f^*, r_f)$. For any $r_f > 0$ this ball contains a vector with non-zero projection on the eigenvector associated with the eigenvalue of e^{AT} of maximum magnitude. It was shown in Corollary III.1 that this vector will be amplified to infinity. This shows that there is no $r_f > 0$ such that $||x^s - x_f^*|| < \epsilon, \forall s > s_0$.■

<u>**Proof of Theorem V.3**</u>

Only the case of open-loop stability (Theorem V.1) is treated. The closed-loop case (Theorem V.2) proof is identical.

1. An equilibrium state is stable by Theorem V.1 when there is some $r_f > 0$ such that

$$||\frac{\partial \chi(T; x, u_f)}{\partial x}|| < 1, \quad \forall x \in \bar{U}(x_f^*, r_f)$$

which implies that

$$||\frac{\partial \chi(T; x_f^*, u_f)}{\partial x_f^*}|| < 1 \tag{V.51}$$

The quantity $\Phi^* = \dfrac{\partial \chi(T; x_f^*, u_f)}{\partial x_f^*}$ is computed next. To this end let $A = \left.\dfrac{\partial f(\varsigma, u_f)}{\partial \varsigma}\right|_{\varsigma = x_f^*}$. From (II.3), Φ^* is the solution at $t = T$ of the initial value problem

$$\frac{\partial \Phi(t)}{\partial t} = \left.\frac{\partial f(\varsigma, \xi)}{\partial \varsigma}\right|_{\substack{\varsigma = \chi(t; x_f^*, u_f) \\ \xi = u_f}} \Phi(t) \tag{V.52}$$

where $\Phi(0) = I$ and $\chi(t; x_f^*, u_f)$ is the solution of the initial value problem

$$\frac{\partial \chi}{\partial t} = f(\chi, u_f), \quad \chi(0) = x_f^* \tag{V.53}$$

($\dagger$) Note that the factor $(e^{AT} - I)$ was cancelled because the assumption $\rho(e^{AT}) < 1$ implies that $(e^{AT} - I)$ is invertible.

(V.53) has the (unique) solution $\chi(t;x_f^*,u_f) = x_f^*$, which upon substitution in (V.52) generates the linear initial value problem

$$\frac{\partial\Phi(t)}{\partial t} = \left.\frac{\partial f(\varsigma,\xi)}{\partial\varsigma}\right|_{\substack{\varsigma = x_f^* \\ \xi = u_f}} \Phi(t) = A\Phi(t) \tag{V.54}$$

(V.54) is analytically integrated (example II.6) to

$$\Phi(t) = e^{At} \quad \text{and} \quad \Phi^* = e^{AT}$$

Then from (V.51) $||e^{AT}|| < 1$ and from Lemma III.6:

$$\rho(e^{AT}) < 1$$

which implies that all the eigenvalues of A have negative real parts.

2. Lemma V.1 implies that all the eigenvalues of e^{AT} are inside the unit circle, therefore $\rho(e^{AT}) < 1$, i.e. $\rho(\frac{\partial\chi(T;x_f^*,u_f)}{\partial x_f^*}) < 1$. Then from Lemma III.6 there exists some induced operator norm $||\cdot||_i$ such that

$$||\frac{\partial\chi(T;x_f^*,u_f)}{\partial x}||_i < 1$$

By the assumptions 1.i-vi of Chapter II, $\Phi(x) = \frac{\partial\chi(T;x,u_f)}{\partial x}$ is a continuous function of x, therefore $||\Phi(x)||_i$ is a continuous function of x also (because the norm function is continuous). It follows that there exists $r > 0$, such that

$$||\Phi(x)||_i < 1, \quad \forall x \in \bar{U}(x_f^*,r)$$

Theorem V.1 asserts that x_f^* is stable.

3. If an equilibrium state is unstable in the sense of Lyapunov, the argumentation of the previous two parts of the proof can be repeated to show that $\rho(\Phi^*) > 1$. By the continuity of Φ and the continuity of the spectral radius, it follows that there is exists $r > 0$ such that

$$\rho(\Phi(x)) > 1, \quad \forall x \in \bar{U}(x_f^*,r)$$

From Lemma III.6 then, for all induced norms

$$||\Phi(x)|| > 1, \quad \forall x \in \bar{U}(x_f^*,r)$$

Consequently for any $r' > 0$ there exist $x \in \bar{U}(x_f^*,r')$ such that $||\Phi(x)|| > 1$ and the conditions of Theorem V.1 cannot be satisfied.

4. The proof is by contradiction. If the equilibrium state was Lyapunov stable, then

$$\rho(\frac{\partial\chi(T;x_f^*,u_f)}{\partial x_f^*}) < 1$$

Then by the result 2 of this theorem there would exist some norm and some ball centered at x_f^* that would satisfy the conditions of Theorem V.1, which is not possible by assumption.■

Proof of Theorem V.4

The uniqueness is established by Corollary V.1. Then it will be shown that $V(x) = ||\chi(T;x,u_f) - x||_\alpha$ is a Lyapunov function.

It is shown that $||\chi(T;x,u_f) - x||$ has the three properties required by Lemma V.2.

i. Follows immediately from the continuity assumption II.1.i-vi.

ii. The case $\alpha = 2$ is treated first:

$$||\chi(T;x,u_f) - x||_2 = ||\chi(T;x,u_f) - \chi(T;0,u_f) - x||_2 \leq ||\chi(T;x,u_f) - \chi(T;0,u_f)||_2 + ||x||_2$$
$$\overset{\text{Lem.III.1}}{\leq} \max_{\lambda\in[0,1]} || \left.\frac{\partial\chi(T;\varsigma,u_f)}{\partial\varsigma}\right|_{\varsigma=\lambda x} || \, ||x-0||_2 + ||x||_2 \leq \theta||x||_2 + ||x||_2$$
$$\leq (1+\theta)||x||_2 \leq \eta_2'||x||_2, \text{ for } \eta_2' = 1+\theta$$

Also

$$||\chi(T;x,u_f) - x||_2 \geq | \, ||x||_2 - ||\chi(T;x,u_f)||_2 \, |$$
$$\geq \left| \, ||x||_2 - \max_{\lambda\in[0,1]} || \left.\frac{\partial\chi(T;\varsigma,u_f)}{\partial\varsigma}\right|_{\varsigma=\lambda x} ||_2 \, ||x||_2 \right|$$
$$= (1-\theta)||x||_2 = \eta_1', \text{ for } \eta_1' = (1-\theta) > 0$$

For any vector norm in a finite dimensional space it is known (Golub and Van Loan, 1984) that there exist positive constants ν_1 and ν_2 such that $\nu_2||x||_\alpha \leq ||x||_2 \leq \nu_1||x||_\alpha$. Setting $\eta_1 = \eta_1'/\nu_1$ and $\eta_2 = \eta_2'\nu_2$ it immediately follows from the above that

$$\eta_1||x||_2 \leq ||\chi(T;x,u_f) - x||_\alpha \leq \eta_2||x||_2$$

iii. It is to be shown that

$$||\chi(T;\chi(T;x,u_f),u_f) - \chi(T;x,u_f)||_\alpha \leq \eta_3||\chi(T;x,u_f) - x||_\alpha$$

Indeed

$$||\chi(T;\chi(T;x,u_f),u_f) - \chi(T;x,u_f)||_\alpha \leq \sup||\frac{\partial\chi(T;x,u_f)}{\partial x}||_{\alpha\to\alpha} \, ||\chi(T;x,u_f) - x||_\alpha$$
$$= \theta||\chi(T;x,u_f) - x||_\alpha$$
$$= \eta_3||\chi(T;x,u_f) - x||_\alpha$$

for $\eta_3 = \theta$.■

Proof of Theorem V.5

The proof is by contradiction. Suppose Ω contained a stable equilibrium state x_f^*. It was shown in parts 1 and 2 of the proof to Theorem V.3 that in this case

$$\rho\left(\frac{\partial\chi(T;x_f^*,u_f)}{\partial x_f^*}\right) < 1, \quad x_f^* \in \Omega$$

This contradicts the assumption of the theorem.

Proof of Theorem V.9

1. The proof is based on Theorem III.2, in exactly the same fashion that Theorem 1 was based on Theorem III.3.

Choose some $p \in I_p$. The existence of a stable equilibrium state of the system will be established for the particular p. Thus the system will be shown to be stable for every $p \in I_p$.

Let $z = v$, $z^0 = v^*_{p_0}$, $F(z) = \pi(v;p)$. Then

$$R \stackrel{\text{def}}{=} \|F(z^0) - z^0)\|/(1-\theta) = \left\| \begin{matrix} \chi(x^*_{p_0}, u^*_{p_0}; p) - x^*_{p_0} \\ Qg(x^*_{p_0}; p) - Qg(x^*_{p_0}) \\ 0 \end{matrix} \right\| \leq r^0$$

Theorem II.2 applies directly to show that there is a unique stable equilibrium state in $\bar{U}(z^0, r^0)$, i.e. in $\bar{U}(v^*_{p_0}, r^0)$.

2. Assume that v^*_p is the equilibrium state of the system. Then, $\bar{U}(v^*_p, r_p) \subseteq \bar{U}(v^*_{p_0}, r)$ for $r_p = r - r^0$. From the assumption it follows that

$$\frac{\partial \pi(v)}{\partial v} \leq \theta < 1, \quad \forall v \in \bar{U}(v^*_p, r_p)$$

Theorem V.6 applies to show that v^*_p is unique in $\bar{U}(v^*_p, r_p)$, it is stable and has $\bar{U}(v^*_p, r_p)$ as a region of attraction.■

Proof of Theorem V.11

The proof of claims 1, 2 and 3 is based on Theorem V.1′. To this end let $x^0 = x^*_f$. Then

$$\begin{aligned} r^0 &= \|\chi(x^0, u_f; d) - x^0\|/(1-\theta) \\ &= \|\chi(x^0, u_f) + d - x^0\|/(1-\theta) \\ &= \|\chi(x^*_f, u_f) + d - x^*_f\|/(1-\theta) \\ &= \|x^*_f + d - x^*_f\|/(1-\theta) \\ &\leq \delta/(1-\theta) \stackrel{\text{def}}{=} R \end{aligned}$$

Theorem V.1′ applies to assert that the disturbed system has a unique equilibrium state (i.e. x^*_d by definition) in $\bar{U}(x^*_f, R)$ which is stable.

4. It is first shown that a trajectory that starts in $\bar{U}(x^*_f, r)$ remains in this same ball. This is done by induction.

For $k = 0$, x^0 is in the ball because the trajectory starts in it.

For $k = s$ assume that x^s is in the ball and therefore

$$\|x^s - x^*_f\| \leq r \tag{V.55}$$

Then

$$\begin{aligned}
\|x^{s+1} - x_f^*\| &= \|\chi(x^s, u_f) + d - x_f^*\| \\
&\leq \|\chi(x^s, u_f) - x_f^*\| + \|d\| \\
&= \|\chi(x^s, u_f) - \chi(x_f^*, u_f)\| + \|d\| && \text{From eq. V.1}' \\
&\leq \theta\|x^s - x_f^*\| + \delta && \text{From the contraction assumption} \\
&\leq \theta r + \delta && \text{Induction assumption} \\
&= \theta r + (1-\theta)R && \text{From the definition of } R \\
&\leq \theta r + (1-\theta)r && (r \geq R) \text{ by assumption} \\
&= r
\end{aligned}$$

Therefore the trajectory lies entirely in $\bar{U}(x_f^*, r)$ where the contraction condition holds. Then

$$\begin{aligned}
\|x^{s+1} - x_d^*\| &= \|\chi(x^s, u_f) + d - x_d^*\| \\
&= \|\chi(x^s, u_f) + d - (\chi(x_d^*, u_f) + d)\| && \text{From (V.39)} \\
&= \|\chi(x^s, u_f) - \chi(x_d^*, u_f)\| \\
&\leq \theta\|x^s - x_d^*\|
\end{aligned}$$

where the last inequality is based on the contraction assumption, as well as the fact that x^s is in the contraction ball as was shown above. It is easy to show then that

$$\|x^s - x_d^*\| \leq \theta^s\|x^0 - x_d^*\|$$

and consequently that every trajectory, starting anywhere in $\bar{U}(x_f^*, r)$ converges to x_d^*. Therefore by the definition of a region of attraction $\bar{U}(x_f^*, r)$ is a region of attraction for x_d^*.■

Proof of Theorem V.12

1. The proof is by induction.

For $k = 0$ it is true, because by assumption the trajectory starts at $x^0 \in \bar{U}(x_f^*, r^0)$.

For $k = s$ assume that

$$\|x^s - x_f^*\| \leq R \tag{V.56}$$

Then

$$\begin{aligned}
\|x^{s+1} - x_f^*\| &= \|\chi(x^s, u_f; d^s) - \chi(x_f^*, u_f; 0)\| \\
&= \|\chi(x^s, u_f; d^s) - \chi(x^s, u_f; 0) + \chi(x^s, u_f; 0) - \chi(x_f^*, u_f; 0)\| \\
&\leq \|\chi(x^s, u_f; d^s) - \chi(x^s, u_f; 0)\| + \|\chi(x^s, u_f; 0) - \chi(x_f^*, u_f; 0)\| \\
&\leq B\|d^s\| + \theta\|x^s - x_f^*\| && \text{From contraction and B definition} \\
&\leq B\delta + \theta\|x^s - x_f^*\| && \text{Bound on disturbance} \\
&= (1-\theta)R + \theta\|x^s - x_f^*\| && R \text{ definition} \\
&\leq (1-\theta)R + \theta R && \text{Induction assumption} \\
&= R
\end{aligned}$$

2. It is enough to show that at the s^{th} sampling interval, any trajectory will be in a ball $\bar{U}(x_f^*, R + \theta^s(r_0 - R))$.This is shown by induction.
For $k = 0$, $x^0 \in \bar{U}(x_f^*, R + r_0 - R)$ by assumption.
For $k = s$ assume that $x^s \in \bar{U}(x_f^*, R + \theta^s(r_0 - R))$. Then

$$\begin{aligned}
\|x^{s+1} - x_f^*\| &= \|\chi(x^s, u_f; d^s) - \chi(x_f^*, u_f; 0)\| \\
&= \|\chi(x^s, u_f; d^s) - \chi(x^s, u_f; 0) + \chi(x^s, u_f; 0) - \chi(x_f^*, u_f; 0)\| \\
&\le \|\chi(x^s, u_f; d^s) - \chi(x^s, u_f; 0)\| + \|\chi(x^s, u_f; 0) - \chi(x_f^*, u_f; 0)\| \\
&\le B\|d^s\| + \theta\|x^s - x_f^*\| && \text{From contraction and B definition} \\
&\le B\delta + \theta\|x^s - x_f^*\| && \text{Bound on disturbance} \\
&\le B\delta + \theta(R + \theta^s(r_0 - R)) && \text{Induction assumption} \\
&= (1-\theta)R + \theta(R + \theta^s(r_0 - R)) && R \text{ definition} \\
&= (1-\theta)R + \theta R + \theta^{s+1}(r_0 - R) \\
&= R + \theta^{s+1}(r_0 - R)
\end{aligned}$$

This proves that $x^{s+1} \in \bar{U}(x_f^*, R + \theta^{s+1}(r_0 - R))$.■

<u>**Proof of Theorem V.13**</u>

1. At $s = \infty$, $d^\infty = 0$, therefore

$$\chi(x_f^*, u_f; 0) = x_f^*$$

2. The conditions of Theorem V.12 are satisfied if the conditions of the current Theorem V.13 are. Therefore, it is inferred that every trajectory that starts inside $\bar{U}(x_f^*, r)$ remains in this ball. Therefore the contraction conditions will hold along any trajectory. Then:

$$\begin{aligned}
\|x^{s+1} - x_f^*\| &= \|\chi(x^s, u_f; d^s) - \chi(x_f^*, u_f; 0)\| \\
&= \|\chi(x^s, u_f; d^s) - \chi(x^s, u_f; 0) + \chi(x^s, u_f; 0) - \chi(x_f^*, u_f; 0)\| \\
&\le \|\chi(x^s, u_f; d^s) - \chi(x^s, u_f; 0)\| + \|\chi(x^s, u_f; 0) - \chi(x_f^*, u_f; 0)\| \\
&\le B\|d^s\| + \theta\|x^s - x_f^*\| \\
&\le \lambda B\delta + \theta\|x^s - x_f^*\|
\end{aligned}$$

In a similar fashion

$$\begin{aligned}
\|x^s - x_f^*\| &\le \lambda B\delta + \theta\|x^{s-1} - x_f^*\| \\
&\vdots \\
\|x^1 - x_f^*\| &\le \lambda B\delta + \theta\|x^0 - x_f^*\|
\end{aligned}$$

Then, backward substitution leads to

$$\begin{aligned}
\|x^s - x_f^*\| &\le B\delta \sum_{i=0}^{i=s-1} \theta^i \lambda^{s-i} + \theta^s\|x^0 - x_f^*\| \\
&\le \theta^s\|x^0 - x_f^*\| + sB\delta max\{\lambda, \theta\}^{s-1}
\end{aligned} \tag{V.57}$$

Since $\lim_{s\to\infty}(sB\delta\eta^s) = 0$ if $0 \leq \eta < 1$ and $0 \leq \eta \stackrel{\text{def}}{=} max\{\lambda, \theta\} < 1$ by assumption, (V.57) shows that the trajectory converges to x_f^*.

3. The previous paragraph established that every trajectory converges to x_f^*, therefore by definition $\bar{U}(x_f^*, r)$ is a region of attraction.■

CHAPTER VI

CONTRACTION PRINCIPLE SYNTHESIS

In Chapter IV control objectives were posed as operator equations. In this Chapter, successive substitution algorithms are employed to solve these equations and generate control laws. The results of Chapter V motivate the search for control laws that generate closed-loop operators with desired contraction properties.

The first section provides the basic synthesis framework. Linear control laws for nonlinear systems are developed in the next section and their properties and implications are investigated in detail. In section 3 nonlinear control laws are proposed by extrapolating the results of section 2. Finally the Chapter concludes with the remarks of section 4.

1. INTRODUCTION

A general synthesis framework for control laws derived from Contraction Mapping arguments is developed. It provides a unified treatment of synthesis problems associated with the three basic types of system operators considered in Chapter V (Def. V.6, V.7 and V.9). Later in the Chapter (section 2) the framework is specialized to each respective operator.

Assume that a particular control objective is expressed in the form of an operator equation :

$$P(v) = 0 \qquad \text{(VI.1)}$$

with v an augmented state vector, possibly containing system and/or model states and control inputs.

Example VI.1 : For the control objective expressed in (IV.5) it is $v = \begin{pmatrix} x & u \end{pmatrix}^T$ *and* $P(v) = S(x,u) - (x, y^*)$.

