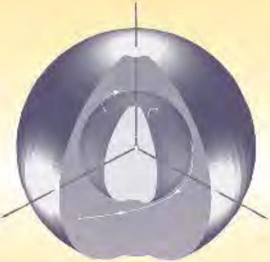
Control Engineering Series

NONLINEAR CONTROL SYSTEMS



Zoran Vukić
Ljubomir Kuljača
Dali Đonlagić
Sejid Tešnjak

NONLINEAR CONTROL SYSTEMS

Zoran Vukić
Ljubomir Kuljača
University of Zagreb
Zagreb, Croatia

Dali Đonlagić
University of Maribor
Maribor, Slovenia

Sejid Tešnjak
University of Zagreb
Zagreb, Croatia

MARCEL DEKKER, INC. NEW YORK • BASEL

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress.

ISBN: 0-8247-4112-9

This book is printed on acid-free paper.

Headquarters
Marcel Dekker, Inc.
270 Madison Avenue, New York, NY 10016
tel: 212-696-9000; fax: 212-685-4540

Eastern Hemisphere Distribution
Marcel Dekker AG
Hutgasse 4, Postfach 812, CH-4001 Basel, Switzerland
tel: 41-61-260-6300; fax: 41-61-260-6333

World Wide Web
http://www.dekker.com

The publisher offers discounts on this book when ordered in bulk quantities. For more information, write to Special Sales/Professional Marketing at the headquarters address above.

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

CONTROL ENGINEERING

A Series of Reference Books and Textbooks

Editor

NEIL MUNRO, PH.D., D.SC.

Professor
Applied Control Engineering
University of Manchester Institute of Science and Technology
Manchester, United Kingdom

1. Nonlinear Control of Electric Machinery, *Darren M. Dawson, Jun Hu, and Timothy C. Burg*
2. Computational Intelligence in Control Engineering, *Robert E. King*
3. Quantitative Feedback Theory: Fundamentals and Applications, *Constantine H. Houpis and Steven J. Rasmussen*
4. Self-Learning Control of Finite Markov Chains, *A. S. Poznyak, K. Najim, and E. Gómez-Ramírez*
5. Robust Control and Filtering for Time-Delay Systems, *Magdi S. Mahmoud*
6. Classical Feedback Control: With MATLAB, *Boris J. Lurie and Paul J. Enright*
7. Optimal Control of Singularly Perturbed Linear Systems and Applications: High-Accuracy Techniques, *Zoran Gajić and Myo-Taeg Lim*
8. Engineering System Dynamics: A Unified Graph-Centered Approach, *Forbes T. Brown*
9. Advanced Process Identification and Control, *Enso Ikonen and Kaddour Najim*
10. Modern Control Engineering, *P. N. Paraskevopoulos*
11. Sliding Mode Control in Engineering, *edited by Wilfrid Perruquetti and Jean Pierre Barbot*
12. Actuator Saturation Control, *edited by Vikram Kapila and Karolos M. Grigoriadis*
13. Nonlinear Control Systems, *Zoran Vukić, Ljubomir Kuljača, Dali Đonlagić, Sejid Tešnjak*

Additional Volumes in Preparation

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Series Introduction

Many textbooks have been written on control engineering, describing new techniques for controlling systems, or new and better ways of mathematically formulating existing methods to solve the ever-increasing complex problems faced by practicing engineers. However, few of these books fully address the applications aspects of control engineering. It is the intention of this new series to redress this situation.

The series will stress applications issues, and not just the mathematics of control engineering. It will provide texts that present not only both new and well-established techniques, but also detailed examples of the application of these methods to the solution of real-world problems. The authors will be drawn from both the academic world and the relevant applications sectors.

There are already many exciting examples of the application of control techniques in the established fields of electrical, mechanical (including aerospace), and chemical engineering. We have only to look around in today's highly automated society to see the use of advanced robotics techniques in the manufacturing industries; the use of automated control and navigation systems in air and surface transport systems; the increasing use of intelligent control systems in the many artifacts available to the domestic consumer market; and the reliable supply of water, gas, and electrical power to the domestic consumer and to industry. However, there are currently many challenging problems that could benefit from wider exposure to the applicability of control methodologies, and the systematic systems-oriented basis inherent in the application of control techniques.

This series presents books that draw on expertise from both the academic world and the applications domains, and will be useful not only as academically recommended course texts but also as handbooks for practitioners in many applications domains. *Nonlinear Control Systems* is another outstanding entry to Dekker's Control Engineering series.

Preface

Nonlinear control designs are very important because new technology asks for more and more sophisticated solutions. The development of nonlinear geometric methods (Isidori, 1995), feedback passivation (Fradkov and Hill, 1998), a recursive design – backstepping (Sontag and Sussmann, 1988; Krstić et al, 1995), output feedback design (Khalil, 1996) and unconventional methods of fuzzy/neuro control (Kosko, 1992), to mention just few of them, enriched our capability to solve practical control problems. For the interested reader a tutorial overview is given in Kokotović, 1991 and Kokotović and Arcak, 2001. Analysis always precedes design and this book is primarily focusing on nonlinear control system analysis. The book is written for the reader who has the necessary prerequisite knowledge of mathematics acquired up to the 3rd or 4th year in an undergraduate electrical and/or mechanical engineering program, covering mathematical analysis (theory of linear differential equations, theory of complex variables, linear algebra), theory of linear systems and signals and design of linear control systems. The book evolved from the lectures given at the University of Zagreb, Croatia and the University of Maribor, Slovenia for undergraduate and graduate courses, and also from our research and development projects over the years.

Fully aware of the fact that for nonlinear systems some new and promising results were recently obtained, we nevertheless decided in favor of classical techniques that are still useful in engineering practice. We believe that the knowledge of classical analysis methods is necessary for anyone willing to acquire new design methods. The topics covered in the book are for advanced undergraduate and introductory graduate courses. These topics will be interesting also for researchers, electrical, mechanical and chemical engineers in the area of control engineering for their self-study. We also believe that those engineers dealing with modernization of industrial plants, consulting, or design will also benefit from this book. The main topics of the book cover classical methods such as: stability analysis, analysis of nonlinear control systems by use of the describing function method and the describing function method applied to fuzzy control systems. Many examples given in the book help the reader to better understand the topic. Our experience is that with these examples and similar problems given to students as project-type homework, they can quite successfully–with the help of some computer programs with viable control system computational tools (such as MATLAB®, by The Mathworks, Inc.)—learn about nonlinear control systems. The material given in the book is extensive and can't be covered in a one-semester course. What will be covered in a particular course is up to the lecturer. We wanted to achieve better readability and presentation of this difficult area, but not by compromising the rigor of mathematics. This goal was not easy to achieve and the reader will judge if we succeeded in that. The basic knowledge necessary for the further study of nonlinear control systems is given, and we hope that readers will enjoy it.

Acknowledgments

It is our privilege to thank the many colleagues and students who were involved and have helped us in the various stages of the book preparation. We would like to thank our colleagues, listed in alphabetical order: Professors, Guy O. Beale, Petar Kokotović, Luka Korkut, Frank Lewis, Manfred Morari, Tugomir Šurina, Boris Tovornik, Kimon Valavanis, and Tine Zorič for their valuable suggestions and help. We are grateful to our research students, in alphabetical order: Vedran Bakarić, Bruno Borović, Krunoslav Horvat, Ognjen Kuljača, Ozren Kuljača, Nenad Muškinja, Edin Omerdić, Henrieta Ožbolt, and Dean Pavleković. They helped in the preparation of figures (Ozren, Nenad, Dean) and in solving examples given in the book (all of them). Useful comments and help in preparation of the camera-ready copy were done by Dean. We would also like to thank B.J. Clark and Brian Black from Marcel Dekker, Inc. for their help and support.

We are grateful to our families for their patience, understanding and support in the course of the preparation of this book. The book is dedicated to them.

Zoran Vukić

Ljubomir Kuljača

Dali Donlagić

Sejid Tešnjak

Contents

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Chapter 1

Properties of Nonlinear Systems

Nonlinear systems behave naturally. Our experience and experiments are proving that every day. However, our students are indoctrinated by the linear theories throughout their first years of study. We have to admit that we also contribute to that through our lectures. So, it seems appropriate, at the beginning of this book, to briefly present properties or dynamical phenomena of nonlinear systems in order to refresh our knowledge about some types of dynamic behavior which we all experience more or less in our surroundings or in the laboratory. Some basic definitions and theorems which will be useful for understanding the subsequent text are also given here. Definitions have to be mathematically rigorous and maybe for some of our readers not so appealing.

1.1 Problems in the Theory of Nonlinear Systems

Physical systems in general, and technical systems in particular, are as a rule nonlinear, i.e. they comprise nonlinearities. By nonlinearities we imply any deflections from linear characteristics or from linear equations of the system's dynamics.

In general, analysis of automatic control system dynamics starts with the mathematical description of the individual elements of the closed-loop system. The description involves linear or nonlinear differential equations, which — combined with the external quantities acting on the system — form the mathematical model of the dynamic performance of the system.

If the dynamic performance of all elements can be described by linear differential equations, then the system as a whole can be described by a linear differential equation. This is called a *linear control system*. Linear control theory

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

based on linear models is useful in the beginning stages of research, but in general cannot comprise all the diversities of the dynamic performance of a real system. Namely, it is well known that the stability of the equilibrium state of the linear system depends neither on initial conditions nor on external quantities which act on the system, but is dependent exclusively on the system's parameters. On the other hand, the stability of the equilibrium states of a nonlinear system is predominantly dependent on the system's parameters, initial conditions as well as on the form and magnitude of acting external quantities. Linear system theory has significant advantages which enable a relatively simple analysis and design of control systems. The system is linear under three presumptions:

(i) Additivity of zero-input and zero-state response,

(ii) Linearity in relation to initial conditions (linearity of zero-input response),

(iii) Linearity in relation to inputs (linearity of zero-state response[1]).

DEFINITION 1.1 (LINEARITY OF THE FUNCTION)
The function $f(x)$ is linear with respect to independent variable x if and only if it satisfies two conditions:

1. *Additivity:* $f(x_1+x_2) = f(x_1)+f(x_2)$, $\forall x_1, x_2$ *in domain of the function* f,
2. *Homogeneity:* $f(\alpha x) = \alpha f(x)$, $\forall x$ *in domain of the function* f, *and all scalars* α.

The systems which don't satisfy the conditions (i), (ii) and (iii) are *nonlinear* systems.

A linear system can be classified according to various criteria. One of the most important properties is stability. It is said that a linear system is:

Stable if after a certain time the system settles to a new equilibrium state,

Unstable if the system is not settled in a new equilibrium state,

Neutral if the system state obtains the form of a periodic function with constant amplitude.

An important class of the linear systems are *time-invariant systems*[2]. Mathematically, they can be described by a linear function which does not depend on time, i.e. a first-order vector differential equation:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t)]; \quad \forall t \geq 0 \\ \mathbf{y}(t) &= \mathbf{h}[\mathbf{x}(t), \mathbf{u}(t)]; \quad \forall t \geq 0\end{aligned}$$

[1] The principle of linear superposition is related only to the zero-state responses of the system.

[2] Many authors use the notion stationary systems instead of time-invariant systems, when the system is discussed through input and output signals. The other term in use is autonomous systems.

where $\mathbf{x}(t)$ is n-dimensional column state vector of the system states, $\mathbf{x} \in \Re^n$; $\mathbf{u}(t)$ is m-dimensional column input vector of the system excitations, $\mathbf{u} \in \Re^m$; $\mathbf{y}(t)$ is r-dimensional column output vector of the system responses, $\mathbf{y} \in \Re^r$; $\mathbf{f}, \mathbf{h}$ are linear vector functions of independent variables $\mathbf{x}$ and $\mathbf{u}$ and t is a scalar variable (time), $t \in \Re_+$.

Among the linear time-invariant systems, of special interest are the systems whose dynamic behavior is described by differential equations with constant coefficients. Such systems are unique since they are covered by a general mathematical theory.

When the physical systems are discussed, it is often not possible to avoid various nonlinearities inherent to the system. They arise either from energy limitations or imperfectly made elements, or from intentionally added special nonlinear elements which improve dynamic properties of the system. All these nonlinearities can be classified either as nonessential, i.e. which can be neglected, or essential ones, i.e. which directly influence the system's dynamics.

Automatic control systems with at least one nonlinearity are called nonlinear systems. Their dynamics are described by nonlinear mathematical models. Contrary to linear system theory, there doesn't exist a general theory applicable to all nonlinear systems. Instead, a much more complicated mathematical apparatus is applied, such as functional analysis, differential geometry or the theory of nonlinear differential equations. From this theory is known that except in some special cases (Riccati or Bernouilli equations, equations which generate elliptic integrals), a general solution cannot be found. Instead, individual procedures are applied, which are often improper and too complex for engineering practice. This is the reason why a series of approximate procedures are in use, in order to get some necessary knowledge about the system's dynamic properties. With such approximate procedures, nonlinear characteristics of real elements are substituted by idealized ones, which can be described mathematically.

Current procedures for the analysis of nonlinear control systems are classified in two categories: exact and approximate. Exact procedures are applied if approximate procedures do not yield satisfying results or when a theoretical basis for various synthesis approaches is needed.

The theory of nonlinear control systems comprises two basic problems:

1. *Analysis problem* consists of theoretical and experimental research in order to find either the properties of the system or an appropriate mathematical model of the system.

2. *Synthesis problem* consists of determining the structure, parameters, and control system elements in order to obtain desired performance of a nonlinear control system. Further, a mathematical model must be set as well as the technical realization of the model. Since the controlled object is usually known, the synthesis consists of defining a controller in a broad sense.

1.2 Basic Mathematical and Structural Models of Nonlinear Systems

Nonlinear time-variant continuous systems can be described by first-order vector differential equations of the form:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}[t, \mathbf{x}(t), \mathbf{u}(t)] \\ \mathbf{y}(t) &= \mathbf{h}[t, \mathbf{x}(t), \mathbf{u}(t)] \end{aligned} \tag{1.1}$$

or in the case of discrete nonlinear systems, by difference equations:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{f}[k, \mathbf{x}(k), \mathbf{u}(k)] \\ \mathbf{y}(k) &= \mathbf{h}[k, \mathbf{x}(k), \mathbf{u}(k)] \end{aligned} \tag{1.2}$$

where $\mathbf{x}$ is $n \times 1$ state vector, $x \in \Re^n$; $\mathbf{u}$ is $m \times 1$ input vector, $\mathbf{u} \in \Re^m$ and $\mathbf{y}$ is $r \times 1$ output vector, $\mathbf{y} \in \Re^r$.

The existence and uniqueness of solutions are not guaranteed if some limitations on the vector function $\mathbf{f}(t)$ are not introduced. By the solution of differential equation (1.1) in the interval $[0, T]$ we mean such $\mathbf{x}(t)$ that has everywhere derivatives, and for which (1.1) is valid for every t. Therefore, the required limitation is that $\mathbf{f}$ is n-dimensional column vector of nonlinear functions which are locally Lipschitz. With continuous systems, which are of exclusive interest in this text, the function $\mathbf{f}$ associates to each value t, $\mathbf{x}(t)$ and $\mathbf{u}(t)$ a corresponding n-dimensional vector $\dot{\mathbf{x}}(t)$. This is denoted by the following notation:

$$\mathbf{f} : \Re_+ \times \Re^n \times \Re^m \mapsto \Re^n$$

and also:

$$\mathbf{h} : \Re_+ \times \Re^n \times \Re^m \mapsto \Re^r$$

There is no restriction that in the expressions (1.1) only differential equations of the first order are present, since any differential equation of n-th order can be substituted by n differential equations of first order.

DEFINITION 1.2 (LOCALLY LIPSCHITZ FUNCTION)
The function $\mathbf{f}$ is said to be locally Lipschitz for the variable $\mathbf{x}(t)$ if near the point $\mathbf{x}(0) = \mathbf{x}_0$ it satisfies the Lipschitz criterion:

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq k \|\mathbf{x} - \mathbf{y}\| \tag{1.3}$$

for all $\mathbf{x}$ and $\mathbf{y}$ in the vicinity of $\mathbf{x}_0$, where k is a positive constant and the norm is Euclidian[3].

[3] Euclid norm is defined as $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Lipschitz criterion guarantees that (1.1) has an unique solution for an initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$.

The study of the dynamics of a nonlinear system can be done through the input-output signals or through the state variables, as is the case with equation (1.1). Although in the rest of this text only systems with one input and output[4] will be treated, a short survey of the mathematical description of a multivariable[5] nonlinear system will be done. For example, the mathematical model with n state variables $\mathbf{x} = [x_1 x_2 \dots x_n]^T$, m inputs $\mathbf{u} = [u_1 u_2 \dots u_m]^T$ and r outputs $\mathbf{y} = [y_1 y_2 \dots y_r]^T$ can be written:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}[\mathbf{x}(t)] + \sum_{i=1}^{m} \mathbf{g}_i(\mathbf{x})\mathbf{u}_i(t) \\ y_i(t) &= \mathbf{h}_i[\mathbf{x}(t)]; \quad 1 \le i \le r \end{aligned} \tag{1.4}$$

where $\mathbf{f}, \mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_m$ are real functions which map one point $\mathbf{x}$ in the open set $\mathbf{U} \subset \Re^n$, while $h_1, h_2, \dots, h_r$ are also real functions defined on $\mathbf{U}$. These functions are presented as follows:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(x_1,\dots,x_n) \\ f_2(x_1,\dots,x_n) \\ \vdots \\ f_n(x_1,\dots,x_n) \end{bmatrix}; \quad \mathbf{g}_i(x) = \begin{bmatrix} g_{1i}(x_1,\dots,x_n) \\ g_{2i}(x_1,\dots,x_n) \\ \vdots \\ g_{ni}(x_1,\dots,x_n) \end{bmatrix};$$

$$h_i(\mathbf{x}) = h_i(x_1,\dots,x_n)$$

The equations (1.4) describe most physical nonlinear systems met in control engineering, including also linear systems which have the same mathematical model, with the following limitations:

1. $\mathbf{f}(\mathbf{x})$ is a linear function of $\mathbf{x}$, given by $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is a square matrix with the dimension $n \times n$,

2. $\mathbf{g}_1(\mathbf{x}), \dots, \mathbf{g}_m(\mathbf{x})$ are constant functions of $\mathbf{x}$, $\mathbf{g}_i(\mathbf{x}) = \mathbf{b}_i$; $i = 1, 2, \dots, m$, where $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ are column vectors of real numbers with the dimension $n \times 1$.

3. Real functions $h_1(\mathbf{x}), \dots, h_r(\mathbf{x})$ are also linear functions of $\mathbf{x}$, i.e. $h_i(\mathbf{x}) = \mathbf{c}_i\mathbf{x}; i = 1, 2, \dots, r$, where $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_r$ are row vectors of real numbers with the dimension $1 \times n$.