Following the guidelines of section III.1, (VI.1) is cast in an equivalent successive substitution form :

$$v = v + Q(P(v)) \stackrel{\text{def}}{=} F(v) \qquad \text{(VI.2)}$$

provided that $Q(\cdot)$ is continuous and invertible. As long as (VI.2) is to be used for control law computations, an additional condition is imposed on $Q(\cdot)$, that of generating consistent algorithms in the sense of (IV.6). The set of consistent operators Q will be denoted by $\mathbf{Q}_c$. The control algorithm becomes :

$$v^{s+1} = F(v^s) = v^s + Q(P(v^s)), \qquad Q \in \mathbf{Q}_c \tag{VI.3}$$

F is the operator that characterizes the behavior of the closed-loop system. In particular, its contraction properties determine the stability properties of the system. The properties of F depend on the operators P and Q. Although P is a fixed operator, Q is adjustable. Motivated by the discussion in Chapters III and V, the design problem is to select Q such that either of the following two objectives is attained.

i) (Performance objective) Given an equilibrium state v^* of F and a ball $\bar{U}(v^*, r)$, find Q such that the rate of settling to equilibrium from any initial state inside $\bar{U}(v^*, r)$, is maximal. By Theorem III.3, in the worst case it is :

$$\|v^s - v^*\| \leq \left[\sup_{v \in \bar{U}(v^*,r)} \|F'(v)\| \right]^s \|v^0 - v^*\| \tag{VI.4}$$

Consequently the performance objective generates the min-max optimization problem :

$$\min_{Q \in \mathbf{Q}_c} \sup_{v \in \bar{U}(v^*,r)} \|F'(v)\| \tag{VI.5}$$

ii) (Stability objective). Given an equilibrium state v^* of F, find Q such that the radius r of a ball of attraction $\bar{U}(v^*, r)$ is maximal. The constrained optimization problem arises :

$$\max_{Q \in \mathbf{Q}_c} r \tag{VI.6}$$

under the condition

$$\sup_{v \in \bar{U}(v^*,r)} \|F'(v)\| < 1 \tag{VI.7}$$

In what follows, emphasis will be placed on the first objective. Treatment of the second objective follows in the same fashion.

If Q_{opt} is the solution of (VI.5) with $\theta_{opt} = \sup \|F'_{Q_{opt}}(v)\| < 1$, the following theorem summarizes the properties of the resulting closed-loop system.

Theorem VI.1 *: Consider the closed-loop system operator F generated by system P and control law Q_{opt}, with Q_{opt} the solution of (VI.5). If $\theta_{opt} \stackrel{\text{def}}{=} \sup\|F'_{Q_{opt}}(v)\| < 1$, the closed-loop system has a unique, stable equilibrium state in $\bar{U}(v^*, r)$. The system returns to v^* after any perturbation to $v^0 \in \bar{U}(v^*, r)$ at least as fast as*

$$\|v^s - v^*\| \leq \theta_{opt}^s \|v^0 - v^*\| \tag{VI.8}$$

Proof : Follows trivially from Theorem V.2.

Solving problems of the type (VI.5) or (VI.6) for general nonlinear operators corresponds to a nonlinear optimization search over an infinite dimensional space of operators, which is not tractable at the current stage. The problems are in principle solvable when Q is restricted to a set of linear operators. Linear Q generate linear control laws for nonlinear systems, which are investigated in the next section.

2. LINEAR CONTROL LAWS

In this section the framework of section 1 is specialized to three important design problems. The structure of the section closely follows the development of the analysis theory in Chapter V for various sets of assumptions.

2.1 The State-Feedback Case

Assumptions The model of the system is exact

The state vector is completely known

Operator Equation

The operator equation expressed by (IV.5), namely to drive the system at steady state with its output at a desired value y^*.

$$\begin{aligned} x &= \chi(x,u) && \text{(System at steady state)} \\ y^* &= g(x) && \text{(Output at desired value } y^*\text{)} \end{aligned} \tag{VI.9}$$

Successive Substitution Form

$$\begin{pmatrix} x^{s+1} \\ u^{s+1} \end{pmatrix} = \begin{pmatrix} x^s \\ u^s \end{pmatrix} + Q \begin{pmatrix} \chi(x^s,u^s) - x^s \\ y^* - g(x^s) \end{pmatrix} \tag{VI.10}$$

The set of consistent Q's

A *candidate* linear operator Q

$$\mathbf{R}^n \times \mathbf{R}^m \longmapsto \mathbf{R}^n \times \mathbf{R}^m$$

$$(\eta,\xi) \overset{Q}{\longmapsto} Q(\eta,\xi)$$

has the properties

i. $Q(\eta,\xi) = (0,0) \iff (\eta,\xi) = (0,0)$

ii. Q is consistent (in the sense that the algorithm (VI.10) generated by Q is consistent).

Consistency requirements impose a special structure on candidate Q's : any Q is a linear operator mapping a finite dimensional space $\mathbf{R}^{n+m}$ into a finite dimensional space $\mathbf{R}^{n+m}$, therefore it has a matrix representation which is unique with respect to a basis of $\mathbf{R}^{m+n}$. This matrix representation is partitioned

$$\mathrm{Q} = \begin{array}{c} \begin{matrix} \;\; n & \;\;\; m \end{matrix} \\ \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \end{array} \begin{matrix} n \\ m \end{matrix}$$

(VI.10) yields then

$$x^{s+1} = x^s + Q_{11}(\chi(x^s,u^s) - x^s) + Q_{12}(y^* - g(x^s)) \quad \text{(VI.11.}a)$$
$$u^{s+1} = u^s + Q_{21}(\chi(x^s,u^s) - x^s) + Q_{22}(y^* - g(x^s)) \quad \text{(VI.11.}b)$$

From (IV.6) $x^{s+1} = \chi(x^s, u^s)$, which upon substitution in (VI.11.a) implies

$$\begin{aligned} &\chi(x^s,u^s) = x^s + Q_{11}(\chi(x^s,u^s) - x^s) + Q_{12}(y^* - g(x^s)) \\ \Longleftrightarrow\ &(Q_{11} - I)(\chi(x^s,u^s) - x^s) + Q_{12}(y^* - g(x^s)) = 0 \\ \Longleftrightarrow\ &\begin{pmatrix} Q_{11} - I & Q_{12} \end{pmatrix} \begin{pmatrix} \chi(x^s,u^s) - x^s \\ y^* - g(x^s) \end{pmatrix} = 0, \quad \forall \begin{pmatrix} \chi(x^s,u^s) - x^s \\ y^* - g(x^s) \end{pmatrix} \end{aligned} \quad \text{(VI.12)}$$

Without loss of generality (because the vectors $\begin{pmatrix} \chi(x^s,u^s) - x^s & y^* - g(x^s) \end{pmatrix}^T$ span $\mathbf{R}^{n+m}$ in non-degenerate cases) (VI.12) implies

$$\begin{pmatrix} Q_{11} - I & Q_{12} \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$$

i.e. $Q_{11} = I$ and $Q_{12} = 0$.
Summing up, a consistent operator Q will be of the form

$$Q = \overset{\quad n \qquad m}{\begin{pmatrix} I & 0 \\ Q_{21} & Q_{22} \end{pmatrix}} \begin{matrix} n \\ m \end{matrix} \qquad Q_{22} \text{ non-singular}$$

where the non-singularity condition is imposed by the requirement that a consistent Q is invertible.

<u>The Control Algorithm</u>

For $Q \in \mathbf{Q}_c$ the successive substitution form (VI.10) generates the control algorithm

$$x^{s+1} = \chi(x^s, u^s) \quad \text{(VI.13.}a)$$
$$u^{s+1} = u^s + Q_{21}(\chi(x^s,u^s) - x^s) + Q_{22}(y^* - g(x^s)) \quad \text{(VI.13.}b)$$

(VI.13.b) is the candidate control law. It has a straightforward interpretation: the second term on the right is state feedback through Q_{21}, the third term is output feedback through Q_{22}, while the first term accounts for the integral action of the control law. Fig. VI.1 shows a block diagram realization of the control law.

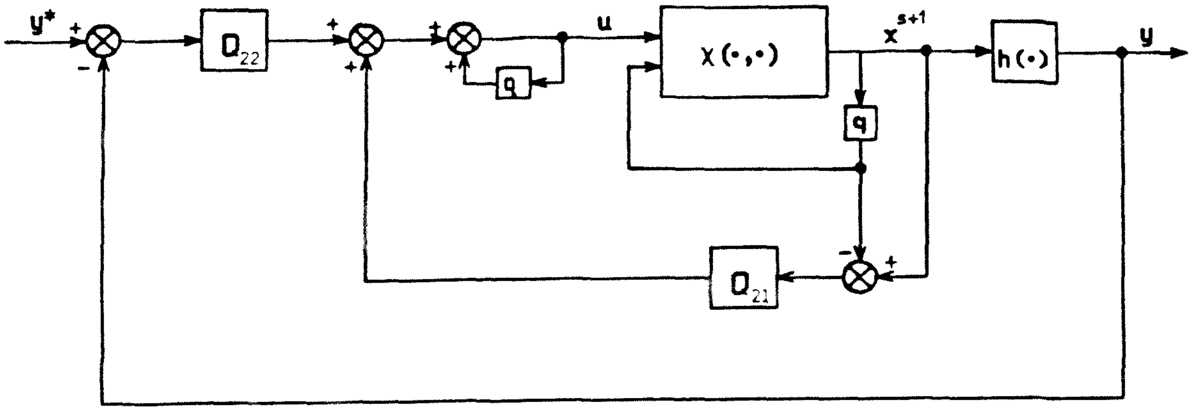

FIGURE VI.1 : *Block structure of the state feedback control law (q: backward shift-delay-operator).*

The Design Problem

Differentiating the right hand side of (VI.13) with respect to $x = x^s$ and $u = u^s$ and letting $\Phi(x,u) = \dfrac{\partial\chi(x,u)}{\partial x}$, $\Gamma(x,u) = \dfrac{\partial\chi(x,u)}{\partial u}$ and $C(x) = \dfrac{\partial g(x)}{\partial x}$, the design problem (VI.5) translates in this case :

The State-Feedback Design Problem

$$\min_{Q\in\mathbf{Q}_c} \max_{(x,u)\in\bar{C}((x^*,u^*),r)} \left\| \begin{matrix} \Phi(x,u) & \Gamma(x,u) \\ Q_{21}(\Phi(x,u)-I) - Q_{22}C(x) & I + Q_{21}\Gamma(x,u) \end{matrix} \right\|$$

Implications for Linear Systems

The design problem is simplified when the system is linear : the entries in the norm are independent of x and u and the problem becomes (using the results of section II.3)

The Linear State-Feedback Design Problem

$$\min_{Q\in\mathbf{Q}_c} \rho \begin{pmatrix} e^{AT} & (e^{AT}-I)A^{-1}B \\ Q_{21}(e^{AT}-I) - Q_{22}C & I + Q_{21}(e^{AT}-I)A^{-1}B \end{pmatrix}$$

The problem is equivalent to a classical pole placement problem, which (assuming observability and controllability) is analytically solvable. The solution yields $\rho = 0$, i.e. all the closed-loop poles are shifted to the origin. The corresponding control law is commonly known as "state dead-beat controller".

Example VI.2 : Consider the linear, continuous system

$$\frac{\partial x_1}{\partial t} = 4x_1 - 3x_2 + u$$
$$\frac{\partial x_2}{\partial t} = x_1$$
$$y = x_2$$

with a sampling time of $T = 0.1$. From (II.5) and (II.8) it is

$$\Phi = e^{AT} = \begin{pmatrix} 1.470 & -0.367 \\ 0.122 & 0.938 \end{pmatrix}, \quad \Gamma = (e^{AT} - I)A^{-1}B = \begin{pmatrix} .122 \\ 5.7 \times 10^{-3} \end{pmatrix}$$

Let

$$Q_{21} = \begin{pmatrix} \alpha & \beta \end{pmatrix}, \quad Q_{22} = \begin{pmatrix} \gamma \end{pmatrix}$$

The linear design problem becomes

$$\min_{\alpha,\beta,\gamma\in\mathbf{R}} \rho \begin{pmatrix} 1.470 & -0.367 & 0.122 \\ .122 & 0.983 & 5.7 \times 10^{-3} \\ 0.47\alpha + 0.122\beta & -0.367\alpha - 0.017\beta - \gamma & 1 + 0.122\alpha + 5.7 \times 10^{-3}\beta \end{pmatrix}$$

The coefficients of the characteristic polynomial of the design problem matrix are equated to zero (so that all eigenvalues, therefore ρ, are at zero). This results in a system of three linear equations for α, β and γ, with the solution

$$\alpha = -18.7$$
$$\beta = -203.7$$
$$\gamma = 81.7$$

These values are substituted in (VI.13.b), yielding the control law

$$u^{s+1} = u^s - 18.7(\chi_1^s - x_1^s) - 203.7(\chi_2^s - x_2^s) + 81.7(y^* - y^s)$$

Fig. VI.2 shows the response of the system under this control law to a step up setpoint command (y^* from 0.0 to 1.0). The system settles after three steps, displaying the state deadbeat characteristic.

Suboptimal Control Laws

In some cases, associated with smooth nonlinear systems, it is possible to obtain simple, yet efficient control algorithms without going into great lengths to solve the optimization design problem of this section. Two simplified procedures are discussed next.

i. (Local linearization design). An equilibrium state is located and the system is linearized around this state. The linear state-feedback case design problem is solved to compute Q_{21}

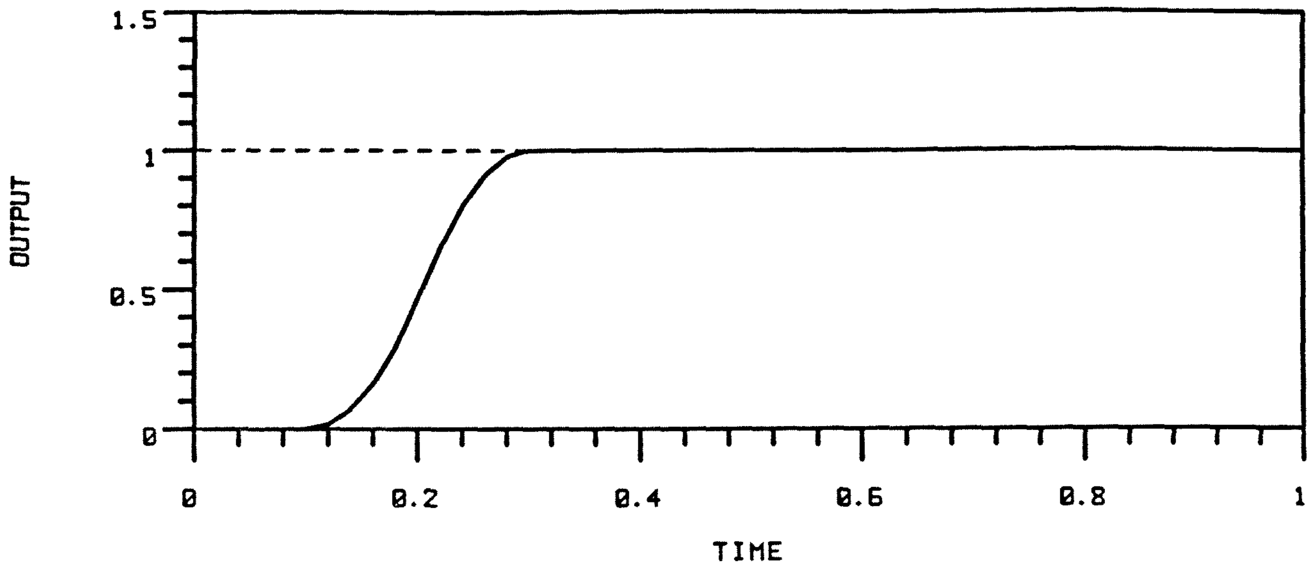

FIGURE VI.2 : *Contraction controller design for linear unstable system. Setpoint tracking. (– – – – : Setpoint command, ——— : Output.)*

and Q_{22}. They are substituted in (VI.13) in turn to generate the control law. Theorem V.2 is used to study the stability of the resulting closed-loop system. An induced norm that approximates the closed-loop spectral radius at the equilibrium state should be used for the analysis.

ii. Some induced norms (for example the 1- and ∞- norms) have the property :

$$\left\| \begin{matrix} A & B \\ 0 & 0 \end{matrix} \right\| \leq \left\| \begin{matrix} A & B \\ C & D \end{matrix} \right\|, \quad \forall \quad A, B, C, D$$

Based on this property, suboptimal control laws are sought by selecting appropriate Q's that make the adjustable blocks (C and D) of the design problem equal to zero. It can be easily shown by substitution, that the following operators have the desired property :

$$\begin{aligned} Q_{21} &= -\left[C(x)(\Phi(x,u) - I)^{-1}\Gamma(x,u)\right]^{-1} C(x)(\Phi(x,u) - I)^{-1} \\ Q_{22} &= -\left[C(x)(\Phi(x,u) - I)^{-1}\Gamma(x,u)\right]^{-1} \end{aligned} \qquad \text{(VI.14)}$$

However, both Q_{21} and Q_{22} in (VI.14) are nonlinear operators in x and u. To obtain linear operators, (VI.14) is computed at a fixed reference state (x_{ref}, u_{ref}) – in most cases the desired equilibrium state is an appropriate reference – yielding the control law:

$$\begin{aligned} Q_{21} &= -\left[C(x_{ref})(\Phi(x_{ref}, u_{ref}) - I)^{-1}\Gamma(x_{ref}, u_{ref})\right]^{-1} C(x_{ref})(\Phi(x_{ref}, u_{ref}) - I)^{-1} \\ Q_{22} &= -\left[C(x_{ref})(\Phi(x_{ref}, u_{ref}) - I)^{-1}\Gamma(x_{ref}, u_{ref})\right]^{-1} \end{aligned} \qquad \text{(VI.15)}$$

Theorem V.2 is then used for stability analysis of the resulting closed-loop system.

2.2 The Model Reference Case

Assumptions The model of the system is exact

Operator Equation

The control objective is to drive the system and model to steady state with the output at y^*. The equation associated with the model will be assumed to be of the form :

$$z^{s+1} = \chi(z^s, u^s) + \bar{Q}(g(x^s) - g(z^s)) \tag{VI.16}$$

because the resulting model evolution exactly matches the system evolution when $z^s = x^s$. In this context, the operator equation is :

$$\begin{array}{ll} x = \chi(x, u) & \text{(System at steady state)} \\ z = \chi(z, u) + \bar{Q}(g(x) - g(z)) & \text{(Model matching of system evolution)} \\ y^* = g(x) & \text{(Output at desired value } y^* \text{)} \end{array} \tag{VI.17}$$

Successive Substitution Form

$$\begin{pmatrix} x^{s+1} \\ z^{s+1} \\ u^{s+1} \end{pmatrix} = \begin{pmatrix} x^s \\ z^s \\ u^s \end{pmatrix} + Q \begin{pmatrix} \chi(x^s, u^s) - x^s \\ \chi(z^s, u^s) + \bar{Q}(g(x^s) - g(z^s)) - z^s \\ y^* - g(x^s) \end{pmatrix} \tag{VI.18}$$

The set of consistent Q's

A *candidate* linear operator Q

$$\mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^m \longmapsto \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^m$$

$$(\nu, \eta, \xi) \overset{Q}{\longmapsto} Q(\nu, \eta, \xi)$$

has the the properties

i. $Q(\nu, \eta, \xi) = (0,0,0) \iff (\nu, \eta, \xi) = (0,0,0)$

ii. Q is consistent. (in the sense that the algorithm (VI.18) generated by Q is consistent).

Consistency requirements impose a special structure on candidate Q's : any Q is a linear operator mapping a finite dimensional space $\mathbf{R}^{n+n+m}$ into a finite dimensional space $\mathbf{R}^{n+n+m}$, therefore it has a matrix representation which is unique with respect to a basis of $\mathbf{R}^{n+n+m}$. This matrix representation is partitioned

$$\mathrm{Q} = \begin{array}{c} \begin{array}{ccc} n & n & m \end{array} \\ \begin{pmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{pmatrix} \end{array} \begin{array}{c} n \\ n \\ m \end{array}$$

(VI.18) becomes then

$$x^{s+1} = x^s + Q_{11}(\chi(x^s, u^s) - x^s) + Q_{12}(\chi(z^s, u^s) + \bar{Q}(g(x^s) - g(z^s)) - z^s) + Q_{13}(y^* - g(x^s))$$

$$(VI.19.a)$$

$$z^{s+1} = z^s + Q_{21}(\chi(x^s,u^s) - x^s) + Q_{22}(\chi(z^s,u^s) + \bar{Q}(g(x^s) - g(z^s)) - z^s) + Q_{23}(y^* - g(x^s)) \quad (VI.19.b)$$

$$u^{s+1} = u^s + Q_{31}(\chi(x^s,u^s) - x^s) + Q_{32}(\chi(z^s,u^s) + \bar{Q}(g(x^s) - g(z^s)) - z^s) + Q_{33}(y^* - g(x^s)) \quad (VI.19.c)$$

The system state is not available for control and consequently Q_{21} and Q_{31} can only be equal to 0. In addition, the system evolution does not explicitly depend on the model state and as a result $Q_{12} = 0$. Furthermore, the model evolution should match the system evolution when $z^s = x^s$, i.e.

$$z^{s+1} = \chi(z^s,u^s) \quad (VI.20)$$

Setting $x^s = z^s$ in (VI.19.b) and equating the right hand sides of (VI.19.b) and (VI.20) :

$$\chi(z^s,u^s) = z^s + Q_{22}(\chi(z^s,u^s) + \bar{Q}(g(z^s) - g(z^s)) - z^s) + Q_{23}(y^* - g(z^s)) \quad (VI.21)$$

which upon rearrangement yields :

$$\begin{pmatrix} Q_{22} - I & Q_{23} \end{pmatrix} \begin{pmatrix} \chi(z^s,u^s) - z^s \\ y^* - g(z^s) \end{pmatrix} = 0, \quad \forall \begin{pmatrix} \chi(z^s,u^s) - z^s \\ y^* - g(z^s) \end{pmatrix} \quad (VI.22)$$

(VI.22) implies that $Q_{22} = I$ and $Q_{23} = 0$.

Finally consistency requirement restricts the set of admissible Q's even further. Following the argumentation in 2.1, it can be established that $Q_{11} = I$ and $Q_{13} = 0$.

Summing up, a consisten operator will be of the form

$$Q = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & Q_{32} & Q_{33} \end{pmatrix} \begin{matrix} n \\ n \\ m \end{matrix} \quad \text{(columns: } n, n, m\text{)} \quad \text{with } Q_{33} \text{ non-singular}$$

where the non-singularity requirement is imposed by the condition that a consistent Q is invertible.

Control Algorithm

For $Q \in \mathbf{Q}_c$, (VI.19) become :

$$x^{s+1} = \chi(x^s,u^s) \quad (VI.23.a)$$

$$z^{s+1} = \chi(z^s,u^s) + \bar{Q}(g(x^s) - g(z^s)) \quad (VI.23.b)$$

$$u^{s+1} = u^s + Q_{32}(\chi(z^s,u^s) + \bar{Q}(g(x^s) - g(z^s)) - z^s) + Q_{33}(y^* - g(x^s)) \quad (VI.23.c)$$

(VI.23.b) and (VI.23.c) comprise the control law. (VI.23.b) represents the model with output feedback through $\bar{Q}$. (VI.23.c) generates the control input by employing system output feedback through Q_{33} and model state feedback through Q_{32}. Finally, the first term of the (VI.23.c) right hand side provides for the integral action. A block diagram realization of the resulting closed-loop system is shown in fig. VI.2.

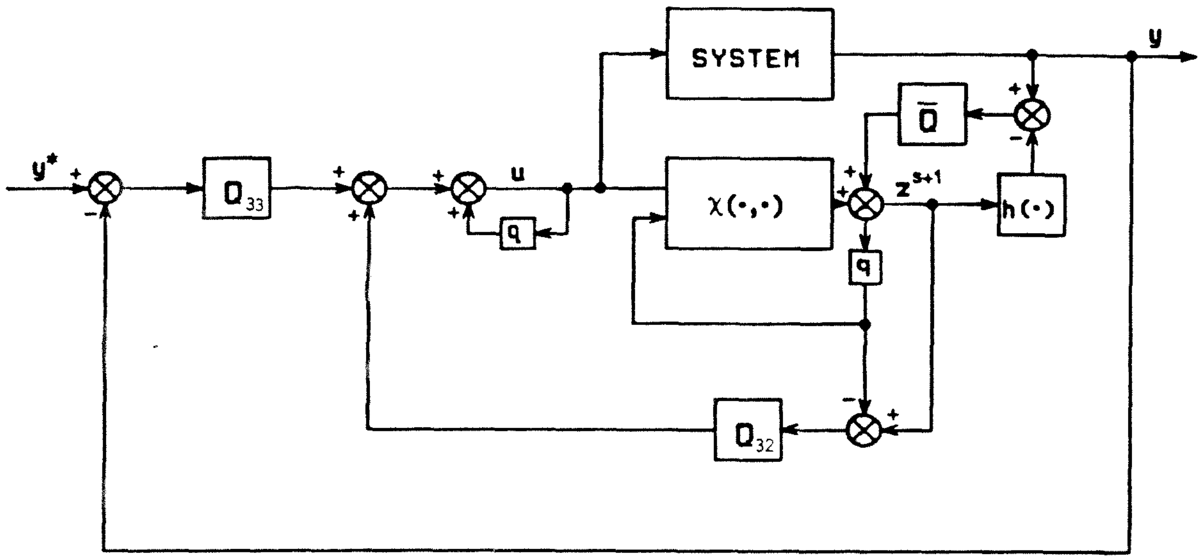

FIGURE VI.3 : *Block structure of the model reference control law (q: Backward shift - delay - operator).*

<u>The Design Problem</u>

Differentiating the right hand side of (VI.23) with respect to $x = x^s$, $z = z^s$ and $u = u^s$ and setting $\Phi(\cdot,u) = \dfrac{\partial\chi(\cdot,u)}{\partial(\cdot)}$, $\Gamma(\cdot,u) = \dfrac{\partial\chi(\cdot,u)}{\partial u}$, and $C(\cdot) = \dfrac{\partial g(\cdot)}{\partial(\cdot)}$, the design problem (VI.15) translates to :

Model Reference Design Problem

$$\min_{\bar{Q},Q_c} \sup_{(x,z,u)\in\mathcal{C}((x^*,z^*,u^*),r)} \left\| \begin{matrix} \Phi(x,u) & 0 & \Gamma(x,u) \\ \bar{Q}C(x) & \Phi(z,u)-\bar{Q}C(z) & \Gamma(z,u) \\ (Q_{32}\bar{Q}-Q_{33})C(x) & Q_{32}(\Phi(z,u)-\bar{Q}C(z)-I) & I+Q_{32}\Gamma(z,u) \end{matrix} \right\|$$

<u>Implications for Linear Systems</u>

If the system is linear the entries in the norm are not dependent on (x,z,u). Using the results of section II.3 then, the problem becomes :

Linear Model Reference Design Problem

$$\min_{\bar{Q},Q_c} \rho \begin{pmatrix} e^{AT} & 0 & (e^{AT}-I)A^{-1}B \\ \bar{Q}C & e^{AT}-\bar{Q}C & (e^{AT}-I)A^{-1}B \\ (Q_{32}\bar{Q}-Q_{33})C & Q_{32}(e^{AT}-\bar{Q}C-I) & I+Q_{32}(e^{AT}-I)A^{-1}B \end{pmatrix}$$

Again a pole-placement problem surfaces, with the objective of shifting the closed-loop poles to the origin, resulting in a stabilizing state dead-beat controller.

Example VI.3 : Consider the linear open-loop system of example VI.2. Let

$$Q_{21} = \begin{pmatrix} \alpha & \beta \end{pmatrix}, Q_{22} = \begin{pmatrix} \gamma \end{pmatrix} \quad \text{and} \quad \bar{Q} = \begin{pmatrix} \delta \\ \varepsilon \end{pmatrix},$$

Then, the linear design problem becomes

$$\min \rho \begin{pmatrix} 1.47 & -0.367 & 0 & 0 & 0.122 \\ .122 & 0.983 & 0 & 0 & 5.7\times 10^{-3} \\ 0 & \delta & 1.47 & -0.367-\delta & 0.122 \\ 0 & \varepsilon & 0.122 & 0.983-\varepsilon & 5.7\times 10^{-3} \\ 0 & \alpha\delta+\beta\varepsilon-\gamma & 0.47\alpha+0.122\beta & -0.367\alpha-\alpha\delta-0.017\beta-\beta\varepsilon & 1+0.122\alpha+5.7\times 10^{-3}\beta \end{pmatrix}$$

The coefficients of the characteristic polynomial of the design problem matrix are equated to zero and the resulting system of equations is solved to obtain

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \varepsilon \end{pmatrix} = \begin{pmatrix} -18.7 \\ -203.7 \\ 81.7 \\ 17.35 \\ 2.46 \end{pmatrix}$$

These values are substituted in (VI.23) yielding the model evolution equations

$$z_1^{s+1} = 1.47z_1^s - 17.72z_2^s + 0.122u^s + 17.35y^s$$
$$z_2^{s+1} = 0.122z_1^s - 1.47z_2^s + 5.7\times 10^{-3}u^s + 2.46y^s$$

and the control law

$$u^{s+1} = u^s - 18.7(z_1^{s+1} - z_1^s) - 203.7(z_2^{s+1} - z_2^s) + 81.7(y^* - y^s)$$

The system response a step up setpoint command (y^* from 0.0 to 1.0) is studied, when the initial model state is significantly different from the system state. The model state "captures" the system state in two sampling intervals (fig. VI.4), after which the system settles in three additional intervals (fig. VI.5).

Suboptimal Control Laws

Suboptimal control laws are sought, to reduce the effort associated with solving the optimization design problem :

i. A two step approach, suggested by classical synthesis, is discussed next.

Step 1: Solve the state feedback design problem of section 2. and compute matrices $Q_{21} = A$ and $Q_{22} = B$.

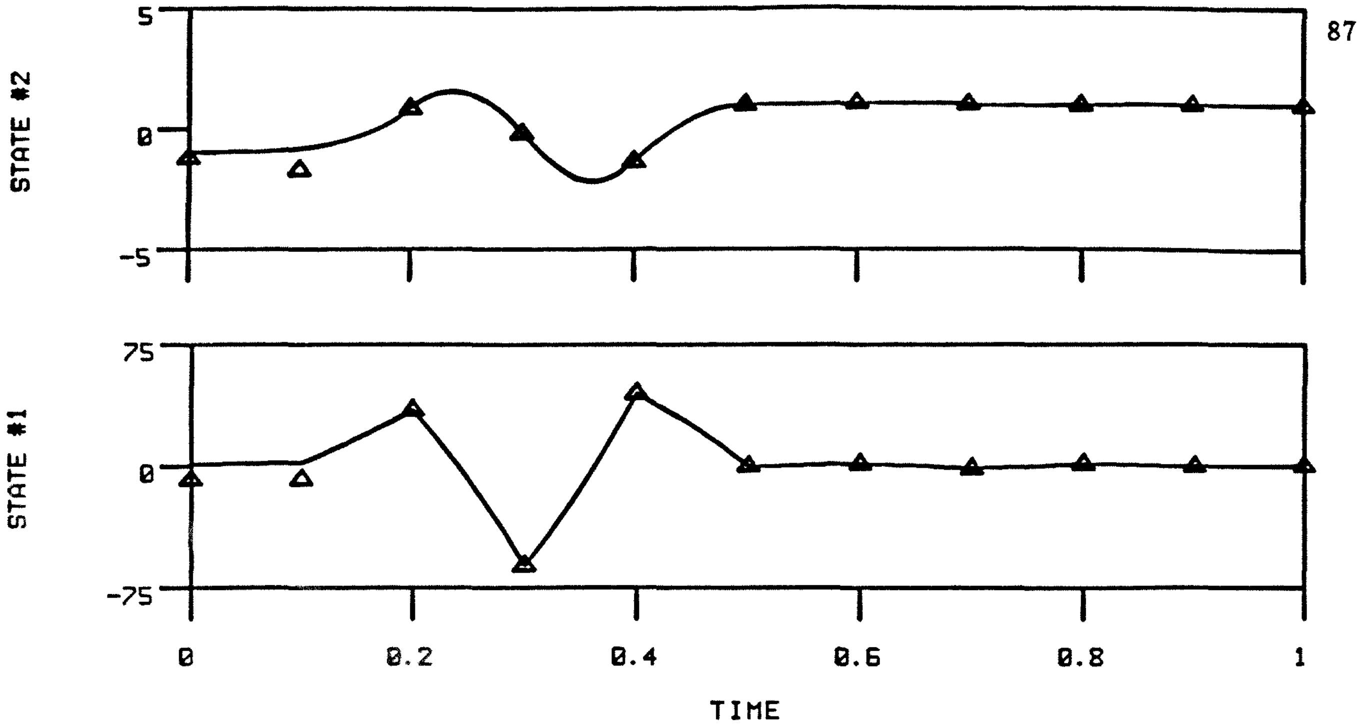

FIGURE VI.4 : *Contraction controller design for linear unstable system, model reference case. State convergence. (——— : System state, △: Model state).*

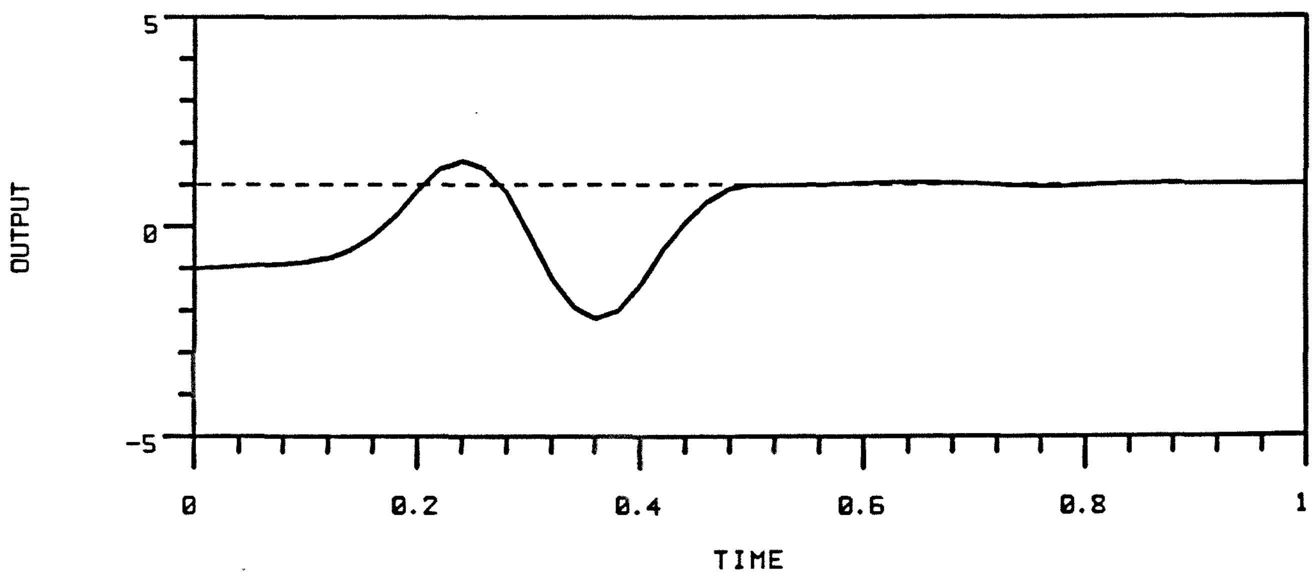

FIGURE VI.5 : *Contraction controller design for linear unstable system, model reference case. Setpoint tracking.(– – – – : Setpoint command, ——— : Output.)*

Step 2: Set $Q_{32} = A$ and $Q_{33} = B$ and then solve the model reference design problem with respect to $\bar{Q}$:

$$\min_{\bar{Q}\in\mathbf{R}^{n\times m}} \sup_{(x,z,u)\in\bar{U}((x^*,z^*,u^*),r)} \left\| \begin{matrix} \Phi(x,u) & 0 & \Gamma(x,u) \\ \bar{Q}C(x) & \Phi(z,u)-\bar{Q}C(z) & \Gamma(z,u) \\ (A\bar{Q}-B)C(x) & A(\Phi(z,u)-\bar{Q}C(z)-I) & I+A\Gamma(z,u) \end{matrix} \right\|$$

ii. (Local linearization design) An equilibrium state is located and the system is linearized around this state. The linear model reference design problem is solved to compute $\bar{Q}$, Q_{32} and Q_{33}. They are substituted in (VI.23) in turn to generate the control law. Theorem V.5 is used to study the stability of the resulting closed-loop system. An induced norm that approximates the closed-loop spectral radius at the equilibrium state should be used for the analysis.

2.3 The Case of Parametric Uncertainty

The discussion parallels that of the previous paragraph 2.2, with the main addition that the contraction conditions should hold for the set of parameters.

Assumptions The system is known up to a set of parameters $p \in I \subset \mathbf{R}^p$

Operator Equation

The control objective is to drive the system and model to steady state with the output at y^*.

$$\begin{aligned} x &= \chi(x,u;p) && \text{(System at steady state)} \\ z &= \chi(z,u;p_0)+\bar{Q}(g(x;p)-g(z;p_0)) && \text{(Model matching of system evolution)} \\ y^* &= g(x;p) && \text{(Output at desired value } y^* \text{)} \end{aligned} \tag{VI.24}$$

Successive Substitution Form

$$\begin{pmatrix} x^{s+1} \\ z^{s+1} \\ u^{s+1} \end{pmatrix} = \begin{pmatrix} x^s \\ z^s \\ u^s \end{pmatrix} + Q \begin{pmatrix} \chi(x^s,u^s;p)-x^s \\ \chi(z^s,u^s;p_0)+\bar{Q}(g(x^s;p)-g(z^s;p_0))-z^s \\ y^*-g(x;p) \end{pmatrix} \tag{VI.25}$$

The set of consistent Q's

The same conditions in 2.2 on candidate Q's apply in this case and as a result the set $\mathbf{Q}_c$ is the same.

Control Algorithm

For $Q \in \mathbf{Q}_c$, (VI.25) become :

$$\begin{aligned} x^{s+1} &= \chi(x^s,u^s;p) \\ z^{s+1} &= \chi(z^s,u^s;p_0)+\bar{Q}(g(x^s;p)-g(z^s;p_0)) \\ u^{s+1} &= u^s+Q_{32}(\chi(z^s,u^s;p_0)+\bar{Q}(g(x^s;p)-g(z^s;p_0))-z^s)+Q_{33}(y^*-g(x^s;p)) \end{aligned} \tag{VI.26}$$

The interpretation and block diagram realization in this case is identical to the model reference case (fig. VI.2).