With these limitations — by means of state variables — a linear system can be described with the well-known model:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t); \quad \mathbf{x} \in \Re^n, \mathbf{u} \in \Re^m \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t); \quad \mathbf{y} \in \Re^r \end{aligned} \tag{1.5}$$

[4] For the system with one input and one output the term scalar system is also used.
[5] In the text the term multivariable system means that there is more than one input or/and output.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The input-output mathematical description of the dynamics of the linear system can be obtained by convolution:

$$\mathbf{y}(t)=\int_{-\infty}^{t}\mathbf{W}(t-\tau)\mathbf{u}(\tau)d\tau=\int_{-\infty}^{t_{0-}}\mathbf{W}(t-\tau)\mathbf{u}(\tau)d\tau+\int_{t_{0-}}^{t}\mathbf{W}(t-\tau)\mathbf{u}(\tau)d\tau \tag{1.6}$$

The first integral in the above expression represents a zero-input response, while the second one stands for the zero-state response of a linear system. The weighting matrix under the integral $\mathbf{W}(t)=\mathbf{C}e^{\mathbf{A}\tau}\mathbf{B}$ becomes for scalar systems the weighting function, which is the response of the linear system to unit impulse input $\delta(t)$. The total response of a linear multivariable system can be written:

$$\mathbf{y}(t)=\mathbf{C}e^{\mathbf{A}(t-t_0)}\mathbf{x}(t_0)+\int_{t_{0-}}^{t}\mathbf{C}(\tau)e^{\mathbf{A}(t-\tau)}\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau\ ;\ t\geq t_{0}\text{-} \tag{1.7}$$

With initial conditions equal to zero, $\mathbf{x}(t_{0-})=\mathbf{0}$, Laplace transform of the above equation yields:

$$\mathbf{Y}_{zs}(s)=\mathbf{G}(s)\mathbf{U}(s) \tag{1.8}$$

where $\mathbf{G}(s)$ is transfer function matrix, with the dimension $r\times m$. It is the Laplace transform of the weighting matrix $\mathbf{W}(t)$. The transfer function matrix of the linear system is:

$$\mathbf{G}(s)=\mathbf{C}(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{B} \tag{1.9}$$

The element $g_{ij}(s)$ of the matrix $\mathbf{G}(s)$ represents the transfer function from u_j to y_i, with all other variables equal to zero. If the monic polynomial[6] $g(s)$ is the least common denominator of all the elements of the matrix $\mathbf{G}(s)$:

$$g(s)=s^q+g_1s^{q-1}+\ldots+g_q \tag{1.10}$$

then the transfer function matrix can be written as:

$$\mathbf{G}(s)=\frac{\mathbf{N}(s)}{g(s)} \tag{1.11}$$

where $\mathbf{N}(s)$ is a polynomial[7] matrix with the dimension $r\times m$. If ν is the highest power of the complex variable s which can occur anywhere in the elements of the polynomial matrix $\mathbf{N}(s)$, then $\mathbf{N}(s)$ can be expressed as:

$$\mathbf{N}(s)=\mathbf{N}_0s^{\nu}+\mathbf{N}_1s^{\nu-1}+\ldots+\mathbf{N}_{\nu-1}s+\mathbf{N}_{\nu} \tag{1.12}$$

where $\mathbf{N}_i$ are constant matrices with dimension $r\times m$, while ν represents the degree of the polynomial matrix $\mathbf{N}(s)$. Transfer function matrix $\mathbf{G}(s)$ is strictly proper if every element $g_{ij}(s)$ is a strictly proper transfer function.

[6]Monic polynomial has the coefficient of the highest power equal to one.
[7]Every element of the polynomial matrix is a polynomial.

EXAMPLE 1.1
For a linear system which has following matrices:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & -8 \end{bmatrix}; \quad \mathbf{C} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

the transfer function matrix is:

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} = \begin{bmatrix} \frac{3}{s-1} & \frac{1}{s-1} \\ \frac{1}{s-1} & \frac{-8}{s-2} \end{bmatrix}$$
$$= \frac{1}{(s-1)(s-2)} \begin{bmatrix} 3(s-2) & s-2 \\ s-2 & -8(s-1) \end{bmatrix}$$
$$\mathbf{G}(s) = \frac{1}{s^2 - 3s + 2} \left\{ \begin{bmatrix} 3 & 1 \\ 1 & -8 \end{bmatrix} s + \begin{bmatrix} -6 & -2 \\ -2 & 8 \end{bmatrix} \right\}$$

where:

$$\mathbf{N}_0 = \begin{bmatrix} 3 & 1 \\ 1 & -8 \end{bmatrix}, \quad \mathbf{N}_1 = \begin{bmatrix} -6 & -2 \\ -2 & 8 \end{bmatrix}$$

It must be observed that $g(s)$ differs from the characteristic polynomial of the matrix $\mathbf{A}$, i.e. $g(s) \neq |s\mathbf{I} - \mathbf{A}| = (s-1)^2(s-2)$. Beside such external description by means of a transfer function matrix (1.9) and (1.11), the linear system can be represented by the external matrix factorized description:

$$\mathbf{G}(s) = \mathbf{N}(s)\mathbf{D}_r^{-1}(s) = \mathbf{D}_l^{-1}(s)\mathbf{N}(s) \tag{1.13}$$

where $\mathbf{N}(s)$ is $r \times m$ coprime[8] polynomial matrix, $\mathbf{D}_r(s)$ is $m \times m$ right coprime polynomial matrix, $|\mathbf{D}_r(s)| \neq 0; \mathbf{D}_r(s) = g(s)\mathbf{I}_m$, $\mathbf{D}_l(s)$ is $r \times r$ left coprime polynomial matrix; $|\mathbf{D}_l(s)| \neq 0; \mathbf{D}_l(s) = g(s)\mathbf{I}_r$ and $g(s)$ is monic polynomial of order q, given by (1.10), which is the least common denominator of all the elements of the transfer function matrix $\mathbf{G}(s)$.

The expression (1.13) is called the right, respectively the left, polynomial factorized representation, depending whether the right or the left coprime polynomial matrix is used. Polynomial matrix $\mathbf{N}(s)$ is the numerator while $\mathbf{D}_r(s)$ or $\mathbf{D}_l(s)$ are called denominators. The notations follow as a natural generalization of the scalar systems for which $\mathbf{G}(s)$ is a rational function of the complex variable s, while $\mathbf{N}(s)$ and $\mathbf{D}_r(s)$ or $\mathbf{D}_l(s)$ are scalar polynomials in the numerator and denominator, respectively.

By applying inverse Laplace transform to the expression (1.13), with left polynomial factorized representation, a set of n-th order differential equations with input-output signals is obtained:

$$\mathbf{D}_l(p)\mathbf{y}_{zs}(t) = \mathbf{N}(p)\mathbf{u}(t) \tag{1.14}$$

[8]Elements of a coprime polynomial matrix are coprime polynomials or polynomials without a common factor.

where $p = \frac{d}{dt}$. The expression (1.14) is an external representation of the linear multivariable system (1.5). For scalar systems (1.14) becomes:

$$\begin{aligned}(a_n p^n + a_{n-1} p^{n-1} + \ldots + a_0) y(t) &= (b_m p^m + b_{m-1} p^{m-1} + \ldots + b_0) u(t) \\ \left[a_n(t) p^n + a_{n-1}(t) p^{n-1} + \ldots + a_0(t)\right] y(t) &= \left[b_m(t) p^m + b_{m-1}(t) p^{m-1} + \ldots + \right. \\ &\quad \left. + b_0(t)\right] u(t)\end{aligned} \tag{1.15}$$

where a_i and b_j, $i = 0, 1, \ldots, n$; $j = 0, 1, 2, \ldots, m$, are constant coefficients, while with time-variant linear systems $a_i(t)$ and $b_j(t)$ are time-variant coefficients. Analysis of the linear time-variant systems (1.15) is much more cumbersome. Namely, applying Laplace or Fourier transform to convolution of two time functions leads to the convolution integral with respect to the complex variable. In general case, solving the equations with variable parameters is achieved in the time domain.

From the presented material, it follows that for linear systems there exist several possibilities to describe their dynamics. Two basic approaches are:

1. External description by input-output signals of a linear system, given by (1.7), (1.8), (1.11), (1.13) and (1.14),

2. Internal description by state variables of the linear system, given by (1.5).

These internal descriptions are not unique, for example they depend on the choice of state variables. However, they may reliably describe the performance of the linear system within a certain range of input signals and initial conditions changes. The conversion from internal to external description is a direct one, as given in (1.9).

If the differential equation for a given system cannot be written in the forms of (1.14) or (1.15), the system is nonlinear. Solving the nonlinear equations of higher order is usually very difficult and time-consuming, because a special procedure is needed for every type of the nonlinear differential equation. Moreover, nonlinear relations are often impossible to express mathematically; the nonlinearities are in such cases presented nonanalytically. The basic difficulty in solving nonlinear differential equations lies in the fact that in nonlinear systems a separate treatment of stationary and transient states is meaningless, since the signal cannot be decomposed to separately treated components; the signal must be taken as a whole. Like linear systems, the dynamics of nonlinear systems can be described by external and internal descriptions, or rather with mathematical model. The internal description (1.1) and (1.4) is the starting point to find the external description, which — like the linear systems — can have several forms.

As the dynamics of the nonlinear systems is "richer" than in the case of linear ones, the mathematical models must be consequently more complex. Nonlinear expansion of the description with the weighting matrix is known as the description with Wiener-Volterra series. Modified description (1.1) by means of a series is also Fliess series expansion of the functionals, which can be used for the description of the nonlinear system. In further text neither Wiener-Volterra series nor Fliess series expansion of functionals will be mentioned, and the interested reader can consult the books of Nijmeijer and Van der Schaft (1990) and Isidori (1995).

DEFINITION 1.3 (FORCED AND UNFORCED SYSTEM)
A continuous system is said to be forced if an input signal is present:

$$\dot{\mathbf{x}}(t) = \mathbf{f}[t, \mathbf{x}(t), \mathbf{u}(t)]; \forall t \geq 0, \mathbf{u}(t) \neq \mathbf{0} \tag{1.16}$$

A continuous system is said to be unforced if there is no input (excitation) signal:

$$\dot{\mathbf{x}}(t) = \mathbf{f}[t, \mathbf{x}(t)]; \forall t \geq 0, \mathbf{u}(t) = \mathbf{0} \tag{1.17}$$

Under the excitation of the system, all external signals such as measurement noise, disturbance, reference value, etc. are understood. All such signals are included in the input vector $\mathbf{u}(t)$. A clear difference between a forced and an unforced system does not exist, since a function $\mathbf{f}_u : \Re_+ \times \Re^n \mapsto \Re^n$ can be always defined with $\mathbf{f}_u(t, \mathbf{x}) = \mathbf{f}[t, \mathbf{x}(t), \mathbf{u}(t)]$ so that a forced system ensues $\dot{\mathbf{x}}(t) = \mathbf{f}_u[t, \mathbf{x}(t)]; \forall t \geq 0$.

DEFINITION 1.4 (TIME-VARYING AND TIME-INVARIANT SYSTEM)
A system is time-invariant[9] *if the function* $\mathbf{f}$ *does not explicitly depend on time, i.e. the system can be described with:*

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t)]; \ \forall t \geq 0 \tag{1.18}$$

A system is time-varying[10] *if the function* $\mathbf{f}$ *explicitly depends on the time, i.e. the system can be described with:*

$$\dot{\mathbf{x}}(t) = \mathbf{f}[t, \mathbf{x}(t), \mathbf{u}(t)]; \ \forall t \geq 0 \tag{1.19}$$

DEFINITION 1.5 (EQUILIBRIUM STATE)
An equilibrium state can be defined as a state which the system retains if no external signal $\mathbf{u}(t)$ *is present. Mathematically, the equilibrium state is expressed by the vector* $\mathbf{x}_e \in \Re^n$*. The system remains in equilibrium state if at the initial moment the state was the equilibrium state, i.e. at* $t = t_{0^-}$ $\mathbf{x}(t_0) = \mathbf{x}_e$*. In other words, if the system begins from the equilibrium state, it remains in this state* $\mathbf{x}(t) = \mathbf{x}_e, \forall t \geq t_0$ *under presumption that no external signal acts, i.e.* $\mathbf{u}(t) = 0$.

[9] In the literature the term autonomous is also used for time-invariant systems. The authors decided to use the latter term for such systems. In Russian literature the term autonomous is used for unforced nonlinear systems.

[10] For such system the term nonautonomous is sometimes used.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

It is common to choose the origin of the coordinate system as the equilibrium state, i.e. $\mathbf{x}_e = \mathbf{0}$. If this is not the case, and the equilibrium state is somewhere else in $\Re^n$, translation of the coordinate system $\mathbf{x}' = \mathbf{x} - \mathbf{x}_e$ will bring any such point to the origin. For an unforced continuous system, the equilibrium state will be $\mathbf{f}(t, \mathbf{x}_e, \mathbf{0}) = \mathbf{0}; \forall t \geq t_0$. Unforced discrete-time systems described by $\mathbf{x}(k+1) = \mathbf{f}[k, \mathbf{x}(k), \mathbf{u}(k)]$ will have the equilibrium state given by $\mathbf{f}(k, \mathbf{x}_e, \mathbf{0}) = \mathbf{x}_e; \forall k \geq 0$.

DEFINITION 1.6 (SYSTEM'S TRAJECTORY—SOLUTION)
An unforced system is described by the vector differential equation:

$$\dot{\mathbf{x}}(t) = \mathbf{f}[t, \mathbf{x}(t)]\,; \forall t \geq 0 \tag{1.20}$$

where $\mathbf{x} \in \Re^n$, $t \in \Re_+$ *and continuous function* $\mathbf{f}: \Re_+ \times \Re^n \mapsto \Re^n$.

It is supposed that (1.20) has a unique trajectory (solution), which corresponds to a particular initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$, where $\mathbf{x}_0 \neq \mathbf{x}_e$. The trajectory can be denoted by $\mathbf{s}(t, t_0, \mathbf{x}_0)$, and the system's state will be described from the moment t_0 as:

$$\mathbf{x}(t) = \mathbf{s}(t, t_0, \mathbf{x}_0); \ \forall t \geq t_0 \geq 0 \tag{1.21}$$

where the function $\mathbf{s}: \Re_+ \times \Re^n \mapsto \Re^n$. If the trajectory is a solution, it must satisfy its differential equation:

$$\dot{\mathbf{s}}(t, t_0, \mathbf{x}_0) = \mathbf{f}[t, \mathbf{s}(t, t_0, \mathbf{x}_0)]; \ \forall t \geq 0; \ \mathbf{s}(t_0, t_0, x_0) = x_0 \tag{1.22}$$

Trajectory (solution) has the following properties:

(a) $\mathbf{s}(t_0, t_0, \mathbf{x}_0) = \mathbf{x}_0; \ \forall \mathbf{x}_0 \in \Re^n$

(b) $\mathbf{s}[t, t_1, \mathbf{s}(t_1, t_0, \mathbf{x}_0)] = \mathbf{s}(t, t_0, \mathbf{x}_0); \ \forall t \geq t_1 \geq t_0 \geq 0; \ \forall \mathbf{x}_0 \in \Re^n$

The existence of the solution of the nonlinear differential equation (1.20) is not guaranteed if appropriate limitations on the function $\mathbf{f}$ are not introduced. The solution of the equation in the interval $[0, T]$ means that $\mathbf{x}(t)$ is everywhere differentiable, and that differential equation (1.20) is valid for every t. If vector $\mathbf{x}(t)$ is the solution of (1.20) in the interval $[0, T]$, and if $\mathbf{f}$ is a vector of continuous functions, then $\mathbf{x}(t)$ statisfies the integral equation:

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}[\tau, \mathbf{x}(\tau)]\, d\tau\,; \ t \in [0, T] \tag{1.23}$$

Equations (1.20) and (1.23) are equivalent in the sense that any solution of (1.20) is at the same time the solution of (1.23) and vice versa.

Local conditions which assure that (1.20) has an unique solution within a finite interval $[0, \tau]$ whenever τ is small enough are given by the following theorem:

THEOREM 1.1 (LOCAL EXISTENCE AND UNIQUENESS)
With the presumption that $\mathbf{f}$ *from (1.20) is continuous on* t *and* $\mathbf{x}$, *and finite constants* h, r, k *and* T *exist, so that:*

$$\|\mathbf{f}(t,\mathbf{x})-\mathbf{f}(t,\mathbf{y})\| \leq k\|\mathbf{x}-\mathbf{y}\|;\ \forall \mathbf{x},\mathbf{y} \in \mathbf{B}, \forall t \in [0,T] \tag{1.24}$$

$$\|\mathbf{f}(t,\mathbf{x}_0)\| \leq h;\ \forall t \in [0,T] \tag{1.25}$$

where $\mathbf{B}$ *is a ball in* $\Re^n$ *of the form:*

$$\mathbf{B} = \{\mathbf{x} \in \Re^n; \|\mathbf{x}-\mathbf{x}_0\| \leq r\} \tag{1.26}$$

then (1.20) has just one solution in the interval $[0,\tau]$ *where the number* τ *is sufficiently small to satisfy the inequalities:*

$$h\tau e^{k\tau} \leq r \tag{1.27}$$

$$\tau \leq \min\left\{T, \frac{\rho}{k}, \frac{r}{h+kr}\right\} \tag{1.28}$$

for some constant $\rho < 1$

Proof of the theorem is in Vidyasagar (1993, p. 34).
Remarks:

1. Condition (1.24) is known as the Lipschitz condition, and k as the Lipschitz constant. If k is a Lipschitz constant for the function $\mathbf{f}$, then so is any constant greater than k.
2. A function $\mathbf{f}$ which statisfies the Lipschitz condition is a Lipschitz continuous function. It is also absolutely continuous, and is differentiable almost anywhere.
3. Lipschitz condition (1.24) is known as a local Lipschitz condition since it holds only for $\forall \mathbf{x}$, $\mathbf{y}$ in a ball around $\mathbf{x}_0$ for $\forall t \in [0,T]$.
4. For given finite constants h, r, R, T, the conditions (1.27) and (1.28) can be always satisfied with a sufficiently small τ.

THEOREM 1.2 (GLOBAL EXISTENCE AND UNIQUENESS)
Suppose that for each $T \in [0,\infty\rangle$ *there exist finite constants* h_T *and* k_T *such that*

$$\|\mathbf{f}(t,\mathbf{x})-\mathbf{f}(t,\mathbf{y})\| \leq k_T\|\mathbf{x}-\mathbf{y}\|;\ \forall \mathbf{x},\mathbf{y} \in \Re^n, \forall t \in [0,T] \tag{1.29}$$

$$\|\mathbf{f}(t,\mathbf{x}_0)\| \leq h_T;\ \forall t \in [0,T] \tag{1.30}$$

Then (1.20) has only one solution in the interval $[0,\infty]$.

Proof of the theorem is in Vidyasagar (1993, p. 38).

It can be demonstrated that all differential equations with continuous functions $\mathbf{f}$ have an unique solution (Vidyasagar, 1993, p. 469), and that Lipschitz continuity does not depend upon the norm in $\Re^n$ which is used (Vidyasagar, 1993, p. 45).