The Design Problem

Differentiating the right hand side of (VI.26) and setting $\Phi(\cdot,u;p) = \dfrac{\partial\chi(\cdot,u;p)}{\partial(\cdot)}$, $\Gamma(\cdot,u;p) = \dfrac{\partial\chi(\cdot,u;p)}{\partial u}$, $C(\cdot,p) = \dfrac{\partial g(\cdot,p)}{\partial(\cdot)}$, and $v=(x,z,u)$, the design problem (VI.5) translates to :

Parametric Uncertainty Design Problem

$$\min_{\bar{Q},Q_c} \sup_{\substack{v\in \bar{U}(v^*,r)\\ p\in I}} \left\| \begin{matrix} \Phi(x,u;p) & 0 & \Gamma(x,u;p) \\ \bar{Q}C(x;p) & \Phi(z,u;p_0)-\bar{Q}C(z;p_0) & \Gamma(z,u;p_0) \\ (Q_{32}\bar{Q}-Q_{33})C(x;p) & Q_{32}(\Phi(z,u;p_0)-\bar{Q}C(z;p_0)-I) & I+Q_{32}\Gamma(z,u;p_0) \end{matrix} \right\|$$

Implications for Linear Systems

If the system is linear the entries in the norm are not dependent on (x,z,u). Using the results of section II.3 then, the problem becomes (after setting $A_p = A(p)$ and $A_0 = A(p_0)$ etc.) :

Linear Parametric Uncertainty Design

$$\min_{\bar{Q},Q_c} \sup_{p\in I} \rho \left\{ \begin{matrix} e^{A_pT} & 0 & (e^{A_pT}-I)A_p^{-1}B_p \\ \bar{Q}C_p & e^{A_0T}-\bar{Q}C_0 & (e^{A_0T}-I)A_0^{-1}B_0 \\ (Q_{32}\bar{Q}-Q_{33})C_p & Q_{32}(e^{A_0T}-\bar{Q}C_0-I) & I+Q_{32}(e^{A_0T}-I)A_0^{-1}B_0 \end{matrix} \right\}$$

Suboptimal Control Laws

(Nominal local linearization design) Assume that $p = p_0$ and subsequently solve the resulting model reference problem . Theorem V.8 is then used for stability analysis.

3. NONLINEAR CONTROL LAWS

An inherent disadvantage of the successive substitution approach is that it does not suggest any computational procedure for obtaining nonlinear operators Q that affect algorithm convergence properties in a desirable manner. As a result, alternative operator equation solution methods should be considered for nonlinear control law synthesis. This is done in the next Chapter by employing the method of Newton.

In the context of successive substitution methods, only the case of linear operators can be dealt with mathematical rigor. Nonlinear control laws can be obtained in the state-feedback case by extrapolating the corresponding linear control laws. The nonlinear operator Q in (VI.13) if substituted in (VI.12) generates the nonlinear control law :

$$\begin{aligned} x^{s+1} &= \chi(x^s,u^s) \\ u^{s+1} &= u^s + \left[C(x)(\Phi(x,u)-I)^{-1}\Gamma(x,u)\right]^{-1} C(x)(\Phi(x,u)-I)^{-1}(x^s-\chi(x^s,u^s)) \\ &\quad + \left[C(x)(\Phi(x,u)-I)^{-1}\Gamma(x,u)\right]^{-1}(g(x^s)-y^*) \end{aligned} \tag{VI.27}$$

The stability of the closed-loop system is determined by Theorem V.2.

Example VI.4 : The Continuous Stirred Tank Reactor of example II.1 will be used as a benchmark test for the proposed nonlinear control laws. Introducing appropriate values for the coefficients, the reactor system is described by the following set of equations:

$$\frac{\partial x_1}{\partial t} = 1 - x_1 - 3 \times 10^5 \exp(-5000/x_3)x_1 + 6 \times 10^7 \exp(-7000/x_3)x_2$$
$$\frac{\partial x_2}{\partial t} = -x_2 + 3 \times 10^5 \exp(-5000/x_3)x_1 - 6 \times 10^7 \exp(-7000/x_3)x_2$$
$$\frac{\partial x_3}{\partial t} = T_i - x_3 + 0.05[3 \times 10^5 \exp(-5000/x_3)x_1 - 6 \times 10^7 \exp(-7000/x_3)x_2]$$
$$y = x_2/(x_1 + x_2)$$

where t is in min, x_1, x_2 in $mol \cdot l^{-1}$, x_3 in $K/100$ and T_i, the inlet temperature, in $K/100$. The temperature - conversion equilibrium diagram of the reactor (fig. V.1, example V.4) shows a well defined conversion maximum. The objective is to safely operate the reactor close to the maximum conversion point, using the feed stream temperature as the control input.
It has been shown (Economou and Morari 1985), that as a result of the nonlinearly varying gain of the reactor, the control problem is inherently difficult to address with a linear controller: integral action controllers are unstable, non-integral action linear controllers lead to unacceptable offsets from the desired operating point.
Fig. VI.6 shows the system response to a step-up setpoint command to maximum conversion, under control law (II.27) (†). A sampling time $T = 1.0$ min is used and the initial reactor state is at the left of the maximum conversion point.
In fig. VI.7 the situation is repeated, only this time the initial reactor state is at the right of the maximum conversion point.

4. CONCLUDING REMARKS

4.1 Linear vs. Nonlinear Controller Design

The design problems for linear and nonlinear systems are conceptually closely related, in the sense that they derive from the same principles. The respective computational aspects however are fundamentally different for the following reasons :

i. For linear systems the derivative of the system operator is independent of the state. As a result, the associated design problem is a simple minimization search. On the other hand, for nonlinear systems the derivative operator is a function of the system state, resulting in a considerably more complicated min-max optimization search.

ii. For linear systems a unique measure (ρ) of the system operator is optimized; whereas for nonlinear systems an additional computational burden is to find the least conservative operator norm – out of an infinite multitude – to optimize.

(†) Program DDASAC (Caracotsios and Stewart 1984) was used for the on-line computation of the quantities involved in the control law computations.

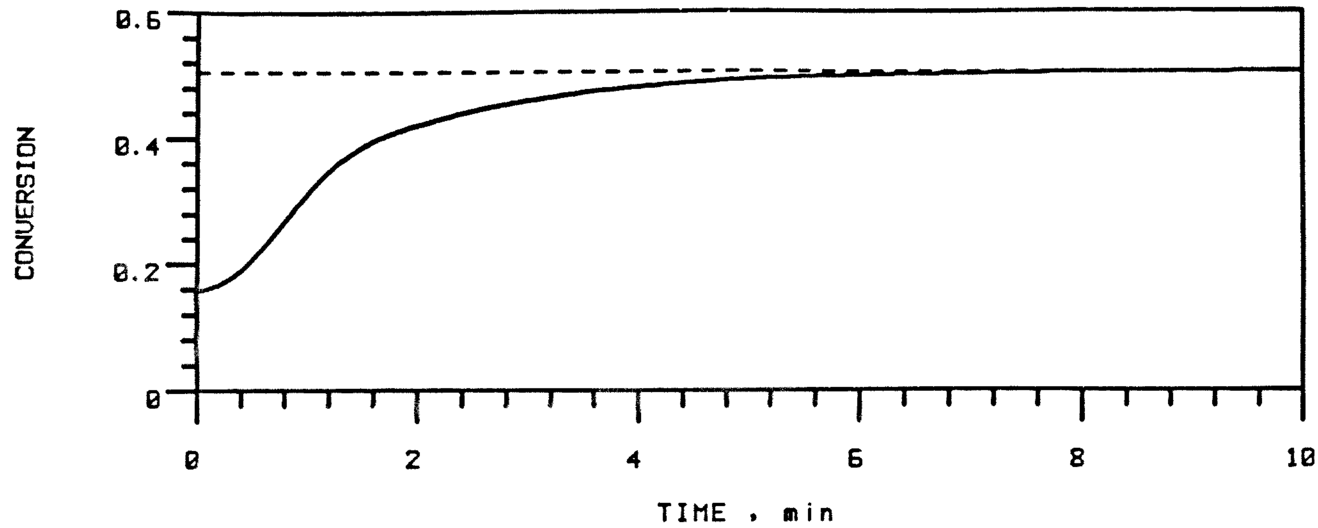

FIGURE VI.6 : *System response to a step-up setpoint command under control law* (II.27) *. Sampling time* $T = 1.0$ *min. Initial reactor state is at the left of the maximum conversion point* ($x_1^0 = 0.84,\ x_2^0 = 0.16,\ x_3^0 = 3.51$). *(– – – – : Setpoint command, ——— : System output, i.e. R concentration).*

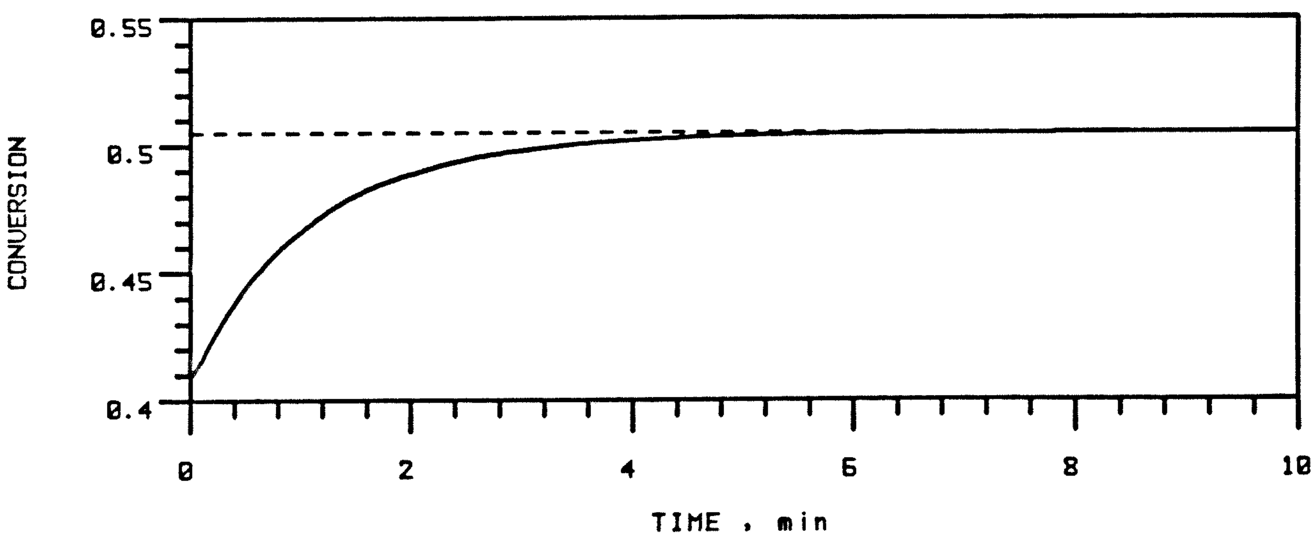

FIGURE VI.7 : *System response to a step-up setpoint command under control law* (II.27) *. Sampling time* $T = 1.0$ *min. Initial reactor state is at the right of the maximum conversion point* ($x_1^0 = 0.59,\ x_2^0 = 0.41,\ x_3^0 = 5.02$). *(– – – – : Setpoint command, ——— : System output, i.e. R concentration).*

iii. For linear systems there is only one equilibrium state in general. This is not the case in nonlinear systems, which can have many equilibrium states; for each equilibrium state a separate stability analysis is necessary.

iv. For linear systems the region of stability (attraction) is the space itself and as a result once stability has been established in some region, it is automatically established for any other region, the space itself included. For nonlinear systems, stability regions can be finite subspaces of the space; at the same time stability in some region does not have any implications for stability in any other region.

The implied computational complexity is inherently tied to nonlinear systems controller design. It is the price to be paid for extending the design problem to a much broader class of systems than the class of linear systems.

4.2 Implementation Aspects

The bulk of computations associated with contraction principle controllers is performed off-line in order to solve the corresponding design problems. On-line implementation is not much different or more complicated than a generic linear control law implementation. At each sampling time (eg. $t = t^s$) the following steps are performed for the implementation of control law (VI.23) :

Step 1 : Measure the system output $y^s = g(x^s)$ at $t = t^s$.
Step 2 : Compute the model output $y^s_M = g(z^s)$ at $t = t^s$.
Step 3 : Simulate the model equations (II.1), with $x(t^s) = z^s$, from $t = t^s$ to $t = t^{s+1}$ to obtain the quantity $\chi(z^s, u^s)$.

At this point all the quantities involved in the control law calculations have been computed. Then,

Step 4 : Update the model state by (VI.23.b).
Step 5 : Compute the input to be injected to the system from $t = t^{s+1}$ to $t = t^{s+2}$ by (VI.23.c).
Step 6 : Set $s = s + 1$. Go to Step 1.

4.3. Integral action

Assuming stability, every control law introduced in this Chapter will have no offset at equilibrium. This is merely a consequence of the way the control objective was expressed as an operator equation. To understand why, note that at equilibrium (VI.2) becomes:

$$
\begin{aligned}
& v_{eq} = v_{eq} + QP(v_{eq}) \\
& \Leftrightarrow QP(v_{eq}) = 0 \qquad \text{(and, since Q is invertible)} \\
& \Leftrightarrow P(v_{eq}) = 0 \qquad \text{(VI.28)}
\end{aligned}
$$

In deriving control laws, the objective expressed by $P(v) = 0$ was to drive a system to steady-state with the output at a desired value. Then (VI.28) shows that, if the system is stable, the control objective is attained at steady-state, i.e. the system output will be at the desired value, with no offset.

CHAPTER VII

NEWTON SYNTHESIS

1. INTRODUCTION

In Chapter VI it was pointed out that the successive substitution approach is not efficient in deriving nonlinear control laws. Alternative operator equation solution methods should be investigated. Nonlinear control synthesis is pursued in this Chapter by employing the Newton method.

The method and its properties have been investigated in detail (Bartle 1955, Kolmogorov and Fomin 1957, Kantorović and Akilow 1964, Curtain and Pritchard 1977, Rall 1979). It has been well established that it is computationally efficient, with fast convergence rates; however it is not very reliable, in the sense that it produces smaller regions of convergence than alternative methods do (Steepest Descent, Successive Substitution etc.). Significant advances in developing modifications of the Newton method with strong convergence properties have been reported (Underrelaxed Damped method of Stoer 1972,1973 and Deufelhard 1974, Ball Newton method of Nickel 1979).

These considerations aside, the Newton method is used here as a prototype for nonlinear controller synthesis. Alternative methods can then be patterned along the proposed synthesis guidelines.

It is shown in Appendix VII that the Newton method does not yield consistent control laws. Hybrid Newton algorithms are constructed in section 2 to account for this situation. Then in sections 2 and 3 pseudo-Newton control laws are developed by considering different operator equation formulations of the control problem. Table VII.1 summarizes the proposed laws, the respective variant control laws, as well as the paragraph they appear.

The structure of the respective sections is common: the operator equation is stated first. Then, in the standard procedure of deriving the Newton method, control laws are obtained by truncating the Taylor series expansion of the operator equation; variants of the control laws follow in the light of the relaxed and simplified Newton methods defined in section III.2; stability conditions for every control law are developed, based either on the hybrid Newton convergence Theorems III.6 and 7, or Theorem V.2 and its corollaries; finally the interpretation, properties and implementation aspects of the control laws are discussed.

Exact modelling and complete access to the state vector will be assumed throughout the

NEWTON CONTROL LAWS REFERENCE TABLE			
TYPE	HYBRID NEWTON	PSEUDO-NEWTON A	PSEUDO-NEWTON B
Basic Law	I. (§2.2)	II. (§3.2)	III. (§4.2)
Relaxed Law (Control updates relaxed by λ)	I.R (§2.4)	II.R (§3.4)	III.R (§4.4)
Simplified Law (Derivative computed at fixed reference state)	I.S (§2.5)	II.S (§3.5)	III.S (§4.5)
Linear Law (Basic Law for Linear Systems)		II.L (§3.6)	III.L (§4.6)

TABLE VII.1 Newton Control Laws: Nomenclature

Chapter. They correspond to the assumptions of sections V.2 and VI.2. Relaxation of these assumption will not be treated at present. It is straightforward conceptually to extend the control laws of this chapter by solving the appropriate operator equations, in the same fashion that the laws of section VI.2 were extended to the model reference case of section VI.3 and the parametric uncertainty case of section VI.4.

The control laws will be expressed in a form $u^{s+1} = \psi(x^s, u^s)$ and $\psi = \psi(x, u)$ will denote this general functional dependence. To facilitate expressing the stability conditions, the notation of Table VII.2 will be used in the section.

2. HYBRID NEWTON CONTROL LAW

2.1 Operator Equation

The objective is to drive the system to its steady-state such that the output is equal to y^*.

$$\begin{aligned} x &= \chi(x, u) && \text{(System at Steady-State)} \\ g(\chi) &= y^* && \text{(Output at desired value } y^* \text{)} \end{aligned} \qquad \text{(VII.1)}$$

2.2 Control Law Derivation

(VII.1) is expanded in its Taylor series around a point $(x, u) = (x^s, u^s)$ yielding

$$\chi(x^{s+1}, u^{s+1}) - x^{s+1} = \chi(T; x^s, u^s) - x^s + \left[\frac{\partial \chi(T; x^s, u^s)}{\partial x^s} - I\right](x^{s+1} - x^s)$$
$$+ \frac{\partial \chi(T; x^s, u^s)}{\partial u^s}(u^{s+1} - u^s) + O\left(\left\| \begin{matrix} x^{s+1} - x^s \\ u^{s+1} - u^s \end{matrix} \right\|^2\right) \qquad \text{(VII.2.a)}$$

$$g(\chi(x^{s+1}, u^{s+1})) - y^* = g(\chi(T; x^s, u^s)) - y^*$$

CONTROL LAW NOTATION	
$\chi = \chi(x,u) = \chi(T;x,u)$	$\Phi_x = \frac{\partial \Phi}{\partial x}$
$\Phi = \Phi(x,u) = \frac{\partial \chi}{\partial x}$	$\Phi_u = \frac{\partial \Phi}{\partial u}$
$\Gamma = \Gamma(x,u) = \frac{\partial \chi}{\partial u}$	$\Gamma_x = \frac{\partial \Gamma}{\partial x}$
$C = C(x,u) = \left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right\|_{\varsigma = \chi}$	$\Gamma_u = \frac{\partial \Gamma}{\partial u}$
$C' = \left.\frac{\partial^2 g(\varsigma)}{\partial \varsigma^2}\right\|_{\varsigma = \chi}$	$\Phi_{ref} = \left.\frac{\partial \chi(\varsigma,\xi)}{\partial \varsigma}\right\|_{\substack{\varsigma = x_{ref} \\ \xi = u_{ref}}}$
$C_{ref} = \left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right\|_{\varsigma = \chi(x_{ref}, u_{ref})}$	$\Gamma_{ref} = \left.\frac{\partial \chi(\varsigma,\xi)}{\partial \xi}\right\|_{\substack{\varsigma = x_{ref} \\ \xi = u_{ref}}}$

TABLE VII.2 Newton Control Laws: Notation

$$
\begin{aligned}
&+ \left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right|_{\varsigma = \chi(T;x^s,u^s)} \frac{\partial \chi(T;x^s,u^s)}{\partial x^s}(x^{s+1} - x^s) \\
&+ \left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right|_{\varsigma = \chi(T;x^s,u^s)} \frac{\partial \chi(T;x^s,u^s)}{\partial u^s}(u^{s+1} - u^s) \\
&+ O\left(\left\| \begin{matrix} x^{s+1} - x^s \\ u^{s+1} - u^s \end{matrix} \right\|^2\right)
\end{aligned} \qquad \text{(VII.2.b)}
$$

Assuming that (x^{s+1}, u^{s+1}) is a solution of (VII.1)

$$
\begin{aligned}
\chi(x^{s+1}, u^{s+1}) - x^{s+1} &= 0 \\
g(\chi(x^{s+1}, u^{s+1})) - y^* &= 0
\end{aligned},
$$

truncating the terms of order two and higher and using the notation of section II.2, (VII.2) becomes

$$0 = \chi^s - x^s + (\Phi^s - I)(x^{s+1} - x^s) + \Gamma^s(u^{s+1} - u^s) \qquad \text{(VII.3.a)}$$

$$0 = g(\chi^s) - y^* + C^{s+1}\Phi^s(x^{s+1} - x^s) + C^{s+1}\Gamma^s(u^{s+1} - u^s) \qquad \text{(VII.3.b)}$$

(VII.3.a) is solved for $(x^{s+1} - x^s)$:

$$x^{s+1} - x^s = (\Phi^s - I)^{-1}[x^s - \chi^s - \Gamma^s(u^{s+1} - u^s)] \qquad \text{(VII.4)}$$

and then (VII.4) is substituted in (VII.3.b) to obtain u^{s+1} after some rearrangement

$$u^{s+1} = u^s + \left[C^{s+1}(\Phi^s - I)^{-1}\Gamma^s\right]^{-1} C^{s+1}\Phi^s(\Phi^s - I)^{-1}(x^s - \chi^s) + \left[C^{s+1}(\Phi^s - I)^{-1}\Gamma^s\right]^{-1}(y^{s+1} - y^*) \qquad \text{(VII.5)}$$

The Hybrid Newton algorithm is then constructed as follows : Consider any successive substitution algorithm (section VI.2) with $Q_{11} = I$, $Q_{12} = 0$, Q_{21} and Q_{22} arbitrary operators. The consistency of the resulting control law was shown in Section VI.2. At the s^{th} iteration this algorithm produces x_c^{s+1} and u_c^{s+1} (equation VI.13). At the same iteration the Newton algorithm produces x_N^{s+1} from (VII.4) and u_N^{s+1} from (VII.5) . Although in general $x_N^{s+1} \neq \chi^s$ – because the Newton algorithm is not consistent – it is always $x_c^{s+1} = \chi^s$ by construction of the consistent successive substitution algorithm. For the next iteration the hybrid algorithm selects

$$x^{s+1} = x_c^{s+1} (= \chi^s) \qquad u^{s+1} = u_N^{s+1} \qquad \text{(VII.6)}$$

By construction the hybrid algorithm is consistent. Summarizing, the control law is :

CONTROL LAW I.

HYBRID NEWTON

$$u^{s+1} = u^s + \left[C^{s+1}(\Phi^s - I)^{-1}\Gamma^s\right]^{-1} C^{s+1}\Phi^s(\Phi^s - I)^{-1}(x^s - \chi^s) + \left[C^{s+1}(\Phi^s - I)^{-1}\Gamma^s\right]^{-1}(y^{s+1} - y^*)$$

2.3 Stability Analysis

According to Theorem III.6 the conditions for stability are:

i. There exist Q_{21}, Q_{22} such that

$$\left\| \begin{matrix} \Phi & \Gamma \\ Q_{21}(\Phi - I) - Q_{22}C & I + Q_{21}\Gamma \end{matrix} \right\| < 1, \quad \forall (x,u) \in \bar{U}((x^*,u^*),r)$$

ii.
$$\sup_{x,u \in \bar{U}((x^*,u^*),r)} \left\| \begin{matrix} \Phi - I & \Gamma \\ C\Phi & C\Gamma \end{matrix} \right\| \sup_{x,u \in \bar{U}((x^*,u^*),r)} \|\{S\}\| \; r < 1$$

where $S_{111} = \{\Phi_x\}$, $S_{121} = \{\Phi_u\}$, $S_{112} = \{\Gamma_x\}$, $S_{122} = \{\Gamma_u\}$, $S_{211} = \{C'\} \bullet \Phi \circ \Phi + C \bullet \{\Phi_x\}$, $S_{221} = \{C'\} \bullet \Phi \circ \Gamma + C \bullet \{\Phi_u\}$, $S_{212} = \{C'\} \bullet \Phi \circ \Gamma + C \bullet \{\Gamma_x\}$ and $S_{222} = \{C'\} \bullet \Phi \circ \Gamma + C \bullet \{\Gamma_u\}$.

Example VII.1 : The Continuous Stirred Tank Reactor of example VI.1 is used to illustrate application of control law I.

Fig. VI.1 shows the system response to a step-up setpoint command to maximum conversion, under control law I., when the initial reactor state is at the left of the maximum conversion point.

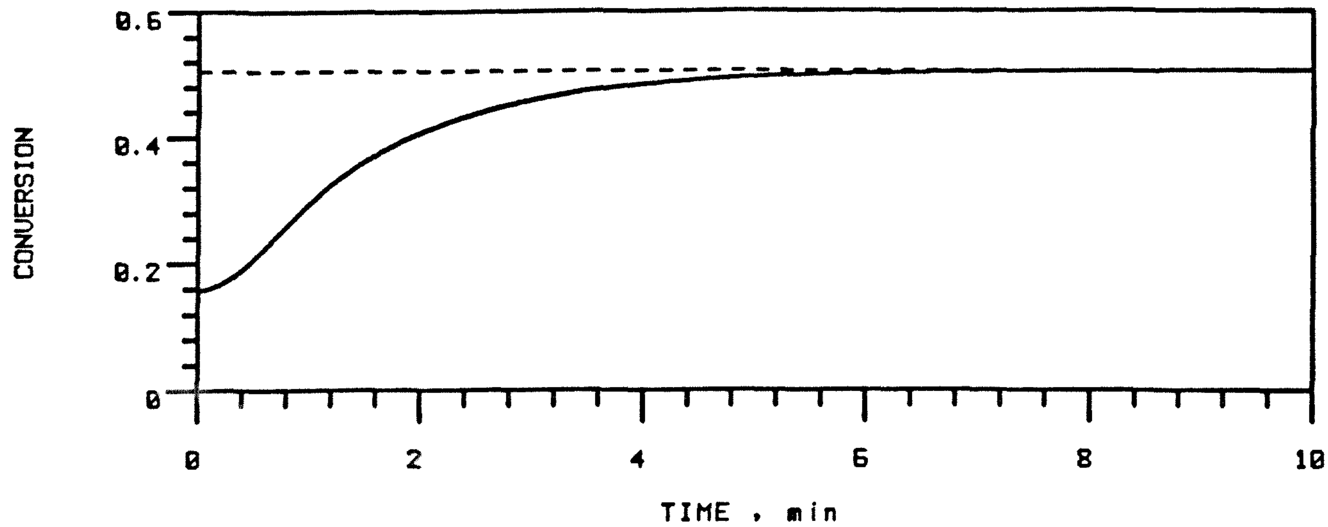

FIGURE VII.1 : *System response to a step-up setpoint command under control law I. Sampling time $T = 1.0$ min. Initial reactor state is at the left of the maximum conversion point ($x_1^0 = 0.84$, $x_2^0 = 0.16$, $x_3^0 = 3.51$). (– – –– : Setpoint command, ——— : System output.)*

In fig. VI.2 the situation is repeated, only this time the initial reactor state is at the right of the maximum conversion point.

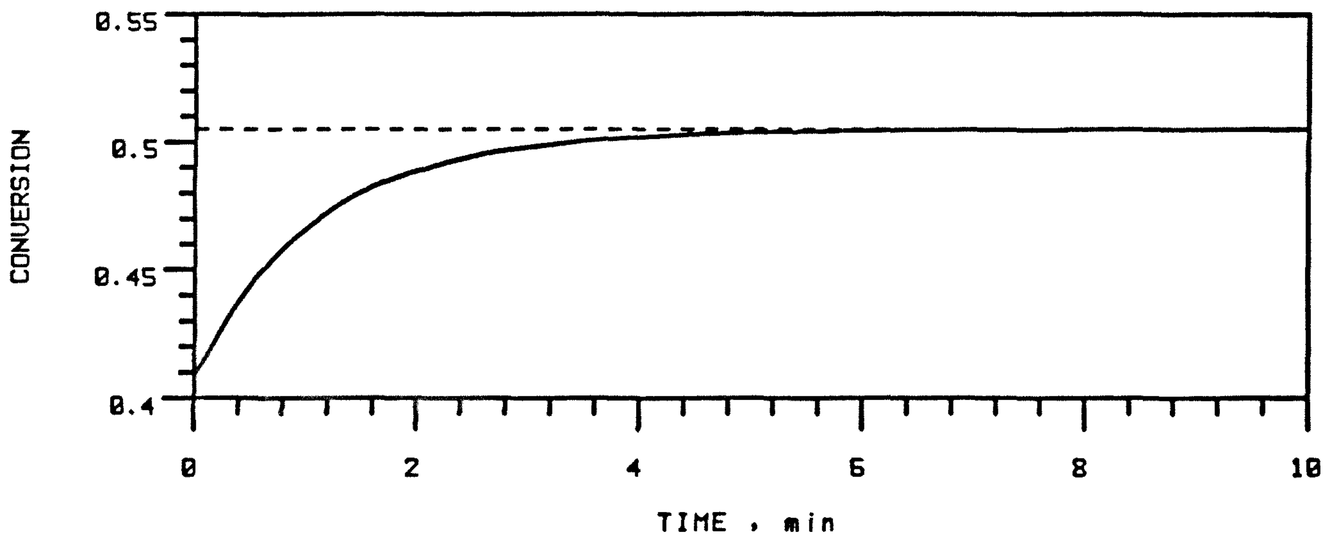

FIGURE VII.2 : *System response to a step-up setpoint command under control law I. Sampling time $T = 1.0$ min. Initial reactor state is at the right of the maximum conversion point ($x_1^0 = 0.59$, $x_2^0 = 0.41$, $x_3^0 = 5.02$). (– – –– : Setpoint command, ——— : System output.)*

2.4 Relaxed Law

Relaxing the updates in law I. by a factor λ, the relaxed law is obtained:

CONTROL LAW I.R

RELAXED HYBRID NEWTON

$$u^{s+1} = u^s + \lambda \left[C^{s+1}(\Phi^s - I)^{-1}\Gamma^s\right]^{-1} C^{s+1}\Phi^s(\Phi^s - I)^{-1}(x^s - \chi^s)$$
$$+\lambda \left[C^{s+1}(\Phi^s - I)^{-1}\Gamma^s\right]^{-1} (y^{s+1} - y^*)$$

Example VII.2 : The effect of the relaxation factor is illustrated in fig. VII.3. The system response to a step-up setpoint command under control law I.R and 3 different relaxation factors is shown, when the initial reactor state is at the left of the maximum conversion point.

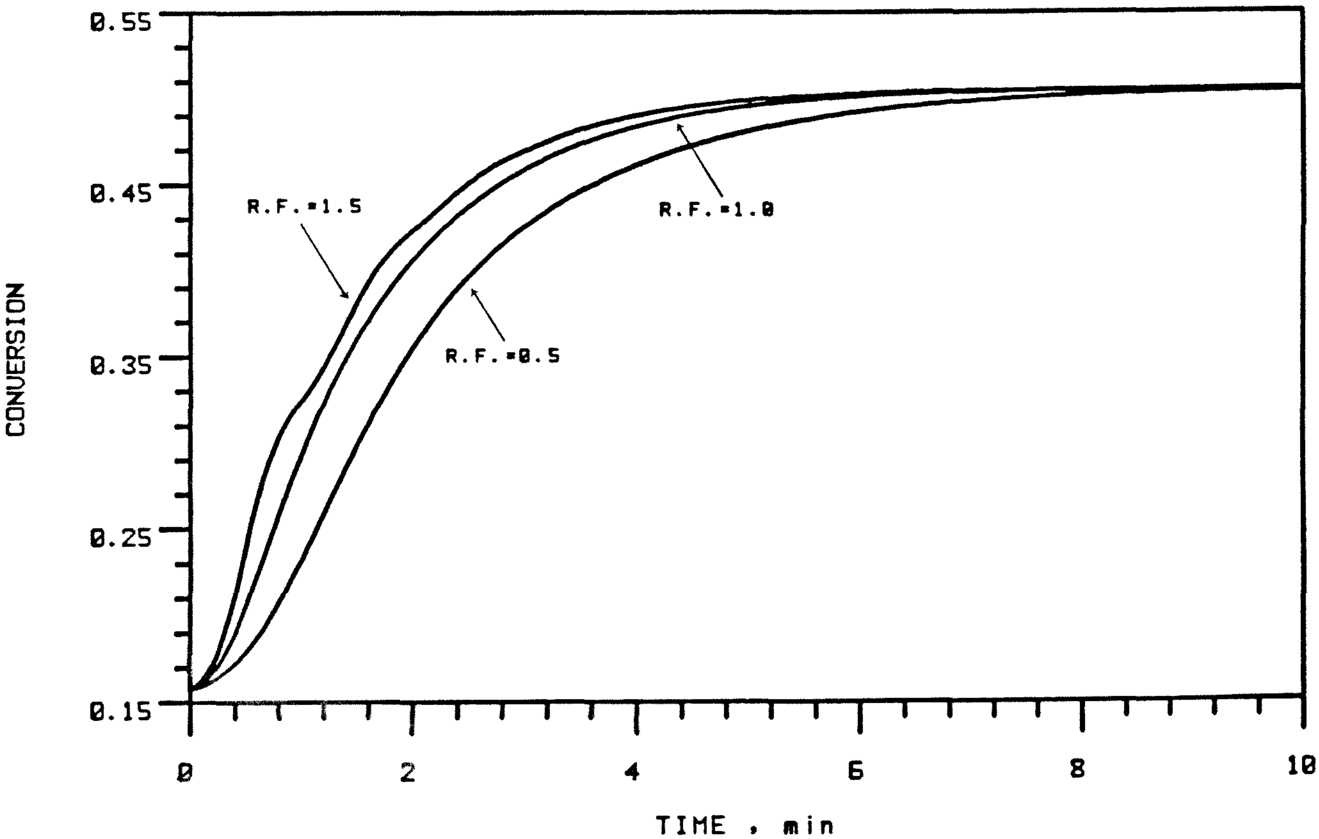

FIGURE VII.3 : *System response to a step-up setpoint command under control law I.R. Relaxation factors: 0.5, 1.0 and 1.5. Sampling time* $T = 1.0$ *min. Initial reactor state is at the left of the maximum conversion point (*$x_1^0 = 0.84$, $x_2^0 = 0.16$, $x_3^0 = 3.51$*). (– – –– : Setpoint command, ——— : System output.)*

2.5 Simplified Law

A reference point (x_{ref}, u_{ref}) is selected (†) and the appropriate quantities Φ_{ref}, Γ_{ref} and C_{ref} are computed. The simplified control law is the linear control law obtained by replacing Φ^s, Γ^s and C^{s+1} in (VII.5) (which are matrix functions of x^s and u^s), by the constant matrices Φ_{ref}, Γ_{ref} and C_{ref} respectively.

CONTROL LAW I.S

SIMPLIFIED HYBRID NEWTON

$$u^{s+1} = u^s + \left[C_{ref}(\Phi_{ref} - I)^{-1}\Gamma_{ref}\right]^{-1} C_{ref}\Phi_{ref}(\Phi_{ref} - I)^{-1}(x^s - \chi^s) + \left[C_{ref}(\Phi_{ref} - I)^{-1}\Gamma_{ref}\right]^{-1}(y^{s+1} - y^*)$$

According to Theorem III.7 if $(x_{ref}, u_{ref}) = (x^*, u^*)$, the conditions for stability are:

i. There exist Q_{21}, Q_{22} such that

$$\left\| \begin{matrix} \Phi & \Gamma \\ Q_{21}(\Phi - I) - Q_{22}C & I + Q_{21}\Gamma \end{matrix} \right\| < 1, \quad \forall (x, u) \in \bar{U}((x^*, u^*), r)$$

ii.
$$\left\| \begin{matrix} \Phi_{ref} - I & \Gamma_{ref} \\ C_{ref}\Phi_{ref} & C_{ref}\Gamma_{ref} \end{matrix} \right\| \sup_{x,u \in \bar{U}((x^*,u^*),r)} \|\{S\}\| \quad r < 1$$

where S as in §2.3.

2.6 Interpretation

The objective of the Hybrid Newton algorithm is to find the equilibrium state and input of the system. The quantity x^{s+1} in (VII.4) is the first order approximation of the equilibrium state. As a result the hybrid algorithm "looks" ahead at $t = \infty$ when it computes the next input to the system. Consequently, the control action will be very conservative, and of practical value only for relatively fast systems where this infinite horizon assumption is reasonable.

The control laws developed later in the Chapter display the opposite behavior, by "looking" only one step ahead.

Comparison with control law (VI.27) of Chapter VI shows that control law I. is identical (a technical difference is that in (VI.27) output feedback at t^s is used instead at t^{s+1} as is the case here). It may be recalled that this control law was obtained by selecting nonlinear operators Q that make the lower blocks of the state feedback stability condition equal to zero.

(†) There are different alternatives for reference selection. The reference point may be the equilibrium state, or any other state around it. Theorem III.7 supports the stability analysis when the equilibrium state is used.

2.7 Implementation

Matrices Q_{21} and Q_{22} need not be explicitly computed for the implementation of the hybrid Newton law, because they do not enter the control law expression. Their existence only suffices for off-line stability analysis.

Computations for the implementation of control law I. are carried out during the s^{th} sampling interval, in the following succession.

Step 1 : The model equations (II.1) are integrated forward from t^s to t^{s+1} to obtain Φ^s and Γ^s (Section II.3).

Step 2 : The quantities C^{s+1}, $\left[C^{s+1}(\Phi^s - I)^{-1}\Gamma^s\right]^{-1}$ and $C^{s+1}\Phi^s(\Phi^s - I)^{-1}$ are computed.

Step 3 : The system output y^{s+1} and state χ^s at $t = t^{s+1}$ are measured.

Step 4 : The new input to be injected into the system at $t = t^{s+1}$ is given by Control Law I.

2.8 Asymptotic Behavior

The Hybrid Newton control law I. (and its derivative laws) display some interesting behavior for large sampling times. It is argued in Appendix VII that for $T \longrightarrow \infty$, it becomes a standard Newton algorithm for the solution of the algebraic equations that describe the steady state of the system.

Specifically, let χ^s be the equilibrium state of the open-loop system that corresponds to some input u^s. Then the control law becomes for $T \longrightarrow \infty$:

$$u^{s+1} = u^s - \left[\left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right|_{\varsigma = \chi^s} \left[\left.\frac{\partial f(\varsigma,\xi)}{\partial \varsigma}\right|_{\substack{\varsigma = \chi^s \\ \xi = u^s}} \right]^{-1} \left.\frac{\partial f(\varsigma,\xi)}{\partial \xi}\right|_{\substack{\varsigma = \chi^s \\ \xi = u^s}} \right]^{-1} (g(\chi^s) - y^*)$$

This is the same expression used for the solution of the algebraic equation $g(x) = y^*$, with x an implicit function of u, given by $f(x, u) = 0$.