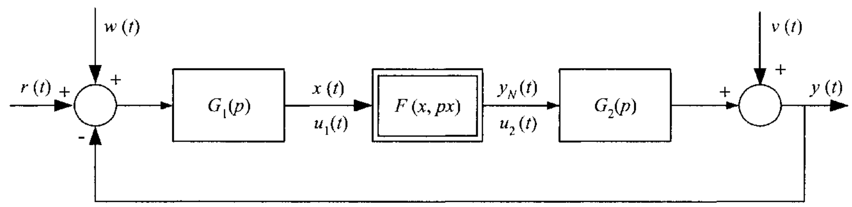


Figure 1.1: Block diagram of a nonlinear system.

Basic Structures of Nonlinear Systems

A broad group of nonlinear systems are such that they can be formed by combining a linear and a nonlinear part, Fig. 1.1. The notations in Fig. 1.1 are:

$r(t)$ – reference signal or set-point value,

$y(t)$ – output signal, response of a closed loop control system to $r(t)$,

$w(t)$ – external or disturbance signal[11],

$x(t), u_1(t)$ – input to nonlinear part of the system[12],

$y_N(t), u_2(t)$ – output of nonlinear part of the system[13],

$v(t)$ – random quantity, noise present by measurement of output quantity,

$G_1(p), G_2(p)$ – mathematical models of linear parts of the system,

$F(x, px)$ – mathematical description of the nonlinear part of the system,

$p = \frac{d}{dt}$ – derivative operator.

The block diagram of the system in Fig. 1.1 can be reduced to the block diagram in Fig. 1.2, which is applicable for the approximate analysis of a large number of nonlinear systems.

[11] Various notation will be used for this signal, depending on the external signal. The so-called dither signal will be $d(t)$ instead of $w(t)$, while the harmonic signal for forced oscillations will be $f(t)$ instead of $w(t)$.

[12] With $x(t)$ is denoted input signal to the nonlinear part — the same notation is reserved for state variables. Which one is meant will be obvious from the context.

[13] For simplicity, the output signal of nonlinear element will also be denoted by $y(t)$ if only the nonlinear element is considered. In the case when closed-loop system is discussed, the output of nonlinear element is denoted by $y_n(t)$ or $u(t)$, while the output signal of the closed loop will be $y(t)$. From the context it will be clear which signal is $y(t)$ or $y_n(t)$, respectively.

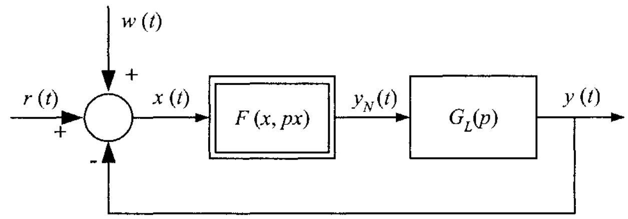


Figure 1.2: Simplified block diagram of a nonlinear system.

The block diagram in Fig. 1.2 is presented by the dynamic equation of nonlinear systems:

$$f(t) = x(t) + G_L(p)F(x, px) \tag{1.31}$$

respectively:

$$A(p)f(t) = A(p)x(t) + B(p)F(x, px) \tag{1.32}$$

where $f(t) = r(t) + w(t)$, for $v(t) = 0$ and $G_L(p) = \frac{A(p)}{B(p)}$.

1.3 Basic Specific Properties of Nonlinear Systems

Dynamic behavior of a linear system is determined by the general solution of the linear differential equation of the system. In a system with time-invariant (constant) parameters, the form and the behavior of the output signal doesn't depend upon the magnitude of the excitation signals — also the magnitude of the output signal of a stable system in the stationary regime does not depend upon the initial conditions.

However, dynamic performance of a nonlinear system depends on the system's parameters and initial conditions, as well as on the form and the magnitude of external actions. The basic solutions of the differential equation of a nonlinear system are generally complex and various. Of special interest are:

Unboundedness of Reaction in Finite Time Interval

The output signal of an unstable linear system increases beyond boundaries, when $t \longrightarrow \infty$. With a nonlinear system, the output signal can increase unboundedly in finite time. For example, the output signal of the nonlinear system described by the differential equation of the first order $\dot{x} = x^2$ with initial condition $x(0) = x_0$ tends to infinity for $t = 1/x_0$.

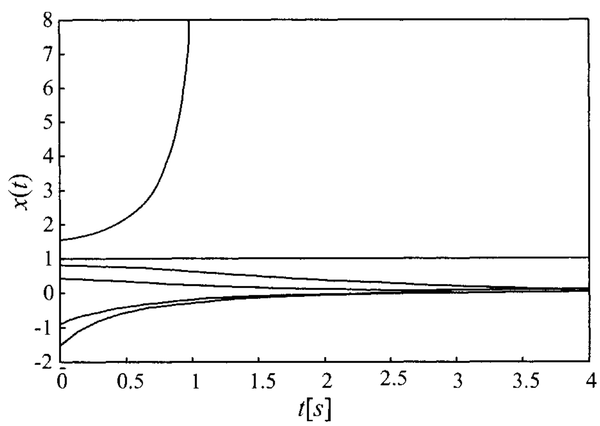


Figure 1.3: State trajectories of a nonlinear system in Example 1.2.

Equilibrium State of Nonlinear System

Linear stable systems have one equilibrium state. As an example, the response of a linear stable system to the unit pulse input is damped towards zero.

Nonlinear stable systems may possess several equilibrium states, i.e. a possible equilibrium state is determined by a system's parameters, initial conditions and the magnitudes and forms of external excitations.

EXAMPLE 1.2
(DEPENDENCY OF EQUILIBRIUM STATE ON INITIAL CONDITIONS)
The first-order nonlinear system described by $\dot{x}(t) = -x(t) + x^2(t)$ *with initial conditions* $x(0) = x_0$, *has the solution (trajectory) given by:*

$$x(t) = \frac{x_0 e^{-t}}{1 - x_0 + x_0 e^{-t}}$$

Depending on initial conditions, the trajectory can end in one of two possible equilibrium states $x_e = 0$ *and* $x_e = 1$, *as illustrated in Fig. 1.3.*

For all initial conditions $x(0) > 1$, *the trajectories will diverge, while for* $x(0) < 1$ *they will approach the equilibrium state* $x_e = 0$. *For* $x(0) = 1$ *the trajectory will remain constant* $x(t) = 1, \forall t$, *and the equilibrium state will be* $x_e = 1$. *Therefore, this nonlinear system has one stable equilibrium state (*$x_e = 0$*) and one unstable equilibrium state, while the equilibrium state* $x_e = 1$ *can be declared as neutrally stable*[14]. *The linearization of this nonlinear system for* $|x| < 1$ *(by dis-*

[14]Exact definitions of the stability of the equilibrium state will be discussed in Chapter 2.

carding the nonlinear term) yields $\dot{x}(t) = -x(t)$ *with the solution (state trajectory)* $x(t) = x(0)e^{-t}$. *This shows that the unique equilibrium state* $x_e = 0$ *is stable for all initial conditions, since all the trajectories — notwithstanding initial conditions — will end at the origin.*

EXAMPLE 1.3 (DEPENDENCE OF EQUILIBRIUM STATE ON THE INPUT)
A simplified mathematical model of an unmanned underwater vehicle in forward motion is:

$$m\ddot{x}(t) + d\,|\dot{x}(t)|\,\dot{x}(t) = \tau(t)$$

where $v(t) = \dot{x}(t)$ *is the velocity of the underwater vehicle and* $\tau(t)$ *is the thrust of the propulsor. The nonlinear term* $|v(t)|\,v(t)$ *is due to the hydrodynamic effect, the so-called added inertia. With* $m = 200[kg]$ *and* $d = 50[kg/m]$, *and with a pulse for the thrust (Fig. 1.4a), the velocity of the underwater vehicle will be as in Fig. 1.4b. The diagram reveals the fact that the velocity change is faster with the increase of thrust than in reverse case — this can be explained by the inertial properties of the underwater vehicle. If the experiment is repeated with ten times greater thrust (Fig. 1.5a), the change of velocity will be as in Fig. 1.5b. The velocity has not increased ten times, as would be the case with a linear system. Such behavior of an unmanned underwater vehicle must be taken into account when the control system is designed. If the underwater vehicle is in the mission of approaching an underwater fixed mine, the thrust of the underwater vehicle must be appropriately regulated, in order that the mission be successful.*

Self-Oscillations — Limit Cycles

The possibility of undamped oscillations in a linear time-invariant system is linked with the existence of a pair of poles on the imaginary axis of the complex plane. The amplitude of oscillations is in this case given by initial conditions.

In nonlinear systems it is possible to have oscillations with amplitude and frequency which are not dependent upon the value of initial conditions, but their occurrence depends upon these initial conditions. Such oscillations are called *self-oscillations* (*limit cycles*)[15] and they belong to one of several stability concepts of the dynamic behavior of nonlinear systems.

Subharmonic, Harmonic or Periodic and Oscillatory Processes with Harmonic Inputs

In a stable linear system, a sinusoidal input causes a sinusoidal output of the same frequency. A nonlinear system under a sinusoidal input can produce an unexpected response. Depending on the type of nonlinearity, the output can be a signal with a frequency which is:

[15] The term limit cycles will be used for the type of a singular point, when the state (or phase) trajectory enters this shape in case of self-oscillatory behavior of a system.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

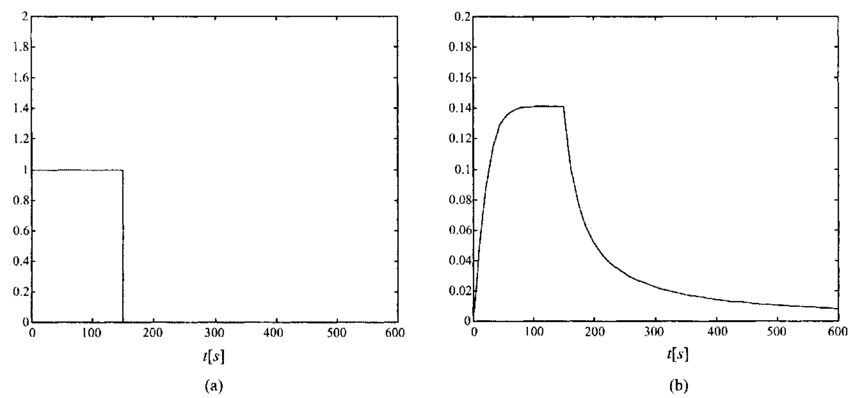


Figure 1.4: Thrust (a) and velocity (b) of an underwater vehicle.

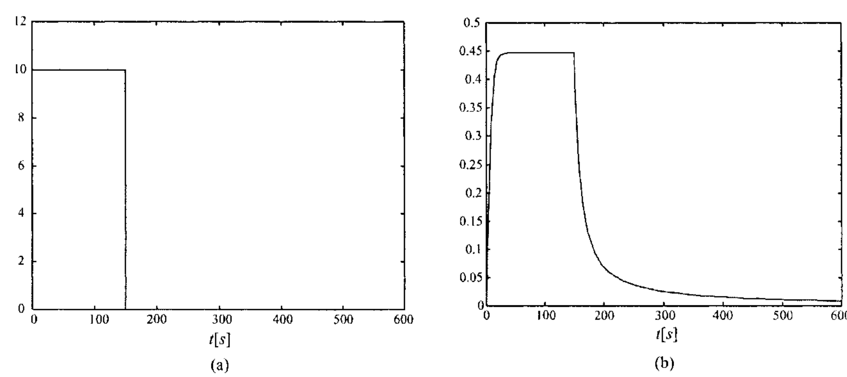


Figure 1.5: Thrust (a) and velocity (b) of an underwater vehicle.

- generally proportional to the input signal frequency,
- higher harmonic of the input signal frequency,
- a periodic signal independent of the input signal frequency, or
- a periodic signal with the same frequency as the input signal frequency.

With the input harmonic signal, the output signal can be a harmonic, subharmonic or periodic signal, depending upon the form, the amplitude and the frequency of the input signal.

Nonuniqueness of Dynamic Performance

In some cases, in a nonlinear system a pulse (short-lived) excitation provokes a response which — under certain initial conditions (energy of the pulse) — tends to one or more stable equilibrium states.

Resonance Jump

Resonance jump was investigated more in the theory of oscillation of mechanical systems than in the theory of control systems. The term "resonance jump" is used in case of a sudden jump of the amplitude and/or phase and/or frequency of a periodic output signal of a nonlinear system. This happens due to a nonunique relation which exists between periodic forcing input signal acting upon a nonlinear system and the output signal from that system. It is believed that resonance jump occurs in nonlinear control systems with small stability phase margin, i.e. with small damping factor of the linear part of the system and with amplitudes of excitation signal that force the system into the operating modes where nonlinear laws are valid, particularly saturation. Higher performance indices such as maximal speed of response with minimal stability degradation, high static and dynamic accuracy, minimal oscillatory dynamics and settling time as well as power efficiency and limitations (durability, resistivity, robustness, dimensions, weight, power) boil down to a higher bandwidth of the system. Thanks to that, input signals can have higher frequency content in them and in some situations can approach closer to the "natural" frequency of the system. As a consequence this favors the occurrence of the resonance jump. Namely, the fulfillments of the aforementioned conditions can bring the forced oscillation frequency (of the nonlinear system operating in the forced oscillations mode) closer to the limit cycle frequency of some subsystems, which together with the amplitude constraints creates conditions for establishing the nonlinear resonance. Resonance jump can occur in nonlinear systems operating in forced oscillations mode and is often not a desirable state of the system. Resonance jump can not be seen from the transient response of the system and can not be defined by solving nonlinear differential equations. It is also not recommended to use experimental tests in plant(s) during operation in order to resolve

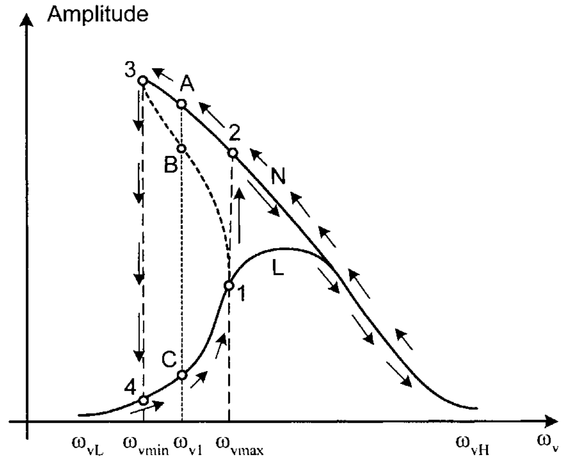


Figure 1.6: Frequency characteristics of linear (L) and nonlinear (N) system.

if the system might have this phenomenon. For that purpose it is best to use the frequency and simulation methods. To reduce or eliminate the resonance jump, higher stability phase margin is needed as well as the widening of the operating region of a nonlinear part of the system where the linear laws are valid.

Synchronization

When the control signal is a sinusoidal signal of "small" amplitude, the output of a nonlinear system may contain subharmonic oscillations. With an increase of the control signal amplitude, it is possible that the output signal frequency "jumps" to the control signal frequency, i.e. synchronization of both input and output frequencies occurs.

Bifurcation[16]

Stability and the number of equilibrium states of a nonlinear system may change as a result either of changing system parameters or of disturbances. Structurally stable systems have the desirable property that small changes in the disturbance yield small changes in the system's performance also. All stable linear systems are structurally stable. This is not the case with nonlinear systems. If there exist

[16] The term originates from the latin *bi* + *furca* – meaning pitchfork, bifurcation, branched. The word was first used by Poincaré to describe the phenomenon when the behavior branches in two different directions in bifurcation point.

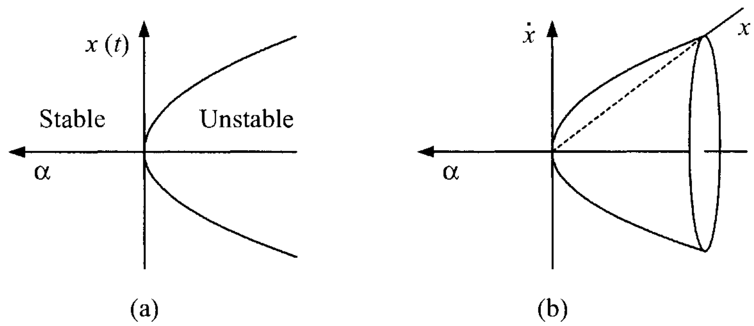


Figure 1.7: Fork bifurcation (a) and Hopf bifurcation (b).

points in the space of system parameters where the system is not structurally stable, such points are called bifurcation points, since the system's performance "bifurcates". The values of parameters at which a qualitative change of the system's performance occurs are called critical or bifurcational values. The theory which encompasses the bifurcation problems is known as bifurcation theory (Guckenheimer and Holmes, 1983).

The system described by Duffing's equation $\ddot{x}(t) + \alpha x(t) + x^3(t) = 0$ has equilibrium points which depend on the parameter α. As α varies from positive to negative values, one equilibrium point (for $\alpha < 0$) is split into three equilibrium points ($x_e = 0, \sqrt{\alpha}, \sqrt{-\alpha}$), as illustrated in Fig. 1.7a. The critical bifurcational value here is $\alpha = 0$, where a qualitative change of system's performance occurs. Such a form of bifurcation is named pitchfork bifurcation because the form of equilibrium points reminds one of a pitchfork. Hopf bifurcation is illustrated in Fig. 1.7b. Here a pair of conjugate complex eigenvalues $\lambda_{1,2} = \sigma \pm j\omega$ crosses from the left to the right half-plane, and the response of the unstable system diverges toward self-oscillations (limit cycle). A very simple example of a system with bifurcation is an inverted rigid pendulum where a spring sustains it in a vertical position. If the mass at the top of the pendulum is sufficiently small, a small disturbance from the vertical position will be corrected by a spring when the disturbance disappears. With greater mass, a critical value m_0 is reached, dependent on the spring constant and length of the pendulum. For $m > m_0$, the spring can no longer correct a small disturbance — the pendulum will move to the left or right side, and after several oscillations will end up at the equilibrium state which is opposite to the starting vertical equilibrium state. In this case m_0 is the bifurcation point, and the branches of the pitchfork bifurcation are to the left or right side along which the pendulum reaches a new equilibrium point, which is—contrary to Fig. 1.7a — stable for both branches.

If the bifurcation points exist in a nonlinear control system, it is important to know the regions of structural stability in the parameter plane and in the phase plane and it is necessary to ensure that the parameters and the states of the system remain within these regions. Structural instability can be expected with control systems whose objects (processes) are nonlinear, with certain types of adaptive control systems, and generally with the systems whose action and reaction forces cannot reach the equilibrium state. If a bifurcation appears, the system can come to a chaotic state — this is of mostly theoretical interest since for safety reasons every control system has built-in activities which prevent any such situation. The consequence of bifurcation can be the transfer to a state with unbounded behavior. For the majority of technical systems this can lead to serious damages if the system has no built in protection.

Chaos

With stable linear systems, small variations of initial conditions can result in small variations in response. Not so with nonlinear systems—small variations of initial conditions can cause large variations in response. This phenomenon is named *chaos*. It is a characteristic of chaos that the behavior of the system is unexpected, an entirely deterministic system which has no uncertainty in the modes of the system, excitation or initial conditions yields an unexpected response. Researchers of this phenomenon have shown that some order exists in chaos, or that under chaotic behavior there are some nice geometric forms which create randomness. The same is true for a card player when he deals the cards or when a mixer mixes pastry for cakes. On one side the discovery of chaos enables one to place boundaries on the anticipation, and on the other side the determinism which is inherent to the chaos implies that many random phenomena can be better foreseen as believed. The research into chaos enables one to see an order in different systems which were up to now considered unpredictable. For example the dynamics of the atmosphere behaves as a chaotic system which prevents a long-range weather forecast. Similar behaviors include the turbulence flow in fluid dynamics, or the rolling and capsizing of a ship in stormy weather. Some mechanical systems (stop-type elements, systems with aeroelastic dynamics, etc.) as well as some electrical systems possess a chaotic behavior.

EXAMPLE 1.4
(CHAOTIC SYSTEM—CHUA ELECTRIC CIRCUIT (MATSUMOTO, 1987))
A very simple deterministic system with just a few elements can generate chaotic behavior, as the example of Chua electric circuit shows.

The schematic of the Chua circuit is given in Fig. 1.8a, while the nonlinear current vs. voltage characteristic of a nonlinear resistor is given in Fig. 1.8b. With nominal values of the components: $C_1 = 0.0053[\mu F]$, $C_2 = 0.047[\mu F]$, $L =$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

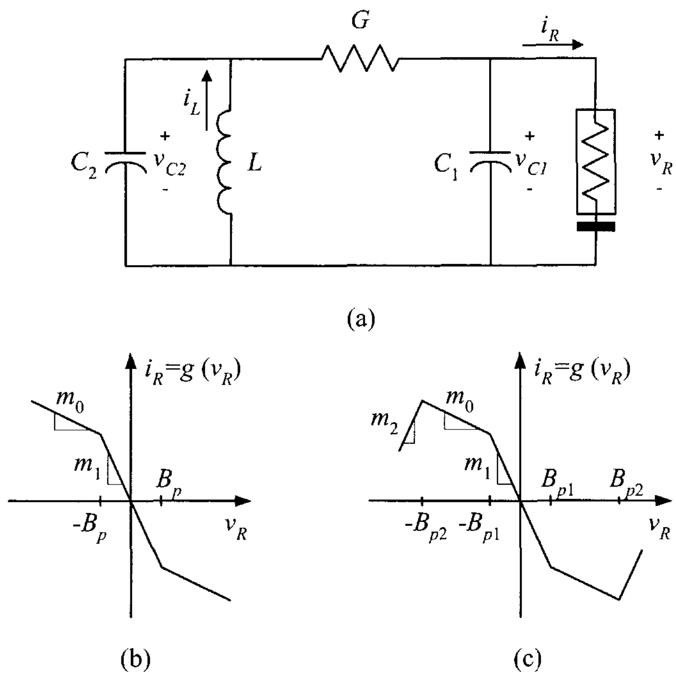


Figure 1.8: Electrical scheme of Chua circuit (a), nonlinear characteristic of the resistor (b), modified resistor characteristic (c).

6.8[mH] and $R = 1.21[k\Omega]$, *the dynamics of the electric circuit can be described by the following mathematical model:*

$$C_1 \frac{dv_{C_1}}{dt} = G(v_{C_2} - v_{C_1}) - g(v_{C_1})$$
$$C_2 \frac{dv_{C_2}}{dt} = G(v_{C_1} - v_{C_2}) + i_L$$
$$L \frac{di_L}{dt} = -v_{C_2}$$

Simulation with normalized parameters $1/C_1 = 9$; $1/C_2 = 1$; $1/L = 7$; $G = 0.7$; $m_0 = -0.5$; $m_1 = -0.8$; $B_p = 1$, *and with initial conditions* $i_L(0) = 0.3$, $v_{c1}(0) = -0.1$ *and* $v_{c2}(0) = 0.5$ *gives the state trajectory of this nonlinear system in Fig. 1.9 (initial state is denoted with point A). The equilibrium state with such attraction of all trajectories has the appropriate name* — attractor. *In the Chua electric circuit the attractor is of the type double scroll attractor (Fig. 1.9). Such a circuit*

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

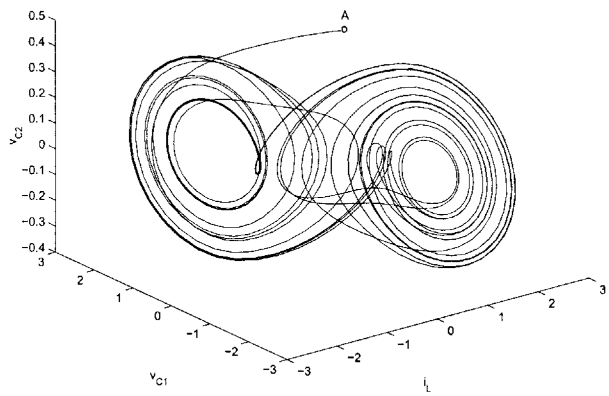


Figure 1.9: State trajectory (i_L, v_{C1}, v_{C2}) of Chua circuit in chaos.

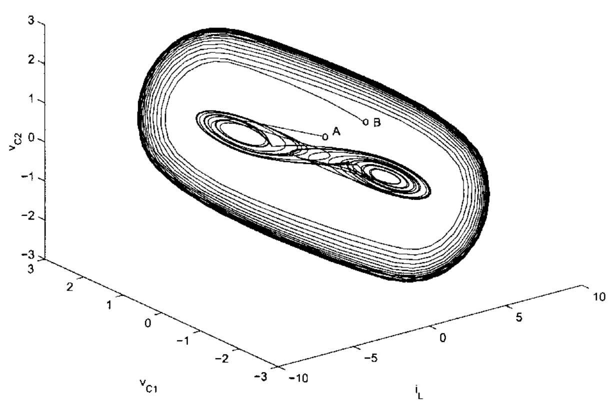


Figure 1.10: State trajectories (i_L, v_{C1}, v_{C2}) of Chua circuit with modified nonlinear resistor characteristics.

with given initial conditions behaves as a chaotic system. Depending on initial conditions, a chaos is formed, and it is inherent to such systems that the trajectory never follows the same path (Chua, Kamoro and Matsumoto, 1986).

By introducing a small modification into the characteristic of the nonlinear resistor, i.e. adding the positive resistance segments as shown in Fig. 1.8c, other interesting aspects of nonlinear systems can be observed. In Fig. 1.10 the trajectory of such system is shown for the same initial conditions as in previous case (point A) where the behavior of the system is chaotic. If we change the initial conditions to $i_L(0) = 3$, $v_{c1}(0) = -0.1$ *and* $v_{c2}(0) = 0.6$ *(point B) the system settles to the stable limit cycle outside the chaotic attractor region. Here one can observe qualitatively completely different behaviors of the nonlinear system which depend only on initial conditions.*

1.4 Stability and Equilibrium States

Very often the dynamic behavior of a system is characterized by the phrase "the system is stable". Under the concept of system stability, in practice it is understood that by small variations of input signals, initial conditions or parameters of the system, the state of the system does not have large deviations, i.e. these are minimal requirements which the system must satisfy. Whereas the linear systems have only one[17] possible stable equilibrium state, nonlinear systems can generally have several stable equilibrium states and dynamic performances.

DEFINITION 1.7 (ATTRACTIVE EQUILIBRIUM STATE — ATTRACTOR)
Equilibrium state $\mathbf{x}_e$ *is attractive if for every* $t_0 \in \Re_+$, *there exists a number* $\eta(t_0) > 0$ *such that:*

$$\|x_0\| < \eta(t_0) \Rightarrow \mathbf{s}(t_0+t, t_0, \mathbf{x}_0) \longrightarrow \mathbf{x}_e \text{ as } t \longrightarrow \infty \tag{1.33}$$

DEFINITION 1.8 (UNIFORMLY ATTRACTIVE EQUILIBRIUM STATE)
Equilibrium state x_e *is uniformly attractive if for every* $t_0 \in \Re_+$, *there exists a number* $\eta > 0$ *so that the following condition is true:*

$$\|\mathbf{x}_0\| < \eta,\ t_0 \geq 0 \Rightarrow \mathbf{s}(t_0+t, t_0, \mathbf{x}_0) \longrightarrow \mathbf{x}_e \text{ as } t \longrightarrow \infty \text{ uniformly in } x_0 \text{ and } t_0 \tag{1.34}$$

Attraction means that for every initial moment $t_0 \in \Re_+$, the trajectory which starts sufficiently close to the equilibrium state $\mathbf{x}_e$ approaches to this equilibrium state as $t_0 + t \longrightarrow \infty$. A necessary (but not sufficient) condition for the equilibrium state to be attractive is that it is an isolated equilibrium state, i.e. that in the vicinity there are no other equilibrium states.

[17] Linear systems with more than one equilibrium state are exceptions. For example, the system $\ddot{x} + \dot{x} = 0$ has infinite number of equilibrium states (signal points) on the real axis.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

DEFINITION 1.9 (REGION OF ATTRACTION)
If an unforced nonlinear time-invariant system described by $\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t)]$ *has an equilibrium state* $\mathbf{x}_e = \mathbf{0}$ *which is attractive, then the region of attraction* $\mathbf{D}(\mathbf{x}_e)$ *is defined by:*

$$\mathbf{D}(\mathbf{0}) = \{\mathbf{x}_0 \in \Re^n : \mathbf{s}(t,0,\mathbf{x}_0) \to \mathbf{0} \text{ when } t \to \infty\} \tag{1.35}$$

Every attractive equilibrium state has its own region of attraction. Such a region in state space means that trajectories which start from any initial condition inside the region of attraction are attracted by the attractive equilibrium state.

The attractive equilibrium states are called *attractors*. The equilibrium states with repellent properties have the name *repellers*, while the equilibrium states which attract on one side and repel on the other side are *saddles*.

EXAMPLE 1.5
(ATTRACTORS AND REPELLERS (HARTLEY, BEALE AND CHICATELLI, 1994))
A nonlinear system of first order is described by $\dot{x}(t) = f(x,u) = x - x^3 + u$. *With the input signal* $u(t) = 0$, *the equilibrium states of the system will be at the points where* $\dot{x}(t) = 0$ *and* $f(x) = x - x^3 = 0$, *respectively (Fig. 1.11).*

Equilibrium states are $x_{e1} = 0$, $x_{e2} = +1$ *and* $x_{e3} = -1$. *Local stability of these states can be determined graphically from the gradient (slope of the tangent) at these points. So at every equilibrium point:*

$$J = \left[\frac{\partial \dot{x}}{\partial x}\right]_{x_e} = 1 - 3x_e^2$$

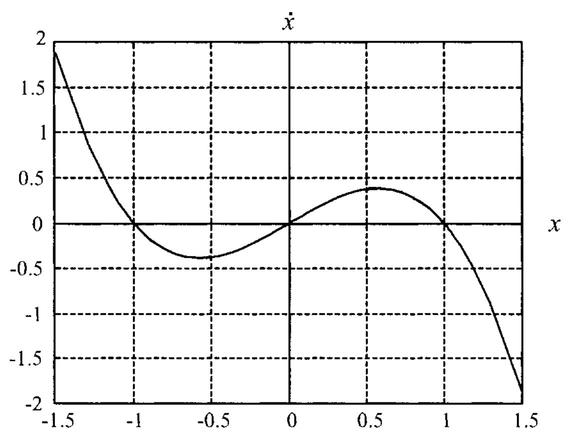


Figure 1.11: State trajectory of the system $\dot{x} = x - x^3$.

and further:

$$\text{for } x_{e1}: \left[\frac{\partial \dot{x}}{\partial x}\right]_{x_{e1}=0} = +1; \text{ for } x_{e2}: \left[\frac{\partial \dot{x}}{\partial x}\right]_{x_{e2}=1} = -2; \text{ for } x_{e3}: \left[\frac{\partial \dot{x}}{\partial x}\right]_{x_{e3}=-1} = -2$$

Therefore x_{e1} is an unstable equilibrium state or repeller, while x_{e2} and x_{e3} are stable equilibrium states or attractors. Every attractor has his own region of attraction. For the equilibrium state $x_{e2} = +1$, the region of attraction is the right half-plane where $x > 0$, while for $x_{e3} = -1$ it is the left half-plane where $x < 0$. It must be remarked that the tangents at equilibrium states are equivalent to the poles of the linearized mathematical model of the system.

1.5 Basic Properties of Nonlinear Functions

Nonlinearities which are present in automatic control system belong to two groups:

1. Inevitably present – inherent nonlinearities of the controlled process,

2. Intended – nonlinearities which are built in the system with the purpose to realize the desired control algorithm or to improve dynamic performance of the system.

In every nonlinear control system one or more nonlinear parts can be separated—these are the nonlinear elements. In analogy to the theory of linear systems, the theory of nonlinear systems presents the system elements with unidirectional action (Fig. 1.12) (Netushil, 1983).

Nonlinear elements with the output as a function of one variable (Fig. 1.12a) comprise typical nonlinear elements described by a nonlinear function which depends on the input quantity $x(t)$ and on its derivative $\dot{x} = \frac{dx}{dt}$:

$$y_N = F(x, \dot{x}) \tag{1.36}$$

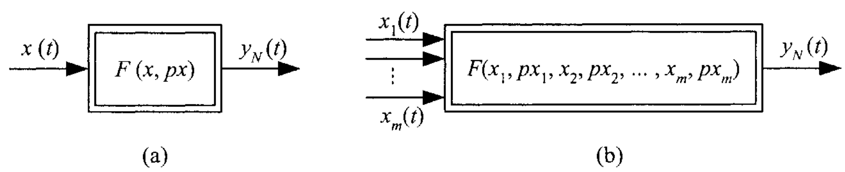


Figure 1.12: Nonlinear element presented in block diagrams of nonlinear systems.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Besides nonlinearities with one input, there are nonlinearities with more input variables (Fig. 1.12b), described by the equation:

$$y_N = F(x_1, \dot{x}_1, x_2, \dot{x}_2, \ldots, x_m, \dot{x}_m) \tag{1.37}$$

When the nonlinear characteristic depends on time, the nonlinear element is called time-varying, and the equations (1.36) and (1.37) take the form:

$$y_N = F(t, x, \dot{x}) \tag{1.38}$$

$$y_N = F(t, x_1, \dot{x}_1, x_2, \dot{x}_2, \ldots, x_m, \dot{x}_m) \tag{1.39}$$

Inherent nonlinearities include typical nonlinear elements: saturation, dead zone, hysteresis, Coulomb friction and backlash. Intended nonlinearities are nonlinear elements of various characteristics of relay type, A/D and D/A converters, modulators, as well as dynamic components designed to realize nonlinear control algorithms.

Nonlinearities can be classified according to their properties: smoothness, symmetry, uniqueness, nonuniqueness, continuity, discontinuity, etc. It is desirable to have a classification of nonlinearities, since the systems with similar properties can be treated as a group or class.

DEFINITION 1.10 (UNIQUENESS)
Nonlinear characteristic $y_N(x)$ is unique or single-valued if to any value of the input signal x there corresponds only one specific value of the output quantity y_N.

DEFINITION 1.11 (NON-UNIQUENESS)
Nonlinear characteristic $y_N(x)$ is non-unique or multi-valued when some value of the input signal x provokes several values of the output signal y_N, depending on the previous state of the system. The number of possible quantities y_N can be from 2 to ∞. Multi-valued nonlinearities include all nonlinearities of the type hysteresis.

DEFINITION 1.12 (SMOOTHNESS)
Nonlinear characteristic $y_N(x)$ is smooth if at every point there exists a derivative $\frac{dy_N(x)}{dx}$, Fig. 1.14 (Netushil, 1983).

DEFINITION 1.13 (SYMMETRY)
From the viewpoint of symmetry, two groups are distinguished:

1. *If the function $y_N(x)$ satisfies the condition:*

$$y_N(x) = y_N(-x) \tag{1.40}$$

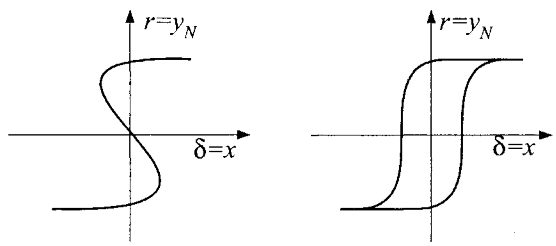


Figure 1.13: Static multi-valued nonlinear characteristics of course unstable ships.

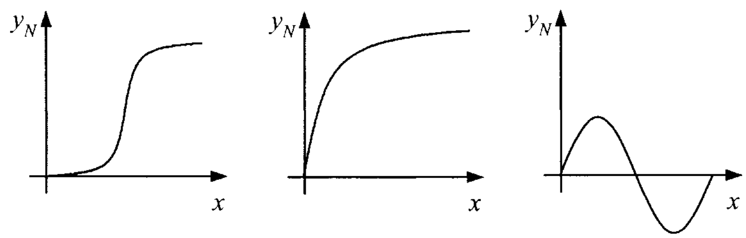


Figure 1.14: Examples of smooth and single-valued static characteristics.

the nonlinearity is symmetrical to the y-axis, i.e. the characteristic is even symmetrical. In the cases when the function (1.40) is unique, the nonlinearity can be described by:

$$y_N(x) = \sum_{i=0}^{\infty} C_{2i}x^{2i} \tag{1.41}$$

where C_{2i} are constant coefficients.

2. *If the function $y_N(x)$ satisfies the condition:*

$$y_N(x) = -y_N(-x) \tag{1.42}$$

nonlinearity is symmetrical to the coordinate origin, i.e. the characteristic is odd symmetrical. When the function (1.42) is unique, the nonlinearity can be written as:

$$y_N(x) = \sum_{i=1}^{\infty} C_{2i+1}x^{2i-1} \tag{1.43}$$

The characteristics which do not satisfy either of the above conditions are asymmetrical ones.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

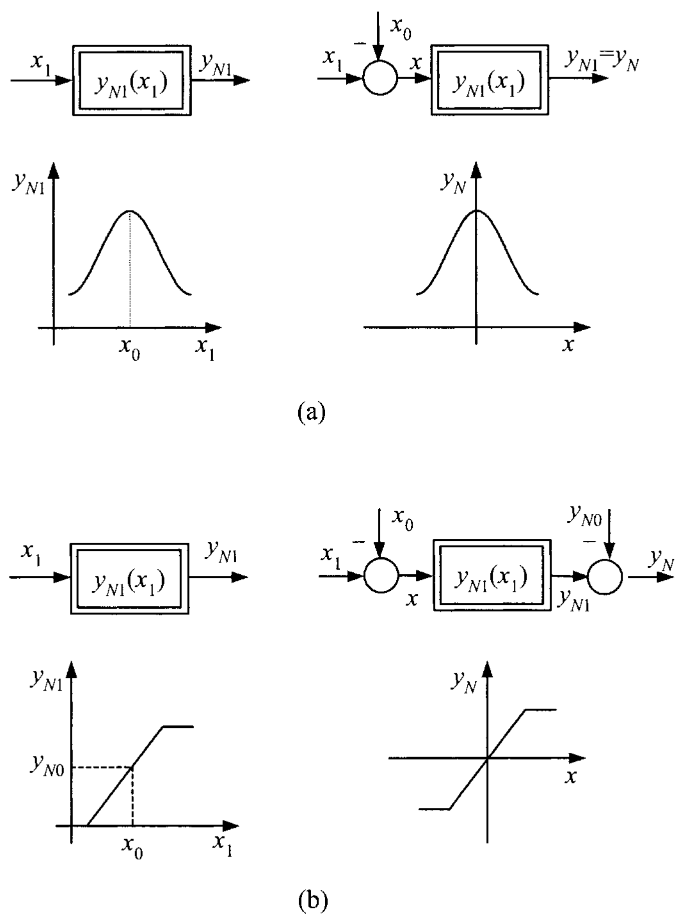


Figure 1.15: Asymmetrical characteristic becomes even (a) and odd (b) symmetrical.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

In the majority of cases, by shift of the coordinate origin, the asymmetrical characteristics can be reduced to symmetrical characteristics. The shift of the coordinate origin means the introduction of additional signals at all inputs and the output of the nonlinear element, Fig. 1.15a and 1.15b.