The implication is that the control law stability for large sampling times can be studied in terms of the convergence properties of a Newton iteration for the solution of a system of algebraic equations.

3. PSEUDO NEWTON CONTROL LAW A

3.1 Operator Equation

The operator equation is discrete this time:

$$\begin{aligned} x^{s+1} &= \chi(x^s, u^s) && \text{(System evolution constraint)} \\ g(x^{s+1}) &= y^* && \text{(System output at } y^* \text{ in one step)} \end{aligned} \qquad \text{(VII.7)}$$

3.2 Control Law Derivation

(VII.7) is expanded in its Taylor series around (x^s, u^s):

$$g(\chi(x^{s+1}, u^{s+1})) - y^* = \; g(\chi(x^s, u^s)) - y^*$$

$$+ \left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right|_{\varsigma = \chi(x^s, u^s)} \frac{\partial \chi(x^s, u^s)}{\partial x^s}(x^{s+1} - x^s)$$
$$+ \left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right|_{\varsigma = \chi(x^s, u^s)} \frac{\partial \chi(x^s, u^s)}{\partial u^s}(u^{s+1} - u^s)$$
$$+ O\left(\left|\begin{matrix} x^{s+1} - x^s \\ u^{s+1} - u^s \end{matrix}\right|^2\right) \qquad \text{(VII.8)}$$

In the standard procedure, the left hand side is set to zero, higher order terms are truncated, the consistency requirement is used to compute $x^{s+1}(=\chi^s)$ and the Section II notation is introduced, after which (VII.8) becomes

$$0 = y^{s+1} - y^* + C^{s+1}\Phi^s(\chi^s - x^s) + \left[C^{s+1}\Gamma^s\right]\Delta u^s \qquad \text{(VII.9)}$$

Solving (VII.9) for $(u^{s+1} - u^s)$ the first pseudo-Newton control law is obtained

CONTROL LAW II.

PSEUDO-NEWTON A

$$u^{s+1} = u^s + \left[C^{s+1}\Gamma^s\right]^{-1} C^{s+1}\Phi^s(x^s - \chi^s) + \left[C^{s+1}\Gamma^s\right]^{-1}(y^* - y^{s+1})$$

3.3 Stability Analysis

Applying Theorem V.2, the stability condition is:

$$\left\|\begin{matrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{matrix}\right\| < 1, \quad \forall (x, u) \in \bar{U}((x^*, u^*), r)$$

where $S_{11} = \Phi$, $S_{12} = \Gamma$,
$S_{21} = -[C\Gamma]^{-1}[C\Phi + \{C'\} \circ \Phi \bullet (\chi - x) + C \bullet \{\Phi_x\} \bullet (\chi - x) + C\Phi(\Phi - I) + \{C'\} \circ \Phi \bullet \Gamma(\psi - u) + C \bullet \{\Gamma_x\} \bullet (\psi - u)]$ and
$S_{22} = I - [C\Gamma]^{-1}[\{C'\} \circ \Gamma \bullet (\chi - x) + C \bullet \{\Phi_u\} \bullet (\chi - x) + C\Phi\Gamma + \{C'\} \circ \Gamma \bullet (\psi - u) + C \bullet \{\Gamma_u\} \bullet (\psi - u)]$

Example VII.3 : The Continuous Stirred Tank Reactor of example VI.1 is used to illustrate application of control law II.
Fig. VI.4 shows the system response to a step-up setpoint command under control law II. when the initial reactor state is at the left of the maximum conversion point.
In fig. VI.5 the situation is repeated, only this time the initial reactor state is at the right of the maximum conversion point.

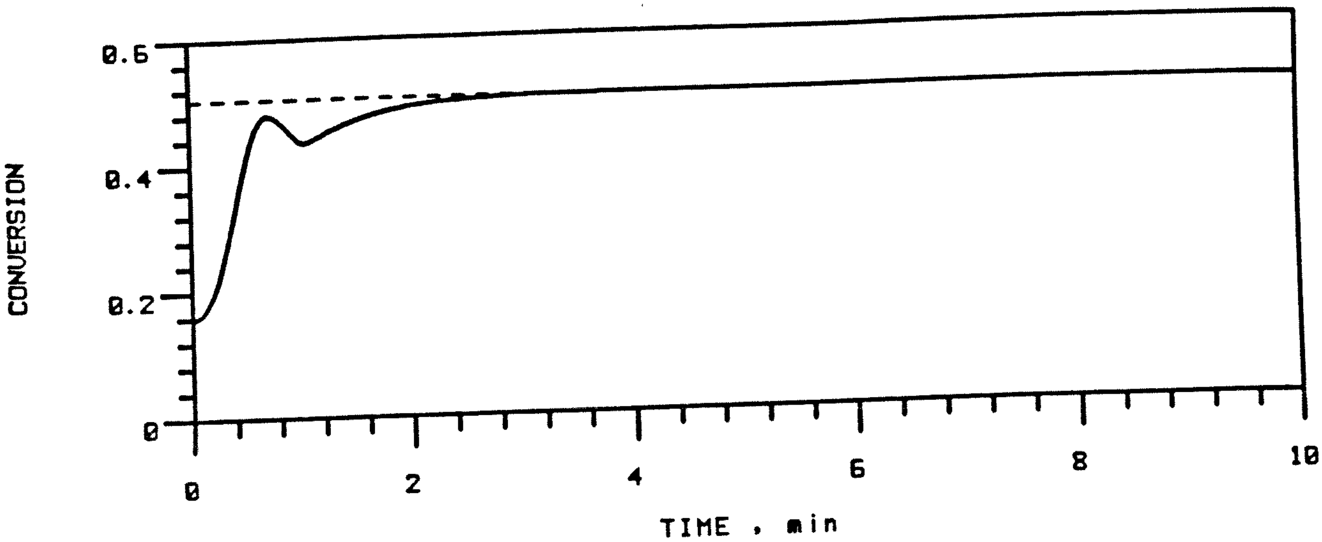

FIGURE VII.4 : *System response to a step-up setpoint command under control law II. Sampling time* $T = 1.0$ *min. Initial reactor state is at the left of the maximum conversion point* ($x_1^0 = 0.84$, $x_2^0 = 0.16$, $x_3^0 = 3.51$). *(– – –– : Setpoint command, ——— : System output.)*

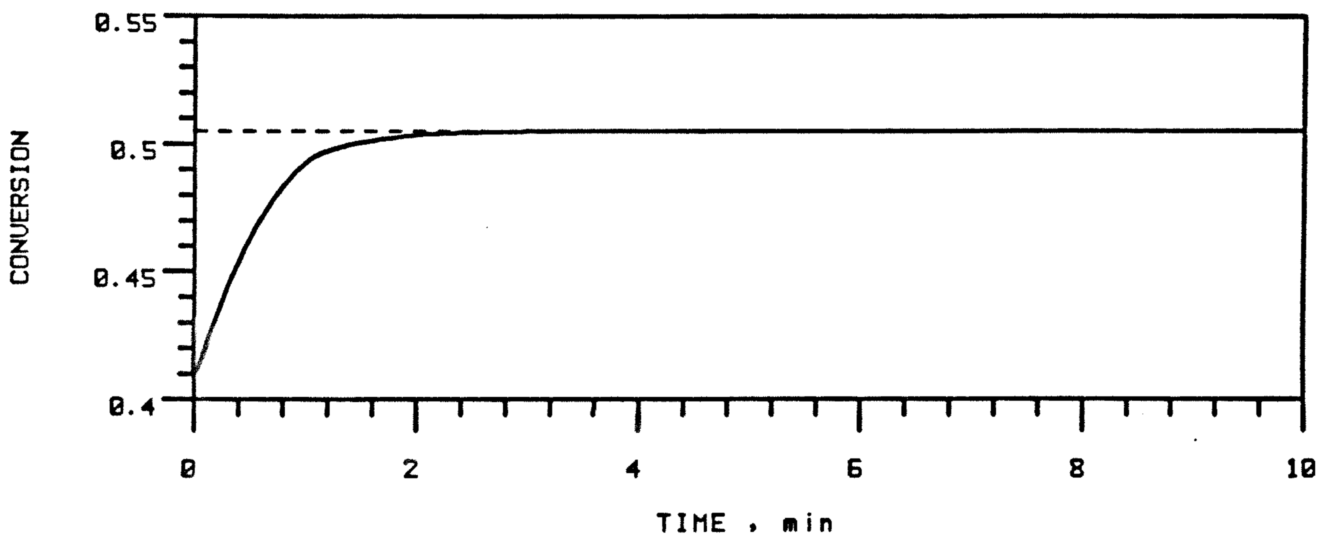

FIGURE VII.5 : *System response to a step-up setpoint command under control law II. Sampling time* $T = 1.0$ *min. Initial reactor state is at the right of the maximum conversion point* ($x_1^0 = 0.59$, $x_2^0 = 0.41$, $x_3^0 = 5.02$). *(– – –– : Setpoint command, ——— : System output.)*

3.4 Relaxed Law

Relaxing the updates in law II. by a factor λ, the relaxed law is obtained:

CONTROL LAW II.R

RELAXED PSEUDO-NEWTON A

$$u^{s+1} = u^s + \lambda \left[C^{s+1}\Gamma^s\right]^{-1} C^{s+1}\Phi^s(x^s - \chi^s) + \lambda \left[C^{s+1}\Gamma^s\right]^{-1} (y^* - y^{s+1})$$

Theorem V.2 characterizes the stability of the relaxed law. The norm condition is obtained from the previous paragraph, after multiplication of S_{21} and $(S_{22} - I)$ by the relaxation factor λ.

Example VII.4 : The effect of the relaxation factor is illustrated in fig. VII.6. The system response to a step-up setpoint command under control law II.R and 3 different relaxation factors is shown, when the initial reactor state is at the left of the maximum conversion point.

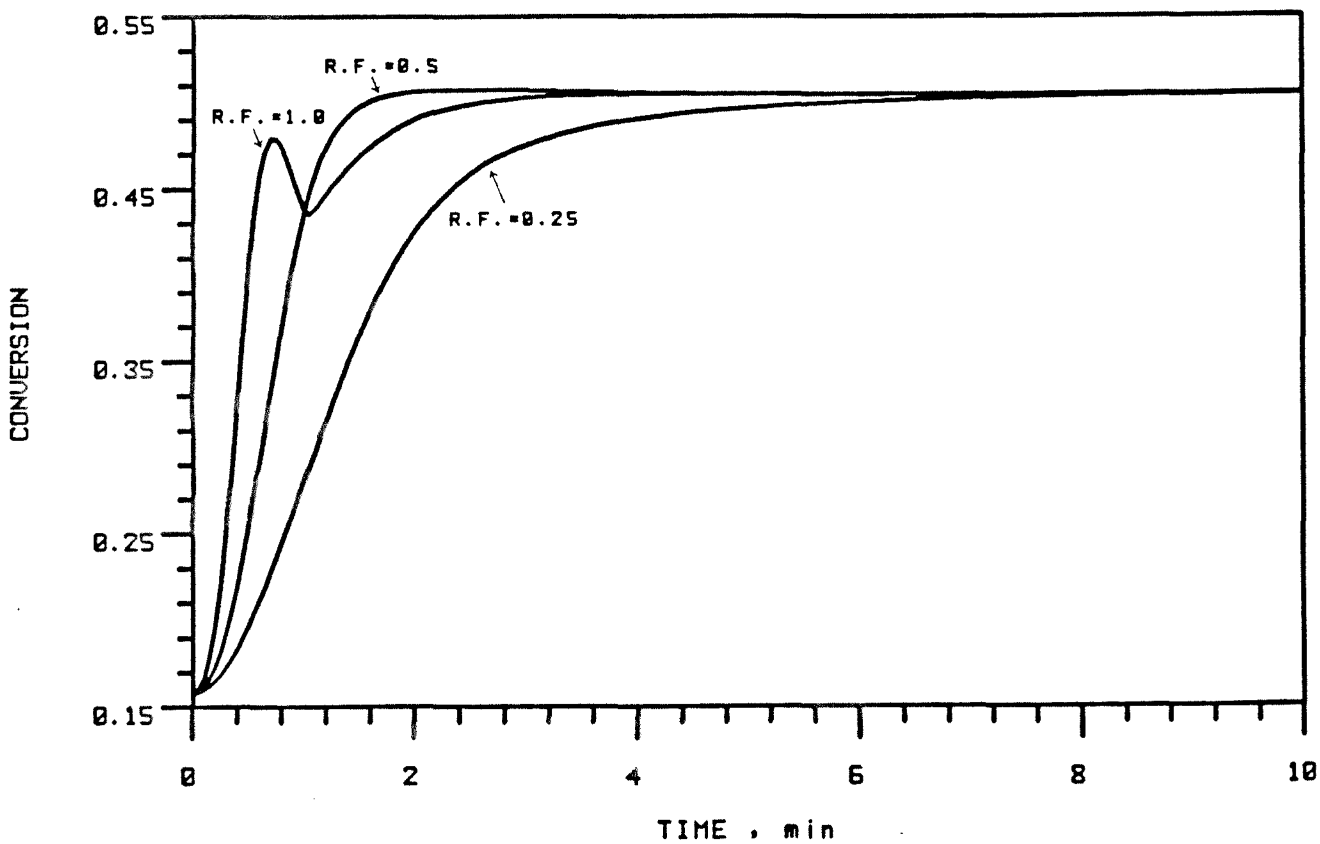

FIGURE VII.6 : *System response to a step-up setpoint command under control law II.R. Relaxation factors: 0.25, 0.5 and 1.0. Sampling time* $T = 1.0$ *min. Initial reactor state is at the left of the maximum conversion point (*$x_1^0 = 0.84$, $x_2^0 = 0.16$, $x_3^0 = 3.51$*). (– – –– : Setpoint command, ——— : System output.)*

3.5 Simplified Law

A reference point (x_{ref}, u_{ref}) is selected (for example the equilibrium state of interest) and the appropriate quantities Φ_{ref}, Γ_{ref} and C_{ref} are computed. The simplified control law is the linear control law obtained by replacing Φ^s, Γ^s and C^{s+1} in (VII.9) (which are matrix functions of x^s and u^s), by the constant matrices Φ_{ref}, Γ_{ref} and C_{ref} respectively.

CONTROL LAW II.S

SIMPLIFIED PSEUDO-NEWTON A

$$u^{s+1} = u^s + \left[C_{ref}\Gamma_{ref}\right]^{-1} C_{ref}\Phi_{ref}(x^s - \chi^s) + \left[C_{ref}\Gamma_{ref}\right]^{-1}(y^* - y^{s+1})$$

From Theorem V.2, the stability condition is

$$\left\| \begin{matrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{matrix} \right\| < 1, \quad \forall (x, u) \in \bar{U}((x^*, u^*), r)$$

where $S_{11} = \Phi$, $S_{12} = \Gamma$,
$S_{21} = -[C_{ref}\Gamma_{ref}]^{-1}C_{ref}\Phi_{ref}(\Phi - I) - [C_{ref}\Gamma_{ref}]^{-1}C\Phi$ and
$S_{22} = I - [C_{ref}\Gamma_{ref}]^{-1}C_{ref}\Phi_{ref}\Gamma - [C_{ref}\Gamma_{ref}]^{-1}C\Gamma$

3.6 Linear Law

For linear systems, it is shown in the appendix that control law II. becomes:

CONTROL LAW II.L

LINEAR PSEUDO-NEWTON A

$$\begin{aligned} u^{s+1} &= u^s + \left[C(e^{AT} - I)A^{-1}B\right]^{-1} \left[Ce^{AT}(x^s - \chi^s) + y^* - y^{s+1}\right] \\ &= u^s + \left[C(e^{AT} - I)A^{-1}B\right]^{-1} \left[y^* - Ce^{AT}\chi^s\right] \end{aligned}$$

The stability condition from Corollary V.4 is

$$\rho \begin{pmatrix} e^{AT} & (e^{AT} - I)A^{-1}B \\ -[C(e^{AT} - I)A^{-1}B]^{-1}Ce^{2AT} & -[C(e^{AT} - I)A^{-1}B]^{-1}Ce^{AT}(e^{AT} - I)A^{-1}B \end{pmatrix} < 1$$

Example VII.5 : Consider the two-input two-output linear system

$$\frac{\partial x_1}{\partial t} = -x_1 + x_2 + 3u_1 - u_2$$
$$\frac{\partial x_2}{\partial t} = -x_1 - 3x_2 + x_3 - u_2$$
$$\frac{\partial x_3}{\partial t} = x_2 - x_3 + u_1$$
$$y_1 = 2x_1 + 2x_2 - x_3$$
$$y_2 = 3x_1 - x_2$$

Using (II.5) and (II.8):

$$\Phi = e^{AT} = \begin{pmatrix} 0.785 & 0.153 & 2.1\times 10^{-2} \\ -0.153 & 0.472 & 0.153 \\ -2.1\times 10^{-2} & 0.153 & 0.800 \end{pmatrix}$$

$$\Gamma = (e^{AT} - I)A^{-1}B = \begin{pmatrix} 0.660 & -0.242 \\ -4.5\times 10^{-2} & -0.153 \\ 0.217 & -2.1\times 10^{-2} \end{pmatrix}$$

$$C = \begin{pmatrix} 2 & 2 & -1 \\ 3 & -1 & 0 \end{pmatrix}$$

Consequently

$$[C\Gamma]^{-1} = \begin{pmatrix} -0.585 & 0.786 \\ -2.07 & 1.03 \end{pmatrix}, \quad C\Phi = \begin{pmatrix} 1.23 & 1.1 & -0.45 \\ 2.43 & -0.013 & -0.09 \end{pmatrix}$$

Then control law II.L becomes

$$\begin{aligned} u_1^{s+1} &= u_1^s + 2.63(x_1^s - \chi_1^s) - 0.65(x_2^s - \chi_2^s) + 0.19(x_3^s - \chi_3^s) \\ &\quad - 0.585(y_1^* - y_1^{s+1}) + 0.786(y_2^* - y_2^{s+1}) \\ u_2^{s+1} &= u_2^s - 0.033(x_1^s - \chi_1^s) - 2.29(x_2^s - \chi_2^s) + 0.84(x_3^s - \chi_3^s) \\ &\quad - 2.07(y_1^* - y_1^{s+1}) + 1.03(y_2^* - y_2^{s+1}) \end{aligned}$$

Fig VII.7 shows the setpoint tracking behavior of the closed-loop system under this control law. The dead-beat action is evident.

3.7 Interpretation

Contrary to control law I. which tries to drive the system output to the desired value y^* at steady state, control law II. tries to drive the system output to y^* (up to first order in accuracy) in one step. As a result the control action is much more aggressive.

For linear systems it becomes a one step ahead output dead-beat controller (see Appendix VII). Its properties are well studied (Franklin and Powell 1981) and will not be repeated here. It

3.9 Asymptotic Behavior

For $T \longrightarrow \infty$, it is argued in Appendix VII that the control law II. becomes identical to control law I. and consequently it displays the same asymptotic behavior.