In Fig. 1.15a, assymetrical characteristic $y_{N_1}(x_1)$ with the coordinate transformation $x_1 = x_0 + x$ becomes even symmetrical, i.e. symmetrical with respect to y-axis. Similarly, asymmetrical characteristic from Fig. 1.15b by additional signals $x_1 = x_0 + x$ and $y_{N_1} = y_{N_0} + y_N$ becomes odd symmetrical, i.e. symmetrical with respect to the coordinate origin.

1.6 Typical Nonlinear Elements

Typical nonlinear elements are most often seen in automatic control systems where static characteristics can be normalized and approximated by straight line segments. If classified, such elements can be symmetric, asymmetric, unique, nonunique and discontinuous functions.

1.6.1 Nonlinear Elements with Single-Valued Continuous Characteristics

Dead zone. Presented with static characteristic as in Fig. 1.16, a dead zone is present in all types of amplifiers when the input signals are small. The mechanical model of the dead zone (Fig. 1.16e) is given by a fork joint of the drive and driven shaft. Vertical (neutral) position of the driven shaft is maintained by a spring. When the drive shaft (input quantity) is rotated by an angle x, the driven shaft remains at a standstill until the deflection x is equal to x_a. With $x > x_a$, the driven shaft (output quantity) rotates simultaneously with the drive shaft. When the drive shaft rotates in the opposite direction, because of the spring action, the driven shaft will follow the drive shaft until the moment when neutral position is reached. During the time the output shaft doesn't rotate, it goes through the dead zone of $2x_a$. Static characteristic of the dead zone element (Fig. 1.16a) is given by the equations:

$$y_N = \begin{cases} 0 & \text{for } |x| \leq x_a \\ K(x - x_a) & \text{for } x > x_a \\ K(x + x_a) & \text{for } x < -x_a \end{cases} \tag{1.44}$$

By introducing the variables $\mu = \frac{x}{x_a}$ and $\eta = \frac{y_N}{Kx_a}$, a normalized characteristic of the nonlinearity is obtained (Fig. 1.16d):

$$\eta = \begin{cases} 0 & \text{for } |\mu| \leq 1 \\ \mu - 1 & \text{for } \mu > 1 \\ \mu + 1 & \text{for } \mu < -1 \end{cases} \tag{1.45}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

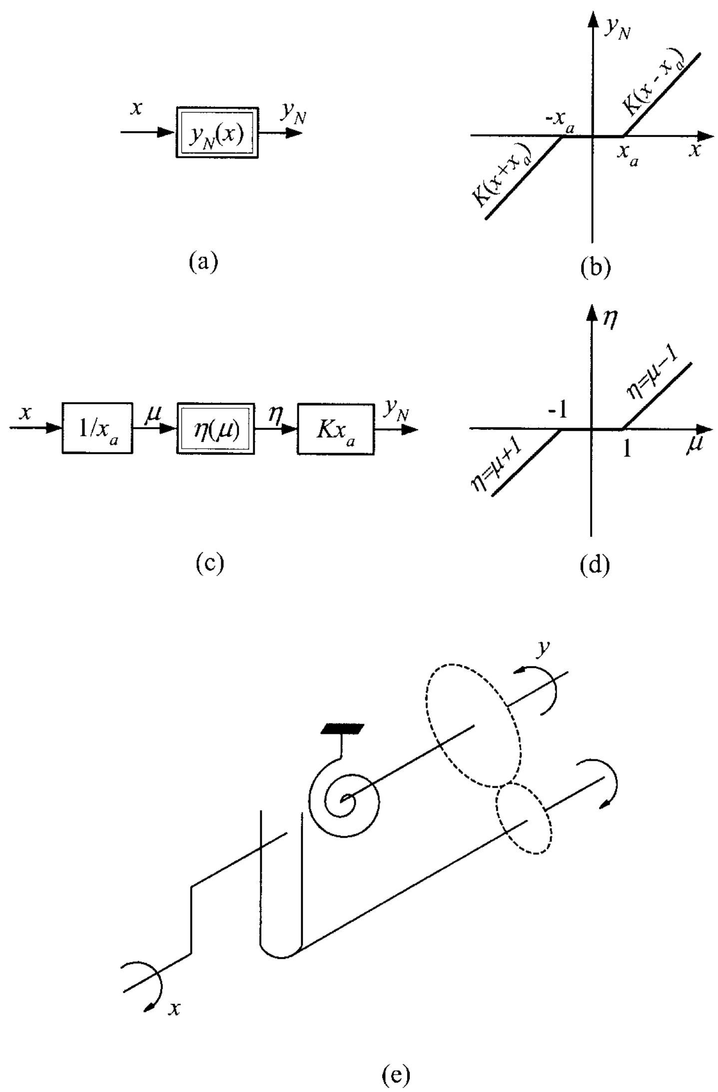
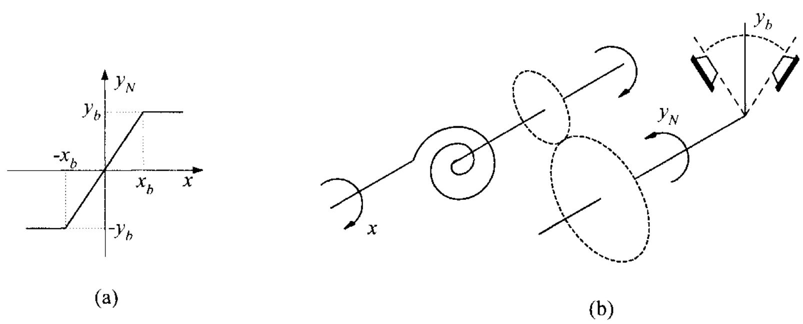


Figure 1.16: Nonlinear characteristic of the type dead zone.



(a) (b)

Figure 1.17: Nonlinear characteristic of the type saturation.

Saturation (limiter). Figure1.17a shows the static characteristic of the nonlinearity of the type saturation (limiter). Similar characteristics are in practically all types of amplifiers (electronic, electromechanic, pneumatic, hydraulic, combined ones) where the limited output power cannot follow large input signals. As an example of a simple mechanical model of the nonlinear element of the type saturation is a system which couples two shafts with a torsional flexible spring and with limited rotation of the driven shaft (Fig. 1.17b). The linear gain of the driven shaft is $2y_b$. The static characteristic of this element is described by the equations:

$$y_N = \begin{cases} Kx & \text{for } |x| \leq x_b \\ y_b \cdot \operatorname{sign} x & \text{for } |x| > x_b \end{cases} \tag{1.46}$$

Introducing in (1.46) the variables $\mu = \frac{x}{x_b}$; $\eta = \frac{y_N}{Kx_b}$, the normalized equations are obtained:

$$\eta = \begin{cases} \mu & \text{for } |\mu| \leq 1 \\ \operatorname{sign} \mu & \text{for } |\mu| > 1 \end{cases} \tag{1.47}$$

Saturation with dead zone. The static characteristic of this type of nonlinearity (Fig. 1.18) describes the properties of all real power amplifiers, i.e. with small input signals such element behaves as dead zone, while large input signals cause the output signal to be limited (saturated).

Static characteristic from Fig. 1.18 is described by the equations:

$$y_N = \begin{cases} 0 & \text{for } |x| \leq x_a \\ K(x - x_a) & \text{for } x_b > x > x_a \\ K(x + x_a) & \text{for } -x_b < x < -x_a \\ y_b \cdot \operatorname{sign} x & \text{for } |x| > x_b \end{cases} \tag{1.48}$$

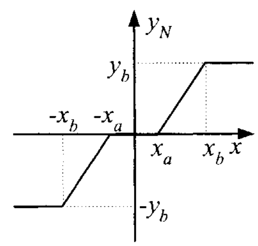


Figure 1.18: Static characteristic of saturation with dead zone.

Normalized form of the equation (1.48) is obtained by introducing $\mu = \frac{x}{x_a}$, $\eta = \frac{y_N}{Kx_a}$ and $m = \frac{x_b}{x_a}$:

$$\eta = \begin{cases} 0 & \text{for } |\mu| \leq 1 \\ \mu - 1 & \text{for } m > \mu > 1 \\ \mu + 1 & \text{for } -m < x < -1 \\ (m-1) \cdot \operatorname{sign} \mu & \text{for } |\mu| > m \end{cases} \tag{1.49}$$

1.6.2 Nonlinear Elements with Single-Valued Discontinuous Characteristics

Typical elements with single-valued discontinuous characteristics are practically covered by the two-position and three-position relay without hysteresis.

Two-position relay without hysteresis. Single-valued static characteristic of two-position polarized relay is shown in Fig. 1.19.

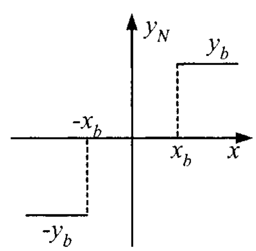


Figure 1.19: Static characteristic of a two-position polarized relay without hysteresis.

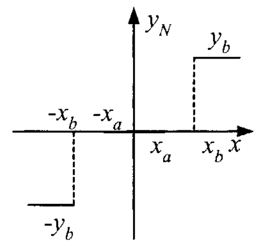


Figure 1.20: Static characteristic of a three-position relay without hysteresis.

From this static characteristic it is seen that the relay contacts are open in the interval $|x| < x_b$, i.e. the output signal y_N is undefined. In the region $|x| > x_b$, the output signal has the value $+y_b$ or $-y_b$, depending on the sign of the input signal x. The corresponding equation for the two-position relay without hysteresis is:

$$y_N = \begin{cases} y_b \cdot \operatorname{sign} x & \text{for } |x| \geq x_b \\ \text{undefined} & \text{for } |x| < x_b \end{cases} \tag{1.50}$$

Three-position relay without hysteresis. Single-valued static characteristic of a three-position relay without hysteresis is shown in Fig. 1.20. The equations which describe the static characteristic from Fig. 1.20 are:

$$y_N = \begin{cases} y_b \cdot \operatorname{sign} x & \text{for } |x| \geq x_b \\ 0 & \text{for } |x| \leq x_a \\ \text{undefined} & \text{for } x_a < |x| < x_b \end{cases} \tag{1.51}$$

Ideal two-position relay. For the analysis and synthesis of various relay control systems, use is made of the idealized static characteristic of a two-position relay. Here at our disposal are equations (1.46), (1.48), (1.50) and (1.51). With the boundary conditions $x_b \to 0$ and $K \to \infty$ for $Kx_b = y_b$ substituted into equation (1.46), we obtain:

$$y_N = \begin{cases} y_b \cdot \operatorname{sign} x & \text{for } |x| > 0 \\ -y_b < y < y_b & \text{for } x = 0 \end{cases} \tag{1.52}$$

A graphical display of an ideal relay characteristic described by (1.52) is given in Fig. 1.21a.

By inserting the boundary conditions in (1.50), the static characteristic in Fig. 1.21b is obtained, together with the equation:

$$y_N = \begin{cases} y_b \cdot \operatorname{sign} x & \text{for } |x| > 0 \\ \text{undefined} & \text{for } x = 0 \end{cases} \tag{1.53}$$

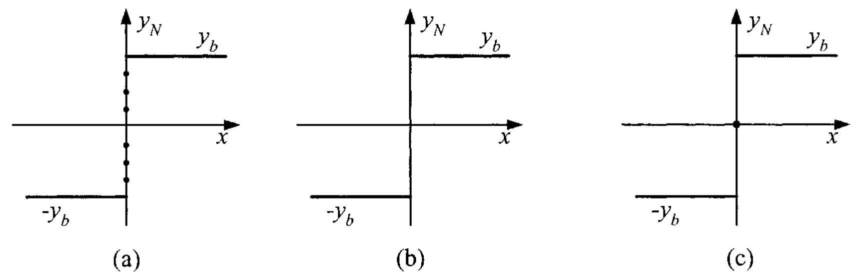


Figure 1.21: Ideal static characteristics of a two-position relay.

By analogy, if the boundary conditions are inserted in (1.48) and (1.51), the static characteristic in Fig. 1.21c ensues and the equation is as follows:

$$y_N = \begin{cases} y_b \cdot \mathrm{sign} x & \text{for } |x| > 0 \\ 0 & \text{for } x = 0 \end{cases} \tag{1.54}$$

For the analysis of relay systems with feedback loops, it is indispensable to examine various conditions of the system behavior for $x = 0$, for every characteristic from Fig. 1.21.

The characteristics from Fig. 1.21 described by the equation (1.52) are called nonlinearity of the signum type. The signum function can be defined analytically as:

$$y_N = y_b \cdot sign\, x = y_b \lim_{n \to \infty} x^{-(2n+1)} \tag{1.55}$$

respectively:

$$x = \mathrm{asign} \frac{y_N}{y_b} = \lim_{n \to \infty} \left(\frac{y_N}{y_b} \right)^{2n+1} \tag{1.56}$$

1.6.3 Nonlinear Elements with Double-Valued Characteristics

Two-position relay with hysteresis. The discussed single-valued relay static characteristic represents an idealized characteristic of real systems. In reality, the values of the input signal at the discontinuity in the value of the output signal y_N are different, depending on whether the relay switches in one or the other direction. For example, with a two-position polarized relay with symmetric control, closing of the contacts in one direction comes with an input voltage x, and the closing in the reverse direction requires the input voltage $-x$. The characteristic of a two-position relay with double-valued characteristic is shown in Fig. 1.22,

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

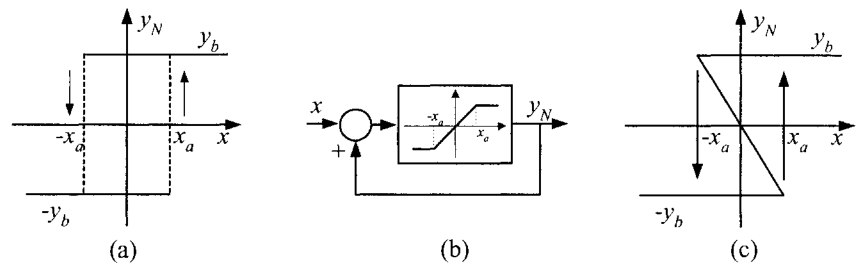


Figure 1.22: Static characteristics of a two-position relay with hysteresis.

and the equivalent mathematical expression is:

$$y_N = \begin{cases} +y_b & \text{for } -x_a < x < \infty \\ -y_b & \text{for } -\infty < x < x_a \end{cases} \tag{1.57}$$

On the part of the characteristic $-x_a < x < x_a$, the output quantity y_N has two values $+y_b$ or $-y_b$, depending on the previous value of the input quantity x. The conditions for a jump of the output quantity from $-y_b$ to $+y_b$ and vice versa are determined by $x = x_a$, $y_N = -y_b$, $dx/dt > 0$ and $x = -x_a$, $y_N = y_b$, $dx/dt < 0$, respectively.

The properties of such an element as two-position relay with hysteresis exhibit the amplifiers with positive feedback (Fig. 1.22b).

For the part of the static characteristic in Fig. 1.22c given by the input signal $-x_a < x < x_a$, $-y_b < y_N < y_b$ has a negative slope and is usually unstable. Namely, in the above-mentioned interval to any value of x correspond three values of y_N, where the values $y_N = \pm y_b$ are stable, while the values $-y_b < y_N < y_b$ are unstable. In such a way, this characteristic is treated as a double-valued nonlinear element.

Three-position relay with hysteresis. The static characteristic of a three-position relay with hysteresis is shown in Fig. 1.23, and the corresponding equation is:

$$y_N = \begin{cases} y_b \cdot \operatorname{sign} x & \text{for } |x| > x_b \\ 0 & \text{for } |x| < x_a \end{cases} \tag{1.58}$$

In the part of the characteristic from Fig. 1.23a, when $|x_b| < |x| < |x_a|$, the output signal y_N has three possible values, 0 and $\pm y_b$. An analogous characteristic is obtained with positive feedback with amplifiers which have dead zone and saturation (Fig. 1.23b). The third characteristic in Fig. 1.23c is a continuous function with two (usually unstable) regions of negative slope, inside which a jump of the output quantity occurs. This is equivalent to the characteristic in Fig. 1.23a.

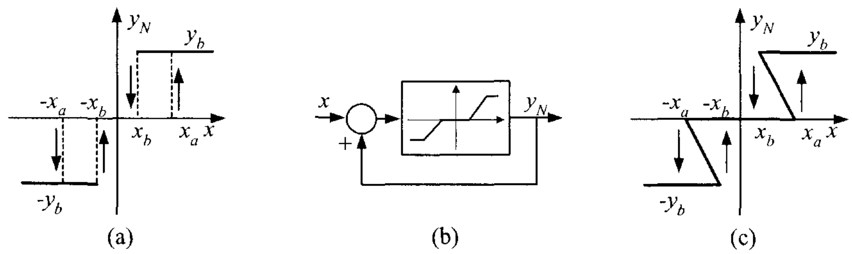


Figure 1.23: Static characteristics of a three-position relay with hysteresis.

1.6.4 Nonlinear Elements with Multi-Valued Characteristics

Backlash (play-type element). This nonlinear element is very often inherent in mechanical systems which transmit motion. The static characteristic and mechanical models of this type of nonlinearity are presented in Fig. 1.24.

The analysis of the equation $y_N = y_N(x)$ on the model of Fig. 1.24a leads to the conclusion that a change in the value of the input quantity x (deflection of the driving shaft) up to $x \leq x_a$ doesn't result in transmission of the motion to the driven shaft (there is no change in the output quantity y_N). For $x > x_a$ the transmission of the motion takes place along the line AB. When the motion of the driving shaft changes direction at any place on the line AB, at the beginning

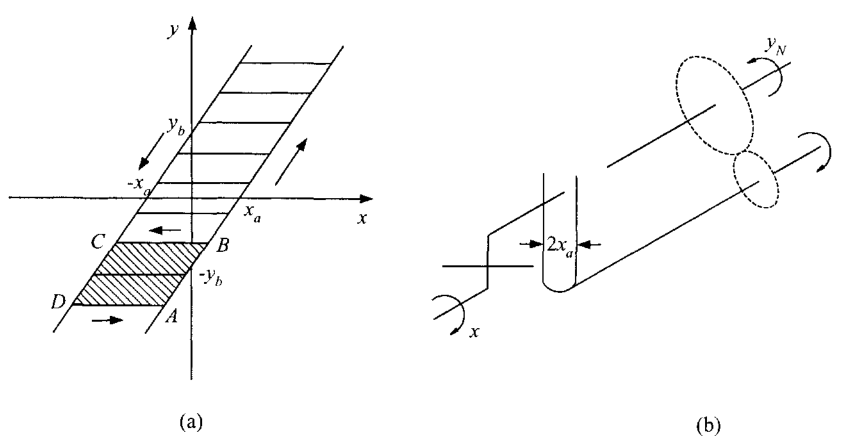


Figure 1.24: Nonlinear element of the type backlash.

the driven shaft will not move, up to the moment that the driving shaft fulfils the rotation on the horizontal part of the characteristic from the line AB to the line CD. Thereafter, the transmission of the angular motion takes place along the line CD. Such a nonlinearity is called backlash or play-type element (Fig. 1.24a) and in the analytical form is given by:

$$y_N = \begin{cases} K(x - x_a) & \text{for } \dot{y}_N > 0 \\ K(x + x_a) & \text{for } \dot{y}_N < 0 \\ \text{const.} & \text{for } \dot{y}_N = 0, \left| x - \frac{y_N}{K} \right| < x_a \end{cases} \tag{1.59}$$

Coulomb friction. The Coulomb friction (also called *dry friction*) is presented in Fig. 1.25. The static characteristic in Fig. 1.25a shows the force or torque as a function of the velocity of the input quantity x.

The performance of such type of nonlinear element can be reasonably approximated by ideal relay characteristic (1.25a) for the situation when the momentary passage of the input quantity $\dot{x}$ through zero results in a momentary change of the sign for force or torque of Coulomb friction.

In the case when the value of the input quantity $\dot{x} = 0$ is not momentary, the motion stops in the Coulomb friction element—the approximation of the static

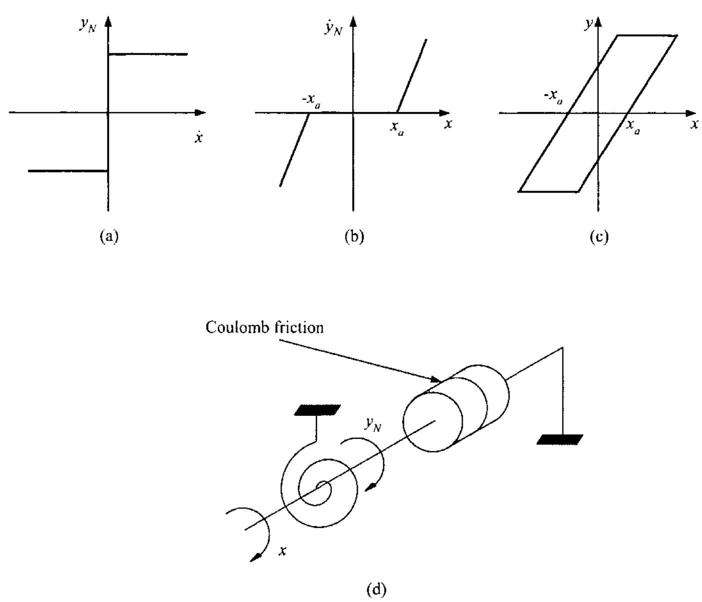


Figure 1.25: Static characteristics of Coulomb friction.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

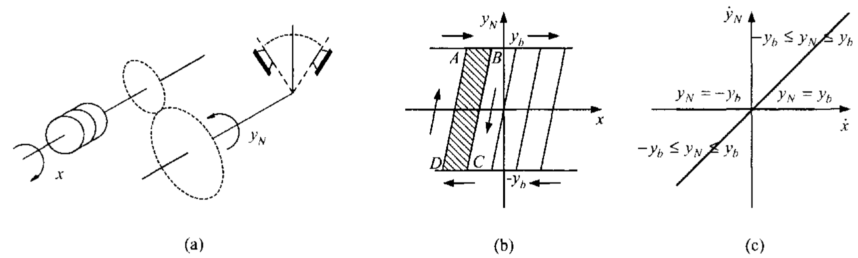


Figure 1.26: Static characteristics of a nonlinear stop-type element.

characteristic of this element by the relay characteristic leads to wrong conclusions in determining the parameters of oscillatory processes. The transmission of motion stops at the moment when the change of the input quantity becomes zero, and forces or torques which act at the input are smaller than forces or torques of the Coulomb friction.

The consequences of stopping, which comes at the moment when $\dot{x} = 0$, can be easily determined in two specific cases.

In the first case the static characteristic of the element of the type Coulomb friction can be approximated by the static characteristic of the element of the type dead zone (Fig. 1.25b). This is justified when the counterforce and the mass of the moving parts of the element are negligible (small inertial forces). The output quantity of the element (velocity) is equal to zero in the region $-x_a \leq x \leq x_a$. When the input quantity $|x| > |x_a|$, the output quantity $\dot{y}_N$ becomes proportional to the difference $|x| - |x_a|$.

In the second case when inertial forces are negligibly small and a balance is established within the element between the equilibrium force/torque and the input force/torque (mechanical model in Fig. 1.25d), the static characteristic of the element Coulomb friction can be approximated by the static characteristic of backlash (Fig. 1.25c). On the mechanical model of the element, the torque of the drive shaft ($x = M$) is at equilibrium with the countertorque of the spring βy_N (β is coefficient of proportionality) and the Coulomb friction torque $\pm x_a$. This is mathematically formulated:

$$x = M = \beta y_N \pm x_a$$

Substitution of $k = 1/\beta$ yields:

$$y_N = k(x \mp x_a) \tag{1.60}$$

Stop-type element. In many mechanical systems, the motion of the output quantity is limited in two directions, while at the sime time the motion of the input

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

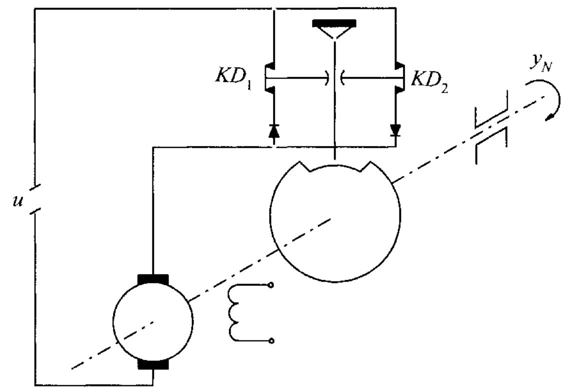


Figure 1.27: Schematic of ship's rudder driven by an electromotor.

quantity is unbounded and can be of constant direction. This is typical for the nonlinearity of the stop-type. Fig. 1.26 gives the mechanical model and the static characteristics.

On the mechanical model (Fig. 1.26a) the motion of the drive shaft is transfered to the driven shaft by means of the friction clutch which disengages at the moment when resisting torque becomes larger than the boundary value. The driven shaft has stops on both sides. When the driven shaft turns to the stop, a large resistance torque arises and the clutch is disengaged from the drive shaft. The latter rotates further while the driven shaft is at a standstill. For a change of direction of the motion of the input quantity x, the resistance torque rises, the clutch is closed and the driven shaft rotates simultaneously with the drive shaft.

Analytical description of the nonlinearity of the stop-type is (Fig. 1.26b and 1.26c):

$$\dot{y}_N = \begin{cases} K\dot{x} & \text{for} \begin{cases} \dot{x} > 0 & -y_b \le y_N < y_b \\ \dot{x} < 0 & -y_b < y_N \le y_b \end{cases} \\ 0 & \text{for} \begin{cases} \dot{x} > 0 & y_N = y_b \\ \dot{x} < 0 & y_N = -y_b \end{cases} \end{cases} \tag{1.61}$$

Nonlinear elements of the stop-type describe processes in pneumatic and hydraulic amplifiers, as well as in electromechanical processes with electromotors and actuating mechanisms with contacts in the armature circuit. Fig. 1.27 illustrates this with the schematic of a ship's rudder driven by the electromotor with contacts KD_1 and KD_2 in the armature circuit.

When one of the contacts is closed, the velocity of the motor shaft is proportional to the armature voltage: $\dot{y}_N = Ku$. When the motor shaft turns to one of

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

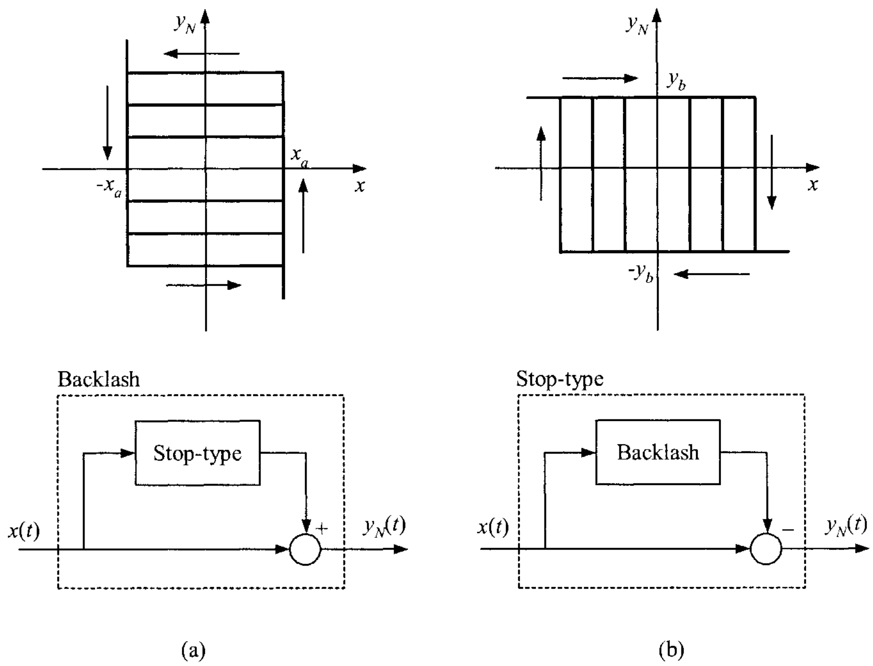


Figure 1.28: Nonlinear characteristics of backlash (a) and stop-type element (b) at high gains.

the stop-elements on the excentricity E, the contacts reopen, the motor is disconnected from the power supply, and the velocity of the motor shaft becomes zero. When the inertia of the motor is neglected, the relation between $u = \dot{x}$ and $\dot{y}_N$ is equal to the characteristic of the nonlinear element of stop-type.

With a high gain $K \longrightarrow \infty$, the characteristics of the nonlinear elements of the backlash and stop-type become rectangular (Fig. 1.28).

The comparison of the characteristic for backlash (Fig. 1.28a) and stop-type element (Fig. 1.28b) shows that any cyclic change of the input quantity x results in the cyclic change of the output quantity y_N with reversed direction of movement arround the characteristic. For the characteristic of the type backlash $\oint y_N dx < 0$, while for the characteristic of the stop-type $\oint y_N dx > 0$.

Comparing these two nonlinearities it follows that the characteristic of the stop-type can be realized from characteristic of the backlash according to the structure from Fig. 1.28c and vice versa, the characteristic for backlash is obtained from the characteristic for stop-type (Fig. 1.28d).

Rectangular characteristics of backlash and stop-type elements (Figures 1.28a

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

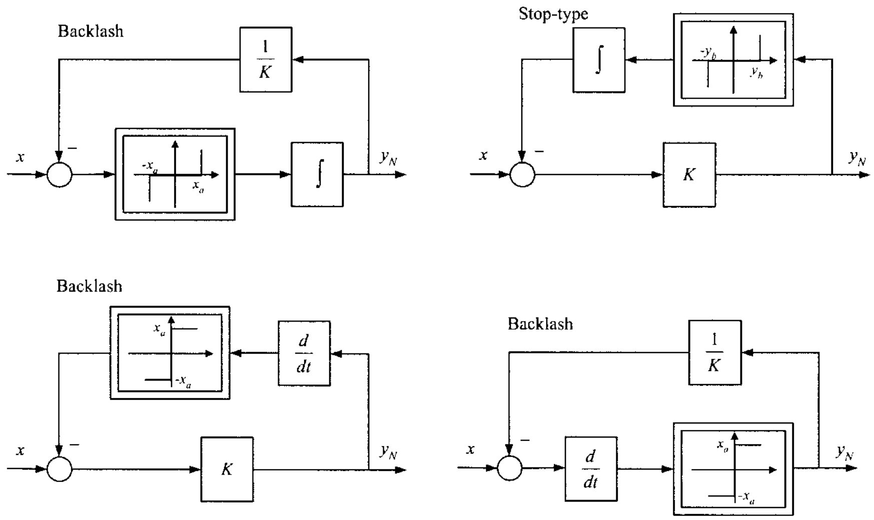


Figure 1.29: Possible realizations of nonlinear characteristics of backlash and stop-type element.

and 1.28b) can be realized with a single-valued dead zone element, or with the nonlinear signum element (Fig. 1.29).

Magnetic hysteresis. Static characteristics of the nonlinearity of the type magnetic hysteresis are presented in Fig. 1.30. Magnetic hysteresis is approximated by piecewise straight lines, and is described by the following equation:

$$\begin{aligned} y_N &= \mu_1(x \pm x_a) \\ y_N &= \mu_2 x + C \end{aligned} \tag{1.62}$$

where μ_1 and μ_2 are constants, $-y_b < C < y_b$. Inside the hysteresis loop, the quantity y_N can assume the values $\mu_2 x - y_b < y_N < \mu_2 x + y_b$ depending on initial conditions. From the static characteristic in Fig. 1.30a for $\mu_1 \to \infty$ and $\mu_2 \to 0$ a rectangular hysteresis loop is obtained (Fig. 1.30b). The structural block diagram of the magnetic hysteresis nonlinear element is shown in Fig. 1.30c.

1.7 Atypical (Non-Standard) Nonlinear Elements

Atypical nonlinear elements are built in control systems in order to realize necessary nonlinear algorithms. One of most used elements is for the operation of

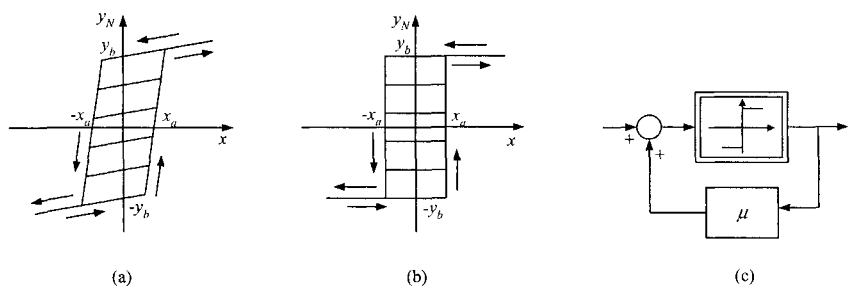


Figure 1.30: Static characteristic of a magnetic hysteresis.

multiplication, which has the name *modulator*. Depending on the input quantites, the output quantity can be linear or nonlinear (Fig. 1.31) (Netushil, 1983).

If at the input there are two mutually independent quantities x_1 and x_2 (Fig. 1.31a), the output quantity $y_N = x_1 x_2$ characterizes any change of system parameters but does not disturb the system linearity. If the input quantities to the modulator are mutually dependent (Fig. 1.31b), the system becomes nonlinear, i.e. the output quantity of the system with linear component $G(s)$ is a nonlinear function $y_N = x^2 G(s)$.

For the realization of extremal control systems, an even-parabolic function (1.32a) is realized by the multiplying element as shown in the block diagram (Fig. 1.32b).

Combining the multiplying element with typical nonlinear elements, a nonlinear element is obtained with an odd-parabolic function (Fig. 1.33). The characteristic of the nonlinearity $y_N = x^2 \text{sign} x$ (Fig. 1.33a) can be realized either by the structure in Fig. 1.33b or the structure in Fig. 1.33c. Various control laws in extremal systems and the systems with variable structure are achieved with the

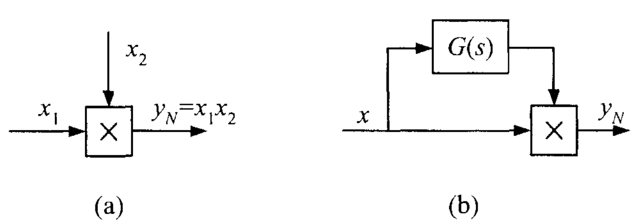


Figure 1.31: Element for multiplication.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

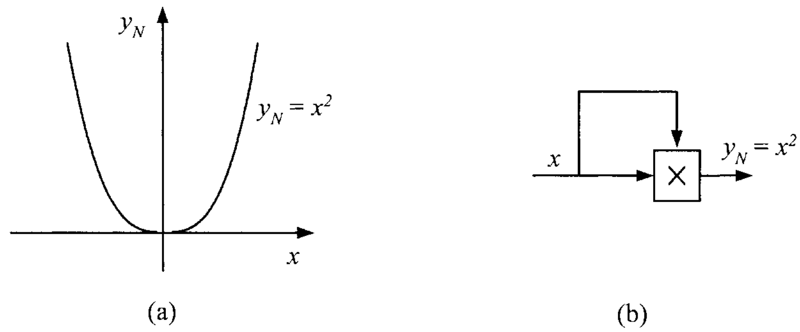


Figure 1.32: Realization of an even-parabolic function.

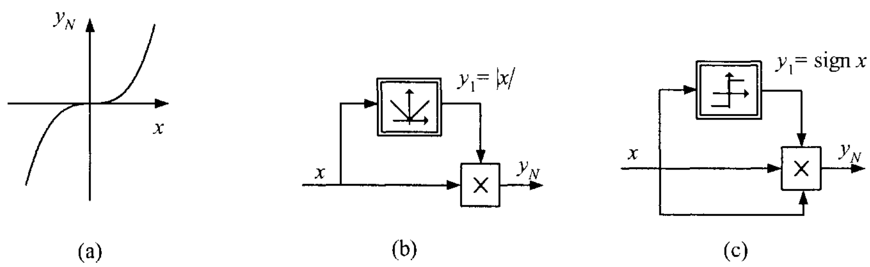


Figure 1.33: Realization of an odd-parabolic function.

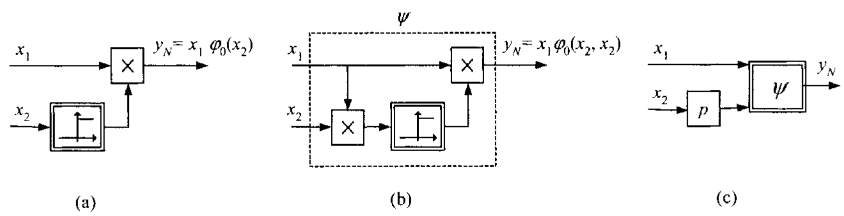


Figure 1.34: Various realizations of special nonlinear elements.

combination of multiplying and relay elements (Fig. 1.34). Nonlinear element in Fig. 1.34b is called ψ-element, $\psi = \frac{y_N}{x_1} = \varphi_0(x_1 x_2)$ — it is applied in the systems with variable structure. With such an element, an element with negative hysteresis (Fig. 1.33c) similar to the stop-type element can be built.