4. PSEUDO NEWTON CONTROL LAW B

4.1 Operator Equation

The operator equation is discrete this time:

$$\begin{aligned} g(x^{s+1}) &= y^* \qquad \text{(System output at } y^* \text{ in one step)} \\ x^{s+1} &= \chi(x^s, u^s) \qquad \text{(System evolution constraint)} \end{aligned} \tag{VII.10}$$

4.2 Control Law Derivation

Consider the variation to first order of the output map around a point x^{s+1}. Then

$$\begin{aligned} g(x^{s+2}) - y^* = g(x^{s+1}) - y^* + \left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right|_{\varsigma = x^{s+1}} (x^{s+2} - x^{s+1}) \\ + O(\|x^{s+2} - x^{s+1}\|^2) \end{aligned} \tag{VII.11}$$

x^{s+2} is the system state at $t = t^{s+2}$, i.e. $x^{s+2} = \chi(x^{s+1}, u^{s+1})$, a nonlinear function of u^{s+1}. If a u^{s+1} is desired which makes the right hand side of (VII.11) zero to first order (and produces $y^{s+2} = y^*$), the equation to be solved (after introducing the usual notation) is:

$$0 = y^{s+1} - y^* + C^{s+1}(\chi(x^{s+1}, u^{s+1}) - x^{s+1}) \tag{VII.12}$$

The nonlinear equation (VII.12) can be solved either by some iterative method (which is to be avoided in lieu of on-line calculations), or its solution can be approximated by the solution of an appropriate linear problem. This linear problem is obtained if $\chi(x^{s+1}, u^{s+1})$ is approximated to first order by expansion around u^s:

$$\chi(x^{s+1}, u^{s+1}) = \chi(x^{s+1}, u^s) + \frac{\partial \chi(x^{s+1}, u^s)}{\partial u^s}(u^{s+1} - u^s) + O(\|u^{s+1} - u^s\|^2) \tag{VII.13}$$

Now the notation is introduced: $\hat{\chi}^s = \chi(x^{s+1}, u^s)$, i.e the system state at $t = t^{s+2}$ if the input were held constant at $u = u^s$ over the $(s+1)^{th}$ sampling interval, and its associated partial derivative with respect to u^s, $\hat{\Gamma}^s \stackrel{\text{def}}{=} \dfrac{\partial \hat{\chi}^s}{\partial u^s}$. Substituting then (VII.13) into (VII.12) yields

$$0 = g^{s+1} - y^* + C^{s+1}(\hat{\chi} + \hat{\Gamma}^s(u^{s+1} - u^s) - \chi^s) \tag{VII.14}$$

Solving (VII.14) for u^{s+1} the second pseudo-Newton control law is obtained

CONTROL LAW III.

PSEUDO-NEWTON B

$$u^{s+1} = u^s + [C^{s+1}\hat{\Gamma}^s]^{-1}C(\chi^s - \hat{\chi}^s) + [C^{s+1}\hat{\Gamma}^s]^{-1}(y^* - y^{s+1})$$

4.3 Stability Analysis

The stability condition is obtained from Theorem V.2 :

$$\left\| \begin{matrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{matrix} \right\| < 1, \quad \forall (x,u) \in \bar{U}((x^*,u^*),r)$$

where $S_{11} = \Phi$, $S_{12} = \Gamma$,

$S_{21} = -[C\hat{\Gamma}]^{-1}[C\Phi + \{C'\} \circ \Phi(\hat{\chi} - \chi) + C(\hat{\Phi} - \Phi) + \{C'\} \circ \Phi\hat{\Gamma}(\psi - u) + C \bullet \{\hat{\Gamma}_x\} \bullet (\psi - u)]$ and

$S_{22} \doteq I - [C\hat{\Gamma}]^{-1}[\{C'\} \circ \Gamma(\hat{\chi} - \chi) + C(\hat{\Gamma} - \Gamma) + \{C'\} \circ \Gamma\hat{\Gamma}(\psi - u) + C \bullet \{\hat{\Gamma}_u\}(\psi - u)]$, $\hat{\Phi} = \dfrac{\partial \hat{\chi}}{\partial x}$.

Example VII.6 : The Continuous Stirred Tank Reactor of example VI.1 is used to illustrate application of control law III.

Fig. VI.8 shows the system response to a step-up setpoint command under control law III. when the initial reactor state is at the left of the maximum conversion point.

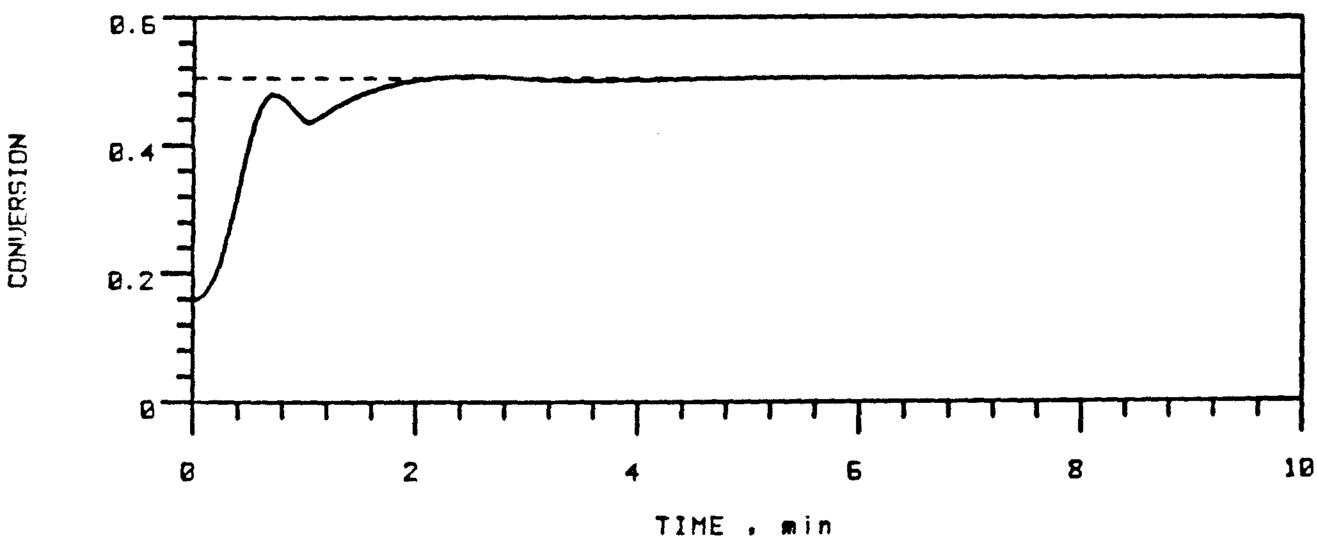

FIGURE VII.8 : *System response to a step-up setpoint command under control law III. Sampling time $T = 1.0$ min. Initial reactor state is at the left of the maximum conversion point ($x_1^0 = 0.84$, $x_2^0 = 0.16$, $x_3^0 = 3.51$). (– – – – : Setpoint command, ——— : System output.)*

In fig. VI.9 the situation is repeated, only this time the initial reactor state is at the right of the maximum conversion point.

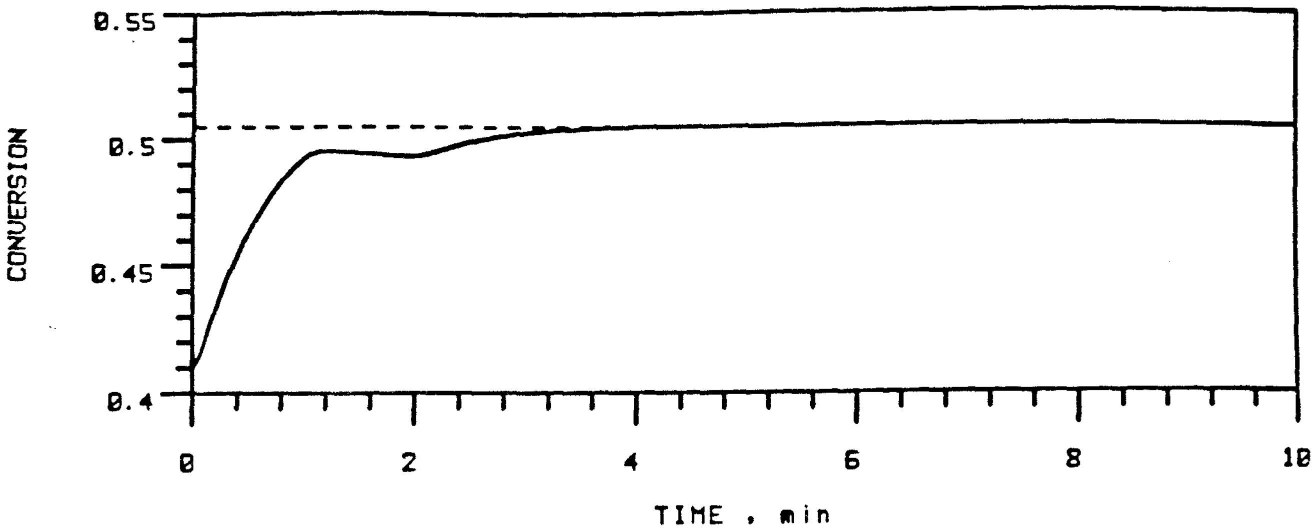

FIGURE VII.9 : *System response to a step-up setpoint command under control law III. Sampling time $T = 1.0$ min. Initial reactor state is at the right of the maximum conversion point ($x_1^0 = 0.59$, $x_2^0 = 0.41$, $x_3^0 = 5.02$). (– – –– : Setpoint command, ——— : System output.)*

4.4 Relaxed Law

Relaxing the updates in law III. by a factor λ, the relaxed law is obtained:

CONTROL LAW III.R

RELAXED PSEUDO-NEWTON B

$$u^{s+1} = u^s + \lambda[C^{s+1}\hat{\Gamma}^s]^{-1}C(\chi^s - \hat{\chi}^s) + \lambda[C^{s+1}\hat{\Gamma}^s]^{-1}(y^* - y^{s+1})$$

Example VII.7 : The effect of the relaxation factor is illustrated in fig. VII.10. The system response to a step-up setpoint command under control law III.R and 3 different relaxation factors is shown, when the initial reactor state is at the left of the maximum conversion point.

4.5 Simplified Law

A reference point (x_{ref}, u_{ref}) is selected (for example the equilibrium state of interest) and the appropriate quantities $\hat{\Gamma}_{ref}$ and C_{ref} are computed. The simplified control law is the linear control law obtained by replacing Φ^s, Γ^s and C^{s+1} in (VII.14) (which are matrix functions of

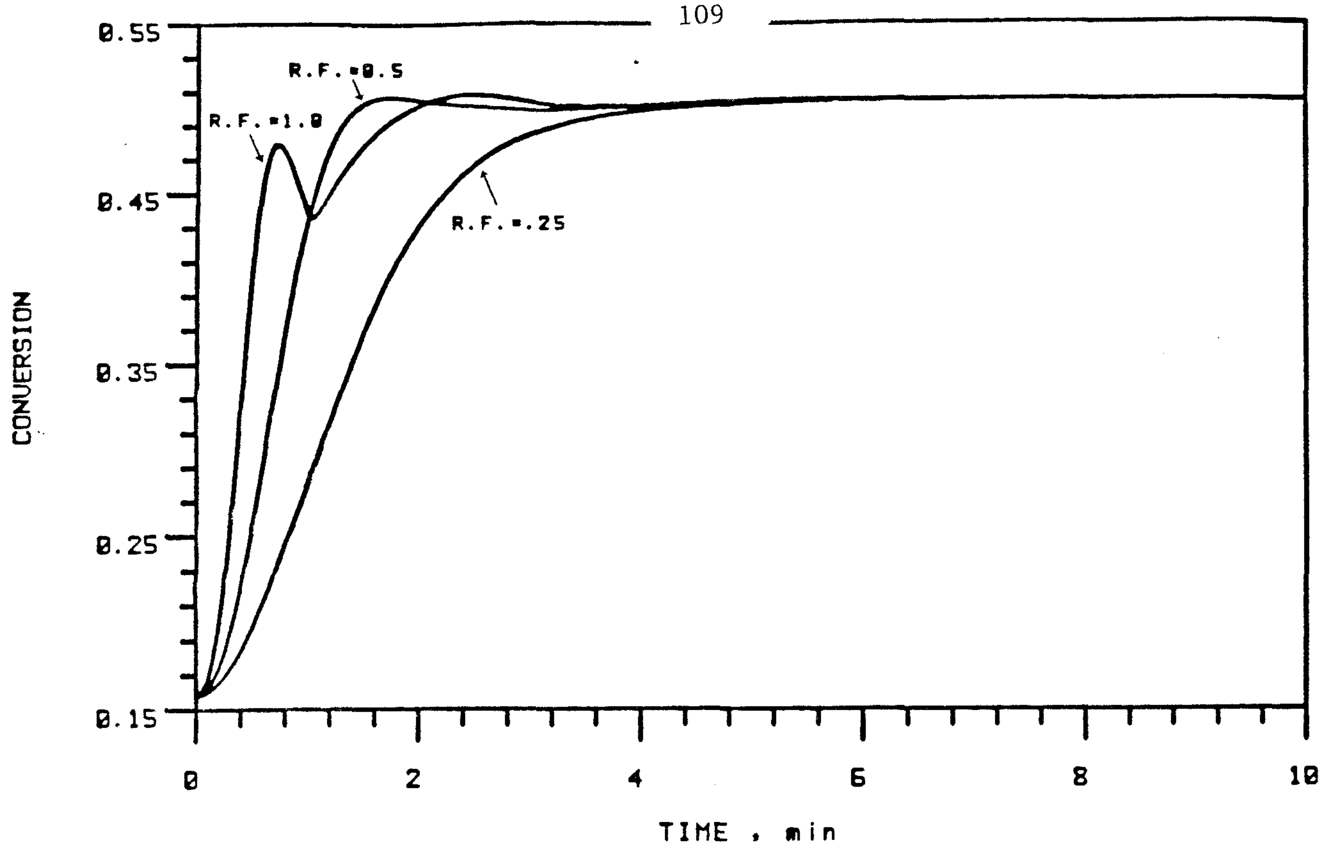

FIGURE VII.10 : *System response to a step-up setpoint command under control law III.R. Relaxation factors: 0.25, 0.5 and 1.0. Sampling time T = 1.0 min. Initial reactor state is at the left of the maximum conversion point ($x_1^0 = 0.84,\ x_2^0 = 0.16,\ x_3^0 = 3.51$). (– – –– : Setpoint command, ——— : System output.)*

x^s and u^s), by the constant matrices Φ_{ref}, Γ_{ref} and C_{ref} respectively.

CONTROL LAW III.S

SIMPLIFIED PSEUDO-NEWTON B

$$u^{s+1} = u^s + [C_{ref}\hat{\Gamma}_{ref}]^{-1}C_{ref}(\chi^s - \hat{\chi}^s) + \left[C_{ref}\hat{\Gamma}_{ref}\right]^{-1}(y^* - y^{s+1})$$

The stability condition is (Theorem V.2) :

$$\left\| \begin{matrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{matrix} \right\| < 1, \quad \forall (x,u) \in \bar{U}((x^*, u^*), r)$$

where $S_{11} = \Phi$, $S_{12} = \Gamma$,
$S_{21} = -[C_{ref}\hat{\Gamma}_{ref}]^{-1}C_{ref}\hat{\Phi}\Phi$ and $S_{22} = I - [C_{ref}\hat{\Gamma}_{ref}]^{-1}C_{ref}(\hat{\Phi}\Gamma + \hat{\Gamma})$.

4.6 Linear Law

For linear systems, it is shown in the Appendix that the resulting control law is

CONTROL LAW III.L

LINEAR PSEUDO-NEWTON B

$$u^{s+1} = u^s + \left[C(e^{AT} - I)A^{-1}B\right]^{-1} \left[Ce^{AT}(x^s - \chi^s) + y^* - y^{s+1}\right]$$
$$= u^s + \left[C(e^{AT} - I)A^{-1}B\right]^{-1} \left[y^* - Ce^{AT}\chi^s\right]$$

The stability condition from Corollary V.4 is :

$$\rho \begin{pmatrix} e^{AT} & (e^{AT} - I)A^{-1}B \\ -[C(e^{AT} - I)A^{-1}B]^{-1}Ce^{2AT} & -[C(e^{AT} - I)A^{-1}B]^{-1}Ce^{AT}(e^{AT} - I)A^{-1}B \end{pmatrix} < 1$$

4.7 Interpretation

The second pseudo-Newton law is only slightly different than the first. Consequently the same comments apply.

The similarity is demonstrated by the fact that the respective linear laws are identical one step ahead dead-beat controllers.

4.8 Implementation

Computations for the implementation of control law III. are carried out during the s^{th} sampling interval, in the following succession.

Step 1: The model equations (II.1) are integrated forward from t^s to t^{s+1} to obtain χ^s.

Step 2: The model equations are integrated forward from t^{s+1} to t^{s+2} with initial condition $x = \chi^s$ and input $u = u^s$ to obtain $\hat{\chi}^s$ and $\hat{\Gamma}^s$ (Section II.3).

Step 3: The quantities C^{s+1} and $[C^{s+1}\hat{\Gamma}^s]^{-1}$ are computed.

Step 4: The system output y^{s+1} and state χ^s at $t = t^{s+1}$ are measured.

Step 5: The new input to be injected into the system at $t = t^{s+1}$ is given by Control Law III.

4.9 Asymptotic Behavior

For $T \longrightarrow \infty$, it is argued in Appendix VII that the control law III. becomes identical to control law I. and consequently it displays the same asymptotic behavior.

5. REMARKS

I. There are quite a few differences between the Contraction Principle Control (CPC) laws of Chapter VI and the Newton Control (NC) laws of this Chapter. First, on the synthesis side, CPC laws are derived by solving norm optimization problems; with the solution depending on the particular norm selection. On the other hand the Newton method gives explicit formulas for the NC laws, without resorting to optimization; there is no norm dependence. Second, on the analysis side, stability analysis of NC laws is more involved than CPC analysis because the respective stability conditions depend explicitly on second derivatives of system operators. Third, the major computational burden associated with CPC laws is carried out off-line, with on-line computations being comparable to a standard linear control algorithm. On the contrary, there is little off-line computation associated with the Newton laws, which are virtually computed on line (Sections 2.7, 3.8 and 4.8). Finally, CPC algorithms are basically linear, while NC algorithms are nonlinear, adding to the complexity of the analysis.

II. It should be stressed that the operators Q used in the construction of the Hybrid algorithms do not enter in the control law expressions. For stability analysis of Hybrid algorithms their explicit form is not needed; it is enough to show the existence of Q that make the system operator a contraction.

III. The remarks concerning integral action in §4.4 of Chapter VI carry over to the case of NC laws: every stable NC law has no offset at steady state.

Appendix VII

Non-consistency of the Newton algorithms

Substitute the consistency requirement (IV.6) ($x^{s+1} = \chi(x^s, u^s)$) into (VII.4) to obtain

$$
\begin{aligned}
&(\Phi^s - I)(\chi^s - x^s) + \Gamma^s \Delta u^s = x^s - \chi^s \\
\Longrightarrow &\Phi^s(\chi^s - x^s) + \Gamma^s(u^{s+1} - u^s) = 0 \\
\Longrightarrow &C^{s+1}\Phi^s(\chi^s - x^s) + C^{s+1}\Gamma^s(u^{s+1} - u^s) = 0 \\
\overset{(VII.3.b)}{\Longrightarrow} &y^* - y^{s+1} = 0
\end{aligned}
$$

This should hold for every x^s, u^s and y^*, which of course is not true and as a result the Newton algorithm is not consistent.