1.8 Basic Nonlinearity Classes

The concept "basic nonlinearity classes" is understood to include single-valued nonlinear functions $y_N = F(x)$, time-varying nonlinear functions $y_N = F(t,x)$ and multi-valued nonlinear characteristics which can be combined from elements with single-valued nonlinear characteristic and from the linear part of the system: Figs.1.22, 1.29, 1.30, 1.33, 1.34, etc.

In the following text a broad class of nonlinear functions, i.e. nonlinear systems which can be represented by the mathematical model-structure as in Fig. 1.2, will be treated. Here the nonlinear part of the system can be generalized by a set of nonlinear functions with some common properties. This basic classes include the nonlinear functions belonging to the sectors $[k_1, k_2]$, $[0, k]$ and $[0, \infty]$ (Fig. 1.35) (Voronov, 1979; Nelepin, 1971).

DEFINITION 1.14 (NONLINEAR FUNCTION OF THE SECTOR $[k_1, k_2]$)
The nonlinear continuous function $y_N = F(x)$ is said to belong to the sector or class $[k_1, k_2]$ if there exist nonnegative constants k_1 and k_2 so that the following equations are satisfied:

$$\begin{aligned} & k_1 \le \frac{y_N}{x} = \frac{F(x)}{x} \le k_2; \quad x \neq 0 \\ & y_N = 0; \quad x = 0 \end{aligned} \tag{1.63}$$

The condition (1.63) is equivalent to the condition:

$$(k_2 x - y_N)(y_N - k_1 x) \ge 0 \tag{1.64}$$

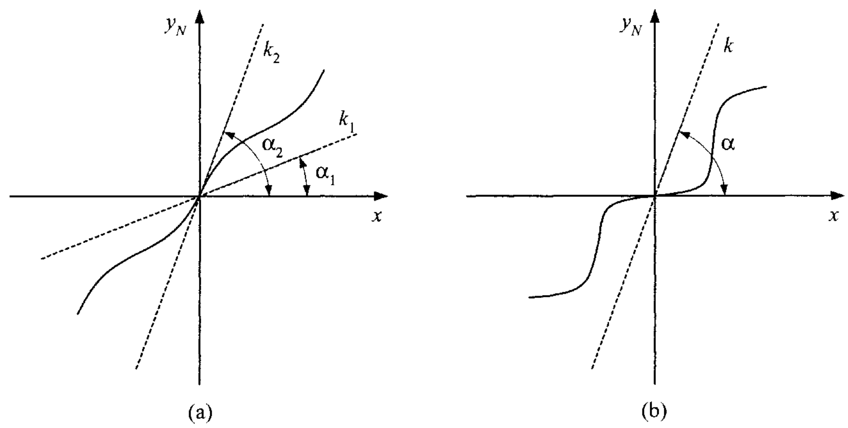


Figure 1.35: Nonlinear functions of class $[k_1, k_2]$ (a) and of class $[0, k]$ (b).

The conditions (1.63) and (1.64) imply two properties. The nonlinear characteristic touches the origin $F(0) = 0$, and $xF(x) \geq 0$, i.e. the curve $F(x)$ lies in the first and the third quadrants (Fig. 1.35a) between two straight lines with slopes k_1 and k_2— both of them crossing the origin.

DEFINITION 1.15 (NONLINEAR FUNCTION OF THE SECTOR $[0, k]$)
The nonlinear continuous function $y_N = F(x)$ is said to belong to the sector or class $[0, k]$ if there exists a nonnegative constant $0 < k \leq \infty$ so that the following equations are satisfied:

$$\begin{aligned} &0 \leq \frac{y_N}{x} = \frac{F(x)}{x} \leq k; \quad x \neq 0 \\ &F(x) = 0; \quad x = 0 \end{aligned} \tag{1.65}$$

The nonlinear characteristic of the sector $[0, k]$ is shown in Fig. 1.35b.

The condition (1.65) is equivalent to the condition:

$$(kx - y_N)y_N \geq 0 \tag{1.66}$$

For $k \rightarrow \infty$, the conditions (1.65) and (1.66) obtain the form:

$$\begin{aligned} &0 \leq \frac{F(x)}{x}; \quad x \neq 0 \\ &F(x) = 0; \quad x = 0 \end{aligned} \tag{1.67}$$

or:

$$x \cdot F(x) \geq 0 \tag{1.68}$$

Characteristics $F(x)$ of the sector $[k_1, k_2]$ can be deduced to the sector $[0, k]$ by substituting in (1.63):

$$F_1(x) = F(x) - kx \tag{1.69}$$

From (1.63) and (1.69) follows:

$$\begin{aligned} &0 \leq \frac{F_1(x)}{x}; \quad x \neq 0 \\ &F(x) = 0; \quad x = 0; \quad k = k_2 - k_1 \end{aligned} \tag{1.70}$$

It is obvious from the above that the nonlinear functions which were analyzed in Sections 1.6 and 1.7, are classified in the nonlinearities of sector $[0, k]$. Thus, a generalized approach to the analysis and the synthesis of the dynamics of the nonlinear systems of this class is made possible.

1.9 Conclusion

Properties (dynamical phenomena) of nonlinear systems briefly presented in this chapter characterize nonlinear systems. Of course not all properties will be present in a particular system. Of all presented properties stability is the most important one, because every control system should at least possess the stability of all equilibrium states in which it can operate. Some properties at first glance look strange and do not seem to possess any useful application. For instance, looking at the chaotic property we will agree that this property is not wanted in any tracking control system. However, pathological state due to fault(s) or some other reason can push the system into chaotic behavior. A control engineer should therefore analyze his system in order to see if this is at all possible. Technical systems are built in such a way that chaos should not be encountered except in some specific situations, like in mixers where chaotic behavior is welcome for the blending purpose.

Typical structure of nonlinear single-input single-output control systems with unity feedback as well as typical nonlinear elements encountered quite often in nonlinear control systems are also presented. Simplicity of the structure enables the use of conventional analysis and design methods. If our nonlinear control system doesn't have the same structure, we have to reduce it, if possible, to recommended structure in order to apply analysis methods presented in this book.

Nonlinear dynamical phenomena impact system behavior. Our goal in control engineering is to obtain specified objectives by altering nonlinear phenomena or by using them to obtain the goal. It should be mentioned here that traditional linear control strategies for nonlinear systems are very often used with good results. However, the linear control theory can not be always used, so in those situations we are left with our knowledge of nonlinear control systems to obtain the specified goal. In the rest of the book we will concentrate on some phenomena which are very common in nonlinear systems such as self-oscillations or oscillatory behavior of nonlinear systems. Stability as the basic and most important property will be covered in the following chapter.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Chapter 2

Stability

Stability is the basic property of any useful system. It is necessary to be cautious here, because when we talk about stability we are talking about the property of an equilibrium state of a system. This distinction should be obvious after reading this chapter. The chapter is fairly extensive and should be digested in small amounts. Hopefully, whoever adopts the material offered in this chapter will improve their knowledge about stability in general.

Stability is the most important qualitative property of automatic control systems. Unstable systems have no practical significance. The concept of stability is so important since every control system must be primarily stable, and only then other properties can be studied. The theory of stability was the preoccupation of scientists from the beginning of the theory of differential equations. The key problem is to obtain information on the system's behavior (state trajectory) without solving the differential equation. The theory considers the system's behavior over a long period of time, i.e. as $t \to \infty$. One of the first scientists to investigate the stability of conservative mechanical systems in the "modern" sense was J.L. Lagrange — his observation was that the equilibrium state of an unforced system is stable if it has a minimum of potential energy. The first discussion about the instability of a control system according to Fuller (1976a) was given in Airy (1840), who was a mathematician and astronomer, first at Cambridge University (between 1826 and 1835) and later at Greenwich Observatory (1835-1881). He proposed that the telescope should rotate contrary to Earth's rotation in order to make possible longer observation times of stars. The prerequisite condition for that was to have a qualitative control system for the telescope's angular velocity. At that time the centrifugal governor was available for control purposes. Airy observed that such a controller can bring the system to an unstable state, with "wild" behavior. According to Fuller, Airy was the first to analyze the control system's dynamics by means of a differential equation. A basic step forward was made by Russian mathematician A.M. Lyapunov (1892)[1], who defined the general concepts of stability, for both linear and nonlinear systems. Beside Lyapunov, there were others who studied stability problems: J.C. Maxwell (1868), E.J. Routh (1877), I.A. Wischnegradskii (1876), A. Stodola (1893) and A. Hurwitz (1895). They studied the conditions which the coefficients of the linear differential equation must fulfil so that the system is stable. While Maxwell and Wischnegradskii considered that such conditions are possible only for systems of third order, Routh in his awarded paper (Adam Prize Essay[2]) generalized this to linear systems of any order. Later followed the works of A.M. Lyapunov (1892), A. Hurwitz (1895), H. Nyquist (1932), et al. A. Hurwitz, not knowing about Routh's work, came to the same conclusion. An interested reader can consult the books of Tsypkin (1951), Fuller (1975), Kharitonov (1979, 1994), Popov (1973), Isidori (1995), Slotine and Li (1991), Vidyasagar (1993), et al., which enable a deeper insight to this interesting research area.

Stability can be treated from the viewpoint of the systems input-output behavior, namely when it is expected that the response is "well-behaved" when the excitation is "well-behaved". The other viewpoint is to observe the asymptotic behavior of the state of the system in the vicinity of the equilibrium state or oscillatory states. Then stability can be referred to as the stability in the Lyapunov sense, as he studied the stability in the vicinity of equilibrium states. "Stability is in itself such a clear concept that it speaks for itself" wrote La Salle and Lefschetz (1961). Nevertheless, mathematical analysis of stability requires a quantitative characteristic. Although the concept of stability is associated with the system, this is not the complete contents of the idea. Stability is a property required for a system in order to function for a long period of time. In dynamics and control theory it is correct to speak of the system's state (equilibrium or motion) and not of the system itself. Only if the system possesses a single equilibrium state does the concept "stability of the system" have meaning. With several equilibrium states, it is appropriate to speak about "stability of equilibrium states". As linear systems have usually[3] only one equilibrium state, in the theory of linear systems it is common to speak about the stability of the system. Nonlinear systems have possibly more equilibrium states, as revealed by the next example of a nonlinear system described by the differential equation:

$$\ddot{x}(t) - 0.5x(t)\dot{x}(t) + x(t) = 0 \tag{2.1}$$

Phase trajectories for different initial conditions are shown in Fig. 2.1. The separatrix (phase trajectory which divides phase space into regions with different

[1]Shcherbakov published in 1992 an interesting article on occasion of the centennial of Lyapunov's dissertation (Shcherbakov, 1992).

[2]It is interesting to note that Maxwell was in the jury for the award. It was a biennal competition for the best scientific contribution to the theme "stability of motion"(1877).

[3]A linear system described with $\ddot{x} + \dot{x} = 0$ has an infinite number of equilibrium states (singular points), all on the real axis. Linear systems with more than one equilibrium state are the exception.

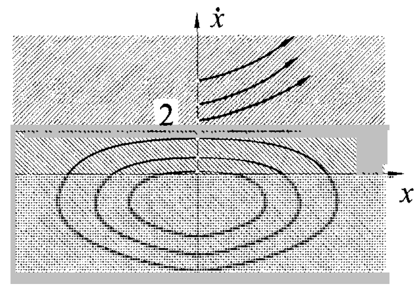


Figure 2.1: Phase trajectories for system (2.1).

behavior) is in this example a straight line parallel to the x-axis and with equation $\dot{x} = 2$. For all initial conditions bellow this line, the phase trajectory is a limit cycle — above the line the phase trajectory is of the type unstable focus[4]. Therefore, in this example, stable and unstable equilibrium states are present. Obviously the stability in this example depends on initial conditions.

In this chapter a short survey of the theory of stability will be given, and especially the part that treats the stability of equilibrium states of nonlinear systems. Input-output stability will not be considered, since such a stability concept requires the reader to be familiar with Lebesque theory. An interested reader should consult Sandberg (1964), Zames (1966a, 1966b), Desoer and Vidyasagar (1975), Kharitonov (1994), Naylor and Sell (1982), Vidyasagar (1993), et al.

2.1 Equilibrium States and Concepts of Stability

The concept of stability can be best understood after the possible equilibrium states are defined and explained. As an example, a simple system is taken with a free rolling ball on an uneven surface (Fig. 2.2). The ball can be at a standstill at the points with zero slope, such as the points A, E, F, G and the interval between points B and D, for example the position C. Every point where the ball can remain stationary is called an *equilibrium point* or *equilibrium state*. An infinitesimal change of position, because of a small (local) disturbance at the points A and F, will result in the ball not returning to the starting position. The equilibrium states A and F are therefore *unstable equilibrium states* (points) of the system. If a small disturbance moves the ball from the position E and G, it can eventually return to the starting equilibrium state—these are *stable equilibrium states* (points)

[4]More about singular points of the nonlinear system will be said in Chapter 5.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

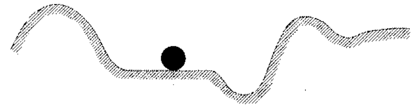


Figure 2.2: Possible equilibrium states of a sphere on a wavelike surface.

of the system. A small disturbance of the ball in position C will bring the ball somewhere between B and D. Such equilibrium states are called *neutrally stable*. In all these situations the disturbances were small, and stability can be referred to as *local stability*. Another situation arises with large disturbances: the ball at position G may not return to the starting position, but to a new equilibrium state. The conclusion is that stability depends on the system characteristics, amplitude of the disturbance, the nature of the disturbance and the initial conditions of the system before the disturbance. It can be envisioned, that the profile is changing with time (such as during an earthquake) or that the mass of the ball changes with time. For example, the E might change in such a way that it is at one moment a local minimum and at another moment local maximum maintaining the zero slope (gradient). The point E will remain the equilibrium point — whether it will be stable or not depends on the moment at which the system is observed. A similar consideration may be applied when the mass of the ball is changing (dynamics is time-varying) and the ball is for instance in equilibrium position G. In this case, the small disturbance will bring the ball to a new equilibrium position (if the mass is small) or back to the same equilibrium position (if the mass is big).

All the discussion up to now emphasizes that analyzing the stability is a very demanding task, especially when everything is subject to change. It is therefore necessary to restrict the observations to certain categories of the system, to which mathematical apparatus enables a solution of the problem to be found. The stability of a linear time-invariant system is easiest to deal with and has been researched for a long time. Such systems possess some advantages — stability depends neither on initial conditions nor on the magnitudes of excitations such as actuating signals, disturbances, measurement noise, etc.

A nonlinear system can be described by a differential equation of the first order:

$$\dot{\mathbf{x}}(t) = \mathbf{f}[t, \mathbf{x}(t), \mathbf{u}(t)];\ \forall t \geq 0,\ \mathbf{u}(t) \neq \mathbf{0} \tag{2.2}$$

where $\mathbf{x}(t)$ is n-dimensional state vector ($\mathbf{x}(t) \in \Re^n$), $\mathbf{u}(t)$ is m-dimensional input vector ($\mathbf{u}(t) \in \Re^m$), t is time ($t \in \Re_x$) and $\mathbf{f}$ is n-dimensional vector of nonlinear functions which are locally Lipschitz ($\mathbf{f} : \Re_+ \times \Re^n \times \Re^m \rightarrow \Re^n$).

The function $\mathbf{f}$ associates to every value of t, $\mathbf{x}(t)$, and $\mathbf{u}(t)$, a corresponding n-dimensional vector $\dot{\mathbf{x}}$; $\mathbf{f} : \Re_+ \times \Re^n \times \Re^m \rightarrow \Re^n$. The fact that (2.2) is a group of first-order differential equations is not a restriction, as every differential equation of n-th order can be rearranged in n differential equations of first order. They can be written in compact form as in (2.2). When the system (2.2) is studied, two problems are possible—analysis and synthesis. If in (2.2) input $\mathbf{u}(t)$ is known, the study of the behavior of $\mathbf{x}(t)$ in (2.2) is called analysis. On the other hand, if (2.2) as well as the desired behavior of $\mathbf{x}(t)$ are known, and input $\mathbf{u}(t)$ is sought, we speak of a synthesis problem. Generally, the input can contain reference signals, disturbances, measurement noise, etc.

If the function $\mathbf{f}(\mathbf{x})$ is locally Lipschitz[5] with respect to the variable $\mathbf{x}(t)$ at the point x_0, then the differential equation (2.2) has an unique solution for an initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$.

2.1.1 Stability of a Nonlinear System Based on Stability of the Linearized System

Stability of a nonlinear system can be analyzed through the stability of a linearized one, but caution must be taken. The linking of the equilibrium states of a linearized system and the original nonlinear system can be seen in an example of an unforced system of second order, with equations:

$$\begin{aligned} \dot{x}_1(t) &= f_1(x_1,x_2) \\ \dot{x}_2(t) &= f_2(x_1,x_2) \end{aligned} \tag{2.3}$$

Linearizing the nonlinear system in the vicinity of the equilibrium state[6], and considering the behavior of the obtained linear system, it is possible — except in specific situations — to analyze the behavior of the nonlinear system in the vicinity of the equilibrium state. Supposing that the equilibrium state of the nonlinear system described by (2.3) is at the origin $[x_{1e}\ x_{2e}]^T = [0\ 0]^T$, and that f_1 and f_2 are continuously differentiable near the origin (equilibrium point), then the Taylor expansion near the origin (equilibrium point) gives:

$$\begin{aligned} \dot{x}_1(t) &= f_1(x_1,x_2) \\ &= f_1(0,0) + a_{11}x_1 + a_{12}x_2 + r_1(x_1,x_2) \\ &= a_{11}x_1 + a_{12}x_2 + r_1(x_1,x_2) \\ \dot{x}_2(t) &= f_2(x_1,x_2) \\ &= f_2(0,0) + a_{21}x_1 + a_{22}x_2 + r_2(x_1,x_2) \\ &= a_{21}x_1 + a_{22}x_2 + r_2(x_1,x_2) \end{aligned} \tag{2.4}$$

where $r_1(x_1,x_2)$ and $r_2(x_1,x_2)$ are higher-order terms of the Taylor series or remainders. As the equilibrium point at the origin is $f_i(0,0) = 0$ for $i = 1,2$ the

[5]See Definition 1.2.
[6]For analytical procedure of linearization see Section 3.3.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Table 2.1: Types of singular points of a nonlinear and a linearized system.