Derivation of control law II.L

Example II.3 shows that for a linear system:

$$
\begin{aligned}
\chi^s &= \Phi x^s + \Gamma u^s \\
y^{s+1} &= C\chi^s
\end{aligned}
\tag{VII.15}
$$

Then control law II. becomes:

$$
\begin{aligned}
u^{s+1} &= (C\Gamma)^{-1}[C\Gamma u^s + C\Phi x^s - C\Phi\chi^s + y^* - C\chi^s] \\
&= (C\Gamma)^{-1}[C(\Phi x^s + \Gamma u^s) - C\chi^s + y^* - C\Phi\chi^s] \\
&= (C\Gamma)^{-1}[C\chi^s - C\chi^s + y^* - C\Phi\chi^s]
\end{aligned}
$$

Controll law II.L is immediately obtained when Φ and Γ in the expression above are replaced from (II.5) and (II.8) respectively.

The dead-beat action is shown next:

$$
\begin{aligned}
y^{s+2} &= Cx^{s+2} = C\Phi\chi^s + C\Gamma u^{s+1} \\
&= C\Phi\chi^s + y^* - C\Phi\chi^s \\
&= y^*
\end{aligned}
$$

Derivation of control law III.L

Example II.2 shows that for a linear system:

$$
\begin{aligned}
\chi^s &= \Phi x^s + \Gamma u^s \\
y^{s+1} &= C\chi^s
\end{aligned}
\tag{VII.16}
$$

By definition then

$$
\begin{aligned}
\hat{\chi}^s &= \Phi\chi^s + \Gamma u^s \\
\hat{\Gamma} &= \left.\frac{\partial \hat{\chi}^s}{\partial u^s}\right|_{\chi^s const.} = \Gamma
\end{aligned}
$$

Then control law III. becomes:

$$\begin{aligned} u^{s+1} &= (C\Gamma)^{-1}[C\Gamma u^s + C\chi^s - C\hat{\chi}^s + y^* - C\chi^s] \\ &= (C\Gamma)^{-1}[C\Gamma u^s - C\Phi\chi^s - C\Gamma u^s + y^*] \\ &= (C\Gamma)^{-1}[-C\Phi\chi^s + y^*] \end{aligned}$$

Controll law III.L is immediately obtained when Φ and Γ in the expression above are replaced from (II.5) and (II.8) respectively.

The dead-beat action is shown next:

$$\begin{aligned} y^{s+2} &= Cx^{s+2} = C\Phi\chi^s + C\Gamma u^{s+1} \\ &= C\Phi\chi^s + y^* - C\Phi\chi^s \\ &= y^* \end{aligned}$$

Asymptotic behavior of control law I.

The discussion is only qualitative at this stage: Assume $T \gg 1$ and that the system is globally stable. For any u^s let χ^s be the corresponding equilibrium state, i.e. $f(\chi^s, u^s) = 0$. Since the system is stable, at times large enough every time derivative will tend to zero. Then (II.3) yields

$$0 = \left.\frac{\partial f(\varsigma,\xi)}{\partial \varsigma}\right|_{\substack{\varsigma = \chi^s \\ \xi = u^s}} \Phi^s$$

Assuming invertibility of the derivative term, it follows that $\Phi^s = 0$.

Also, (II.6) yields

$$0 = \left.\frac{\partial f(\varsigma,\xi)}{\partial \varsigma}\right|_{\substack{\varsigma = \chi^s \\ \xi = u^s}} \Gamma^s + \left.\frac{\partial f}{\partial \xi}\right|_{\substack{\varsigma = \chi^s \\ \xi = u^s}}$$

and solving for Γ^s:

$$\Gamma^s = -\left[\left.\frac{\partial f(\varsigma,\xi)}{\partial \varsigma}\right|_{\substack{\varsigma = \chi^s \\ \xi = u^s}}\right]^{-1} \left.\frac{\partial f(\varsigma,\xi)}{\partial \xi}\right|_{\substack{\varsigma = \chi^s \\ \xi = u^s}}$$

Therefore, for large sampling times control law I. becomes

$$u^{s+1} = u^s - \left[\left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right|_{\varsigma = \chi^s} \left[\left.\frac{\partial f(\varsigma,\xi)}{\partial \varsigma}\right|_{\substack{\varsigma = \chi^s \\ \xi = u^s}}\right]^{-1} \left.\frac{\partial f(\varsigma,\xi)}{\partial \xi}\right|_{\substack{\varsigma = \varsigma \\ \xi = u^s}}\right]^{-1} (y^{s+1} - y^*) \quad \text{(VII.17)}$$

Consider now the algebraic equations that describe the steady-state of the system:

$$\begin{aligned} f(x,u) &= 0 \\ g(x) &= y^* \end{aligned}$$

Consider also the Newton algorithm for the solution:

1. Select u^s.
2. Solve $f(\chi^s, u^s) = 0$ to obtain $\chi^s(u^s)$.
3. Update u^s by the Newton formula

$$u^{s+1} = u^s - J^{-1}(g(\chi^s(u^s)) - y^*) \qquad \text{(VII.18)}$$

where the Jacobian J is

$$\begin{aligned} J &= \frac{\partial[g(x(u)) - y^*]}{\partial u} \\ &= \left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right|_{\varsigma = \chi^s} \frac{\partial \chi^s}{\partial u^s} \\ &= \left.\frac{\partial g(\varsigma)}{\partial \varsigma}\right|_{\varsigma = \chi^s} \left[\left.\frac{\partial f(\varsigma, \xi)}{\partial \varsigma}\right|_{\substack{\varsigma = \chi^s \\ \xi = u^s}} \right]^{-1} \left.\frac{\partial f(\varsigma, \xi)}{\partial \xi}\right|_{\substack{\varsigma = \chi^s \\ \xi = u^s}} \end{aligned}$$

Substitute J in (VII.18) to obtain (VII.17) .

CHAPTER VIII

CONCLUSIONS

There is a strong duality between the theory of feedback control and operator equation theory, is the thesis of this dissertation. The research objective was to establish the duality in a rigorous context and show that control theory can benefit from a well defined operator theory, with the emphasis in the area of nonlinear systems analysis and controller design.

To this end a framework was developed that formulated the control problem as an operator inversion problem. The principles and concepts involved were shown to be basically identical. Application of operator equation methodologies to the fundamental control problems was then straightforward. Inversion algorithm development led to control law synthesis and algorithm convergence to stability analysis.

Original results of theoretical value and practical significance were claimed. On the analysis part, a general stability theory for nonlinear discrete systems was developed. Conditions for local, finite and global stability were stated and proved and the respective computational aspects were stressed. On the synthesis part, control laws were derived and shown to possess desirable stability and performance characteristics. When condensed to the case of linear systems, the results were found to conform well with traditional linear systems theory concepts and methodologies.

The structure of the dissertation proceeded along the following lines:

The importance of research in the area of nonlinear systems control was advocated in Chapter I. The argument was supported by a literature review stressing the unavailability of a general design theory. The lack of practically applicable analysis methods and synthesis techniques was emphasized.

The theoretical and computational background was set in the following two Chapters. In Chapter II the basic set of assumptions was laid out (namely, well behaved dynamic systems described by coupled ordinary differential equations, with certain continuity and differentiability properties of the solutions). A computational theory for differentiation of nonlinear operators was worked out by transforming the associated problems to initial value problems. In Chapter III fundamental results from operator equation theory were assembled to build the theoretical infrastructure. Some new results of unique importance to control considerations were added.

Chapter IV set the conceptual background of the proposed theory. It was shown how control problems can be transformed to operator equation problems for different sets of objectives. Both the similarities and unique features of operator inversion and feedback control problems were discussed to set the pace for the basic theoretical developments to follow.

A general stability theory was detailed in Chapter V. Conditions that assert (exponential) stability in the face of infinitessimal, finite and infinite perturbations were established. Both the open and closed-loop cases were investigated. The results were found in good agreement with traditional concepts of linear and nonlinear stability. Computational procedures associated with the analysis theorems were given.

The final Chapters addressed the synthesis aspects. Stemming from the analysis theory, in Chapter VI linear control laws for nonlinear systems were shown to be solutions to optimization problems minimizing a norm of the closed-loop system derivative operator. Although possible extrapolations to nonlinear control laws were proposed, it was pointed out that the successive substitution method is not suitable to generate nonlinear controllers.

The method of Newton was employed in Chapter VII as the prototype in deriving nonlinear control laws for nonlinear systems. It was not chosen on the basis of its properties, rather than for its clarity in providing guidelines for extension of the approach to alternative operator equation solution methods.

The original material of the thesis is believed to be in: Chapter II, section 3. Chapter III the parts associated with Theorems III.3, III.6, as well as section 3; finally, Chapters IV, V, VI and VII in their entirety.

The main objective of the work was to establish the duality of control and operator theories respectively and show how it can be profitably applied to control problems of theoretical and practical importance. It is not claimed that the treatment given is complete, because of the generality of the approach and its far reaching consequences. The development stops at the point where theoretical work and application to real life systems can be independently continued.

A number of important problems and promising extensions stand out at present.

A basic problem is associated with developing more efficient computational procedures to a) check the stability conditions of the Chapter V theorems and b) solve the optimization problems of Chapter VI.

A second problem is to reduce the conservativeness of the norm stability conditions of Chapter V and the associated conservativeness of the Chapter VI design problems. In particular, norm optimization by scaling should be investigated in detail.

Finally the effect of the sampling time was not covered in any detail, although it affects both the analysis and synthesis and consequently should be dealt with in future efforts.

Following the Newton prototype of Chapter VII, control law synthesis can be straightforwardly extended to alternative operator equation methods to benefit from their respective desirable properties. The "robust" Newton methods of Stoer, Deufelhard and Nickel, the family of steepest descent methods (where the inversion problem is transformed to a functional minimization problem) and continuation methods pose as particularly promising alternatives.

Another road to extending the current results is by enlarging the class of systems considered. For once, the theory is directly applicable to systems of coupled differential and algebraic equations arising in chemical reactors, realizations of partial differential equations etc. At the same time extension of the concept to distributed parameter and adaptive systems is possible, after the appropriate computational tools of Chapter II are extended in the respective classes.

Finally, extension to continuous systems should be investigated. It was not attempted at this stage, because operator solution methods are inherently discrete event processes, naturally befitting to discrete control considerations.

REFERENCES

[1] B.D.O. Anderson and J.B. Moore, "Optimal Filtering", Prentice-Hall, Englewood Cliffs NJ, 1979

[2] M. Athans, P.L. Falb, "Optimal Control", McGraw Hill, 1966

[3] K.J. Åström and B. Wittenmark, "Computer Controlled Systems: Theory and Design", Prentice – Hall, N.J., 1984

[4] S. Banach, "Sur les opérations dans les ensembles abstraits et leur applications aux equations integrales", Fund. Math., **3**, 133-181, 1922

[5] N.E. Barabanov, "Stability, Instability and Dichotomy of Control Systems with Gradient Nonlinearities" Avtomat. Telemekh., No **4**, p. 5-8, 1982

[6] R.G. Bartle, "Newton's Method in Banach Spaces", Proc. Amer. Math. Soc., **6**, 827-831, 1955

[7] A.E. Bryson, Y.C. Ho, "Applied Optimal Control", John Wiley, N.Y., 1975

[8] R. Caccioppoli, "Sugli elementi uniti delle transformazioni funzionali: un' osservazione sui problemi di valori ai limiti", Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez. I, **6**, 498-502, 1931

[9] M. Caracotsios and W.E. Stewart, "Sensitivity Analysis of Initial Value Problems with Mixed ODE's and Algebraic Equations", Comp. Chem. Engng., 1984 (in press).

[10] R.F. Curtain, A.J. Pritchard, "Functional Analysis in Modern Applied Mathematics", Academic Press, London, 1977

[11] C.A. Desoer, M. Vidyasagar, "Feedback Systems: Input–Output Properties", Academic Press, 1975

[12] P. Deufelhard, "A Modified Newton Method for the Solution of ill - conditioned Systems of Nonlinear Equations with Application to Multiple Shooting", Numer. Math., **22**, 289–315, 1974

[13] J. C. Doyle and G. Stein, "Robustness with observers", IEEE Trans. Auto. Cntrl., August, 1979

[14] J.C. Doyle, "Guaranteed margins for LQG regulators, IEEE Trans. Autom. Cntrl., AC-**23**, pp 756-757, Aug. 1978

[15] J. C. Doyle and G.Stein, "Multivariable feedback design: Concepts for a classical / modern synthesis", IEEE Trans. Automat. Contr., vol AC-**26**, pp. 4–16, 1981

[16] J.C. Doyle, "Lecture Notes", ONR/Honeywell Workshop on Advances in Multivariable Control, Minneapolis MN, Oct. 8-10, 1984

[17] G.C. Economou and M. Morari, "Internal Model Control 4: Extension to Nonlinear Systems", Ind. & Eng. Chem., Proc. Des. Dev., 1985 (In Press).

[18] G.F. Franklin and J.D. Powell, "Digital Control of Dynamic Systems", Addison-Wesley Publishing Co., Reading MA, 1981

[19] C.E. Garcia and M. Morari, "Internal model control 1: A unifying review and ome new results", Ind. & Eng. Chem., Proc. Des. Dev., **21**, pp. 308-323, 1982

[20] G.H. Golub and C.F. Van Loan, "Matrix Computations", John Hopkins, Baltimore MD, 1984

[21] G.C. Goodwin, K.S. Sin, "Adaptive Filtering Prediction and Control", Prentice Hall, N.J., 1984

[22] L.T. Grujić and D.D. Šiljak, "On Stability of Discrete Composite Systems", IEEE Trans. Auto. Cntrl., AC-**18**, October 1973, pp. 522-524

[23] J.M. Holtzman, "Nonlinear Systems Theory", Prentice-Hall, N.J., 1970

[24] L.R. Hunt, R. Su and G. Meyer, "Global Transformations of Nonlinear Systems", IEEE Trans. Autom. Cntrl, AC-**28**, pp 24-30, January 1983

[25] T. Kailath, "Linear Systems", Prentice Hall, N.J., 1980

[26] R. E. Kalman and J. E. Bertram, "Control System Analysis and Design via the 'Second Method' of Lyapunov. Parts I and II", ASME transactions, J. Basic Engng., June 1960, pp. 371–400

[27] L. Kantorović and G. Akilow, "Functional analysis in normed spaces", Pergamon Press, 1964

[28] D.E. Knuth, "The TeXbook", Addison-Wesley, Reading MA, 1984

[29] A.N. Kolmogorov, S.V. Fomin, "Elements of the Theory of Functions and Functional Analysis Vol. 1: Metric and Normed Spaces", Graylock Press, Rochester NY, 1957

[30] N.N. Krassovskii, "Some Problems of the Theory of the Stability of Motion", (in Russian), Moscow: Fizmatgiz, 1959 (English Translation: "Stability of Motion", Stanford Univ. Press, Stanford, CA, 1963)

[31] I.D. Landau, "Adaptive Control – The Model Reference Approach", Marcel Dekker, New York, 1979

[32] J.P. LaSalle, S. Lefschetz, "Stability by Liapunov's Direct Method", Academic Press, London, 1961

[33] N. A. Lehtomaki, "Practical robustness measures in multivariable control systems analysis", Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge MA, 1981

[34] J.R. Leis and M.A. Kramer, "Sensitivity Analysis of Systems of Differential and Algebraic Equations", AIChE Annual Meeting, Paper No. 27c., 1984

[35] D.G. Luenberger, "Introduction to Linear and Nonlinear Programming", Addison-Wesley, Reading MA, 1973

[36] A.A. Lyapunov, "Problèm général de la stabilité du mouvement", Photoreproduction as Annals of Mathematical Society No. 17 (Princeton Univ. Press, Princeton NJ) of the 1907 French translation from the original Russian version published in 1892.

[37] G. Meyer, "Flight Test Results of a Nonlinear Control strategy", Paper FP2, ACC conference, Boston MA, 1985

[38] A.P. Molchanov, E.S. Pyatinskii, "Absolute instability of Nonlinear nonstationary systems I", Avtomat. Telemekh., No 1, p. 19-27, 1982

[39] A.P. Molchanov, E.S. Pyatinskii, "Absolute instability of Nonlinear nonstationary systems II", Avtomat. Telemekh., No 2, p. 17-28, 1982

[40] A.P. Molchanov, E.S. Pyatinskii, "Absolute instability of Nonlinear nonstationary systems III", Avtomat. Telemekh., No 3, p. 29-43, 1982

[41] N. Munro, "Modern approaches to control system design", IEE Control Engineering Series No. 9, Peter Peregrinus Ltd, London 1979

[42] K.L. Nickel, "A Globally Convergent Ball Newton Method", Mathematics Research Center Summary Report #2022, Madison WI, 1979

[43] J.M. Ortega, W.C. Rheinboldt, "Iterative Solution of Nonlinear Equations in Several Variables", Academic Press, New York 1970

[44] L.R. Petzold, "A description of DASSL: a differential / algebraic system solver", Sandia National Laboratories Report, SAND82-8637, 1982

[45] V.M. Popov, "Hyperstability of Automatic Control Systems", Springer-Verlag, New York, 1973

[46] L.B. Rall, "Computational Solutions of Nonlinear Operator Equations", R.E. Krieger, N.Y., 1979

[47] R.D. Richtmeyer and K.W. Morton, "*Difference Methods for Initial Value Problems*", 2^{nd} edition, Interscience Publications, J. Wiley, New York NY, 1969

[48] H.H. Rosenbrock, "Computer Aided Control System Design", Academic Press, 1974

[49] M.G. Safonov, "Stability and Robustness of Multivariable Feedback Systems", MIT Press, Cambridge MA, 1980

[50] V.I. Skorodinskii, "Absolute stability and absolute instability of control systems with two nonstationary nonlinear elements, Part I", Avtomat. Telemekh., No 9, p. 21-29, 1981

[51] V.I. Skorodinskii, "Absolute stability and absolute instability of control systems with two nonstationary nonlinear elements, Part II", Avtomat. Telemekh., No 6, p. 87-94, 1982

[52] I. Stakold, "Green's Functions and Boundary Value Problems", John Wiley, NY, 1979

[53] G. Stein, "LQG-Based multivariable Design: Frequency domain interpretation", AGARD Lecture Series No. 117, NATO, 1981

[54] J. Stoer, "Einführung in die Numerische Mathematik I." Heidelberger Taschenbuch 105, Berlin - Heidelberg - New York, Springer, 1972 (in german)

[55] J. Stoer and R. Bulirsch, "Einführung in die Numerische Mathematik II." Heidelberger Taschenbuch 114, Berlin - Heidelberg - New York, Springer, 1973 (in german)

[56] M. Vidyasagar, "Nonlinear Systems Analysis", Prentice Hall, N.J., 1978

[57] G. Zames, "On the input-output stability of time-varying nonlinear systems, Parts I and II", IEEE Trans. Autom. Contr., vol AC-**11**, pp. 228-239 and 465-476, 1966

[58] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms and approximate inverses", IEEE Trans. Auto. Contr., AC-**26**, 1981