Eigenvalues (λ_i) of linear system (2.5)	Equilibrium state $[z_{1e}\ z_{2e}]^T$ of linear system (2.5)	Equilibrium state $[x_{1e}\ x_{2e}]^T$ of nonlinear system (2.3)
Real and negative	Stable node	Stable node
Real and positive	Unstable node	Unstable node
Real of opposite sign	Saddle	Saddle
Conjugate complex with negative real part	Stable focus	Stable focus
Conjugate complex with positive real part	Unstable focus	Unstable focus
Imaginary (single)	Center	Undefined

linearized model follows:

$$\begin{aligned}\dot{z}_1 &= a_{11}z_1 + a_{12}z_2\\ \dot{z}_2 &= a_{21}z_1 + a_{22}z_2\end{aligned} \tag{2.5}$$

or:

$$\dot{z}(t) = Az(t) \tag{2.6}$$

where:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad z = [z_1\ z_2]^T\ a_{ij} = \left[\frac{\partial f_i}{\partial x_j}\right]_{x=0}; \quad i,j = 1,2.$$

The analytical procedure of linearization is based on the fact that if the matrix A has no eigenvalues with $Re\{\lambda_i\} = 0$, the trajectories of the nonlinear system (2.3) in the vicinity of the equilibrium state $[x_{1e}\, x_{2e}]^T = [0\ 0]^T$ have the same form as the trajectories of the linear system (2.5) in vicinity of the equilibrium state $[z_{1e}\, z_{2e}]^T = [0\ 0]^T$. Table 2.1 shows the types of equilibrium states (points) that are determined from the singular points[7].

If the equilibrium state of the linearized mathematical model (2.5) is of the type center, then the linearized system oscillates with constant amplitude. The behavior of the trajectory of the original nonlinear system (2.3) is determined by the remainder of Taylor series $r_1(x_1,x_2)$ and $r_2(x_1,x_2)$ that were neglected during the linearization process. Analysis of the linearized system alone gives in this case no final answer about the behavior of the nonlinear system. In order to clarify this situation, the following example can be offered.

[7] See Chapter 5.

EXAMPLE 2.1 (VAN DER POL OSCILLATOR)
The Van der Pol oscillator is described with the following nonlinear differential equation:

$$\ddot{y} - \mu(1-y^2)\dot{y} + y(t) = 0; \quad \mu = \text{const.} > 0 \tag{2.7}$$

After choosing the state variables $x_1 = y$, $x_2 = \dot{y}$ *the state-space model is:*

$$\begin{aligned} \dot{x}_1(t) &= x_2 \\ \dot{x}_2(t) &= -x_1 + \mu(1-x_1^2)x_2 \end{aligned} \tag{2.8}$$

Linearization at the equilibrium point $[x_{1e}\; x_{2e}]^T = [0\; 0]^T$ *gives:*

$$\begin{aligned} \dot{z}_1(t) &= z_2 \\ \dot{z}_2(t) &= -z_1 + \mu z_2 \end{aligned}$$

and the linearized system matrix is:

$$A = \begin{bmatrix} 0 & 1 \\ -1 & \mu \end{bmatrix}$$

Eigenvalues of the matrix $\mathbf{A}$ *follow from the characteristic equation* $\mu^2 - \mu\lambda + 1 = 0$. *If* $\mu > 0$, *the eigenvalues have a positive real part, and the equilibrium point of the linearized mathematical model* $[z_{1e}\, z_{2e}]^T = [0\,0]^T$ *is of the type unstable focus* ($\mu \neq 1 > 0$)*. The original nonlinear system (2.8) will have, according to Table 2.1, also an equilibrium state at the origin* $[x_{1e}\, x_{2e}]^T = [0\; 0]^T$ *of the type unstable focus. However, the phase trajectory of the Van der Pol oscillator drawn in Fig. 2.3 demonstrates that the phase trajectory ends in a limit cycle for all initial conditions except for those which are at the origin* $[0\,0]^T$. *For* $\mu = 1$ *a MATLAB program for the simulation of Van der Pol equation with Euler's integrator is:*

```
% Simulation of Van der Pol oscillator
% (Euler's integrator)

axis('square', 1);
axis([-4 4 - 4 4]);
plot(0, 0, '.');
hold

T  = 0.05;        % Integration Interval
mi = 1.0;         % Constant

for x10 = -1.5:1:1.5
  for x20= - 1.5:1:1.5
    x1 = x10;
```

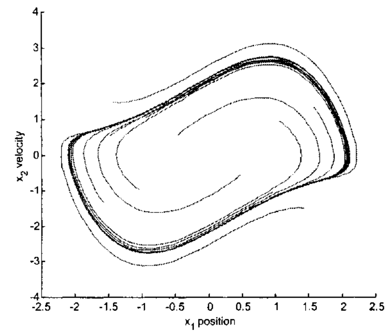


Figure 2.3: Phase trajectories of Van der Pol equation.

```
    x2 = x20;

    % Main loop
    for n = 1:200
      f1 = x2
      f2 =-x1 + mi*(1 - x1*x1);
      x1 = x1 + T*f1;
      x2 = x2 + T*f2;
      plot(x1, x2, '.');
      end
   end
end
xlabel('Position');
ylabel('Velocity')
title ('Simulation of Van der Pol equation
      (Euler method: T=0.05 with different x1(0) and x2(0)');
```

2.2 Lyapunov Stability

Lyapunov introduced the concept of stability in the vicinity of an equilibrium state. It is known that if the unforced system starts from initial conditions which are equal to equilibrium states, the state of the system will not change, i.e. the system will permanently retain its initial equilibrium state under the assumption that no excitation disturbs the system. Therefore if analysis of stability is to be made, the system must start in the vicinity of the equilibrium state. This is justified by the fact that any system can move from the equilibrium state under external signals (measurement noise, disturbances, etc). The Russian mathematician A.M. Lyapunov (1892)[8] proved his theorems by which different stability concepts are defined. Up to now more than twenty definitions of stability have been established, the most important are local stability[9], global stability[10], local asymptotic stability, global asymptotic stability, uniform stability and exponential stability.

The dissertation of A.M. Lyapunov offers two methods for the analysis of stability. The first is the linearization method, where the stability of the nonlinear system is found by analyzing the stability of a linear model that is obtained by linearization of the nonlinear system in the vicinity of the equilibrium state. This method is apt to analyze only the local stability of the nonlinear system. The other method is known as the direct method and is not limited just to local stability. With this method the analysis of stability is made indirectly by analyzing the time behavior of a scalar function that is defined for the given system and that must have certain properties. As the direct method is more general, it will be elaborated in detail in the later text. The first method of linearization was presented in Section 2.1, where in the example of the Van der Pol oscillator (Example 2.1) some problems when applying this method are mentioned. However, it must be emphasized that the method of linearization is nowadays used to justify theoretically the application of the theory of linear systems to real (nonlinear) processes, when the conditions of local deflections from the operating point and of the existence of only one equilibrium state are met.

2.2.1 Definitions of Stability

In this subsection the most important definitions of stability will be presented, accompanied by examples for better understanding. Unforced systems will be described by the vector differential equation (1.17). Lyapunov stability theory considers the behavior of the solution of differential equation (1.17) given by $\mathbf{s}(t,t_0,\mathbf{x}_0)$ when the initial state is not an equilibrium state (i.e. when $\mathbf{x}_0 \neq \mathbf{x}_e$), but is in the vicinity of the equilibrium state. In such cases, local stability or the stability in the small of the equilibrium state is observed. A case can be also considered when the initial state is quite far from the equilibrium state, or when the so-called stability in the whole is treated.

DEFINITION 2.1 (STABILITY IN THE LYAPUNOV SENSE)
Equilibrium state $\mathbf{x}_e = \mathbf{0}$ is stable in the Lyapunov sense if for every $\varepsilon > 0$ and for every $t_0 \in \Re_+$, there exists a positive number $\delta = \delta(\varepsilon, t_0) > 0$, so that

$$\|\mathbf{x}(t_0)\| < \delta(\varepsilon,t_0) \Rightarrow \|\mathbf{s}(t,t_0,\mathbf{x}_0)\| < \varepsilon;\ \forall t \geq t_0 \tag{2.9}$$

[8]English translation of the dissertation was published as late as 1949 and again in 1992, on occasion of the centennial of this significant work (Fuller, 1992).

[9]Local stability is also known as stability in the small.

[10]Global stability is often called stability in the large or in the whole.

is valid.

DEFINITION 2.2 (ASYMPTOTIC STABILITY IN THE LYAPUNOV SENSE)
Equilibrium state $\mathbf{x}_e = \mathbf{0}$ is asymptotically stable in the Lyapunov sense if:

1. *It is stable in the Lyapunov sense and*
2. *a positive number $\delta = \delta(t_0) > 0$, $t_0 \in \Re_+$ exists, so that*

$$\|\mathbf{x}(t_0)\| < \delta(t_0) \Rightarrow \|\mathbf{s}(t,t_0,\mathbf{x}_0)\| \longrightarrow 0;\ \forall t \geq t_0 \tag{2.10}$$

In other words, the disturbed state of the system has a tendency to return to the equilibrium state from which it was disturbed, as $t \longrightarrow \infty$. For asymptotic stability we can write:

$$\lim_{t \longrightarrow \infty} \|\mathbf{x}(t)\| = 0 \tag{2.11}$$

Asymptotic stability is more rigorous than stability in the Lyapunov sense, since it requires that the system returns to the equilibrium state from which it was disturbed at t_0 when observation of the unforced system starts.

In the definitions of stability, the norm $\|x\|$ represents any norm in $\Re^n$, as all the norms in $\Re^n$ are topologically equivalent. That means that the stability of the equilibrium state doesn't depend on type of the norm that is used in the condition of stability.

Stability in the Lyapunov sense can be interpreted graphically in three-dimensional space in the way shown in Fig. 2.4. In the vicinity of the equilibrium state (origin) one can conceive two spheres, one inside the other. The smaller one has radius δ, which depends on the initial time of observations t_0 as well as on the radius of the larger sphere ε. Therefore $\delta = \delta(\varepsilon, t_0) > 0$, where the necessary condition is $\delta = \delta(\varepsilon, t_0) \leq \varepsilon$. It follows that the initial disturbance of the equilibrium state which is by norm smaller than $\delta = \delta(\varepsilon, t_0)$, leads to the conclusion that the initial state is somewhere inside or on the smaller sphere. The equilibrium state of the system (origin) is stable in the Lyapunov sense if the state of the system (state trajectory) remains inside the sphere of larger radius ε. This means that the state trajectory will remain in the vicinity of the equilibrium state, and will not depart from it — it will not increase in norm. In other words, a small excitation (disturbance) around the equilibrium state will result in a small disturbance of the state trajectory.

If the system is asymptotically stable in the Lyapunov sense, then a small disturbance (excitation) will result in the return of the state of the system to its initial equilibrium state from where it was disturbed.

Therefore, the smaller sphere B_δ in Fig. 2.4 may be considered as the region of initial conditions or the permissible region for the excitation, while the larger sphere B_ε is the permissible region of new equilibrium states of the system. Depending on δ, we can speak of local stability or stability in the small if δ is a small

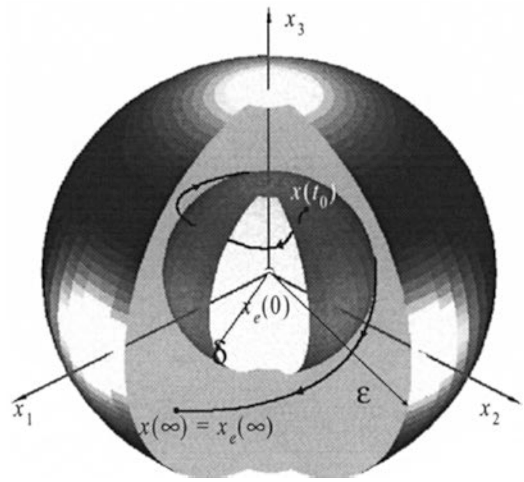


Figure 2.4: Spheres of initial conditions (B_δ) and of new equilibrium states (B_ε).

positive number, while the global stability or stability in the whole is linked with a large positive number δ.

DEFINITION 2.3 (UNIFORM STABILITY)
Equilibrium state $\mathbf{x}_e = \mathbf{0}$ is uniformly stable if for every $\varepsilon > 0$ there exists a positive number $\delta = \delta(\varepsilon)$ so that:

$$\|\mathbf{x}(t_0)\| < \delta(\varepsilon), t_0 \geq 0 \Rightarrow \|\mathbf{s}(t, t_0, \mathbf{x}_0)\| < \varepsilon; \ \forall t \geq t_0 \tag{2.12}$$

If $\delta \neq \delta(t_0)$, we speak of uniform stability. If the system is time-invariant, there is no difference between stability and uniform stability, since the change of initial time t_0 only translates the solution or the trajectory $\mathbf{s}(t, t_0, \mathbf{x}_0)$, respectively, by the same amount. Therefore, nothing changes except for its initial time.

EXAMPLE 2.2
(STABILITY AND UNIFORM STABILITY FOR A TIME-INVARIANT SYSTEM ARE THE SAME)
If a pendulum (Fig. 2.5) is examined, then the motion of the pendulum with mass m can be described by:

$$\ddot{\Theta}(t) + \frac{g}{l}\sin\Theta = 0$$

where g is Earth's gravity, l is the length of the pendulum and Θ is the deflection angle of pendulum from the equilibrium state. By choosing the state variables

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

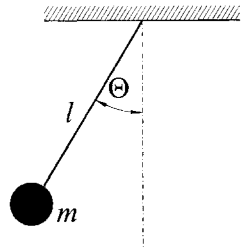


Figure 2.5: Pendulum with the length l and mass m.

$x_1 = \Theta$, $x_2 = \dot{\Theta}$, the equation transforms to:

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= -\frac{g}{l}\sin x_1 \end{aligned} \tag{2.13}$$

The solution (trajectory) has the form:

$$\frac{x_2^2}{2} - \frac{g}{l}\cos x_1 = \frac{x_{20}^2}{2} - \frac{g}{l}\cos x_{10} = a_0 \tag{2.14}$$

The stability of the equilibrium state $[\Theta_e\ \dot{\Theta}_e]^T = [0\ 0]^T$ can be examined by the direct verification of conditions from Definition 2.3. If $\varepsilon > 0$ is given, then the constant $a_0 > 0$ can be chosen so that the trajectory (2.14) lies completely inside the sphere B_ε. If $\delta > 0$ is chosen so that the sphere B_δ lies completely within the curve (trajectory), then the condition from the definition of uniform stability is satisfied:

$$\|\mathbf{x}_e\| < \delta(\varepsilon),\ t_0 \geq 0 \Rightarrow \|\mathbf{s}(t, t_0, \mathbf{x}_0)\| < \varepsilon;\ \forall t \geq t_0 \tag{2.15}$$

The equilibrium state is stable in the Lyapunov sense, since the procedure can be repeated for every $\varepsilon > 0$. Figure 2.6 shows the state trajectory of the pendulum with $x_1(0) = 3$, $x_2(0) = 0$. This example also shows that if a_0 is constant (time-invariant), then the condition from Definition 2.1 will also be satisfied because the state trajectory will not be different for various t_0. We can conclude that stability and uniform stability are the same for time-invariant systems. MATLAB program for simulation of pendulum with double Halijak integrator:

```
% Simulation of pendulum (Halijak double integrator)

axis ([-4 4-4 4 ]);
T = 0.08;          % integration step [s]
```

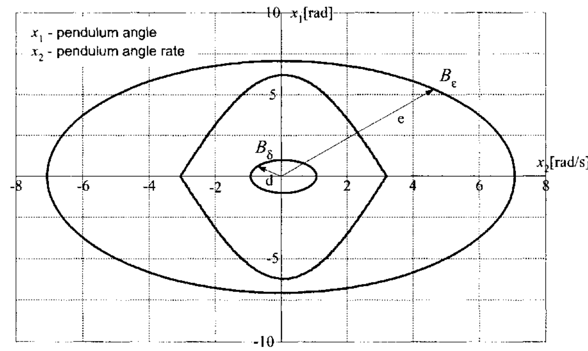


Figure 2.6: State trajectory of a pendulum.

```
g = 9.81;           % Earth's gravity [m/s^{2}]
l = 1;              % length of pendulum [m]

% Initial conditions

th = 3;             % theta(0) - initial angle
thdot=0;            % thetadot(0) - angle rate
plot(theta, (theta - thdot)/T, '.');
hold

% Main loop

for n = 1:300
   f=-(g^{*} sin(theta))/l;
   theta=2^{*} theta -th+T^{*}T^{*}f;
   th=theta;
   theta=theta;
   thdot=(theta-th)/T;
   plot (theta, thdot, '.');
end

xlabel ('Angular deflection [rad]');
ylabel ('Change of angle [rad/s]');
title  ('Simulation of pendulum
         (Halijak integrator: T=0.08) with theta (0)=3');
```


MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

EXAMPLE 2.3
(STABILITY AND UNIFORM STABILITY DIFFER FOR A TIME-VARYING SYSTEM)
The system is given by Massera (1949):

$$\dot{x}(t) = (6t\sin t - 2t)x(t)$$

The solution (trajectory) is of the form:

$$x(t) = x(t_0)\exp\left[6\sin t - 6t\cos t - t^2 - 6\sin t_0 + 6t_0\cos t_0 + t_0^2\right] \tag{2.16}$$

With a selected $t_0 \geq 0$ *we have:*

$$\left|\frac{x(t)}{x(t_0)}\right| = \exp[6sint - 6tcost - t^2 - 6sint_0 + 6t_0cost_0 + t_0 2]$$

If $t-t_0 > 6$, *the exponential function on the right side of the equation will have an upper limit given by:*

$$\exp[12 + T(6-T)]$$

where $T = t - t_0$. *Since this is a continuous function of* t, *it is also limited in the interval* $[t_0, T]$. *Therefore, if the function* $c(t_0)$ *is defined by:*

$$c(t_0) = \sup_{t \geq t_0} \exp\left[6\sin t - 6t\cos t - t^2 - 6\sin t_0 + 6t_0\cos t_0 + t_0^2\right]$$

$c(t_0)$ *will be finite for every fixed* t_0. *As a result, for any* $\varepsilon > 0$, *the condition (2.9) is satisfied if* $\delta = \varepsilon/c(t_0)$ *is chosen, which shows that the origin is a stable equilibrium point in the Lyapunov sense.*

On the other hand, if $t_0 = 2n\pi$, *from (2.16) follows:*

$$x[(2n+1)\pi] = x(2n\pi)\exp[(4n+1)(6-\pi)\pi]$$

and

$$c(2n\pi) \geq \exp[(4n+1)(6-\pi)\pi]$$

$c(2n\pi)$ *is not limited as the function of* t_0. *The result is that for a given* $\varepsilon > 0$ *it is not possible to find such* $\delta(\varepsilon)$, *independent of* t_0, *so that (2.12) is valid. That means the origin is not a uniformly stable equilibrium point.*

When we speak of instability, then the equilibrium state of the system is unstable if Definition 2.1 is not satisfied or the expression (2.9) is not valid. It is wrong to consider that the instability is only when $\|\mathbf{x}(t)\| \to \infty$ as $t \to \infty$ because that is not the only possible behavior of the unstable equilibrium state of the system. Namely, if the origin is an unstable equilibrium state — with $\varepsilon > 0$, $\delta > 0$ cannot be determined so that $\delta \leq \varepsilon$ and (2.9) is valid. The trajectory which leaves the smaller sphere B_δ, where $\mathbf{x}(t_0) = \mathbf{x}_0 \neq \mathbf{x}_e$ in the vicinity of the equilibrium state, will leave the larger sphere B_ε if the equilibrium state is unstable. It can happen

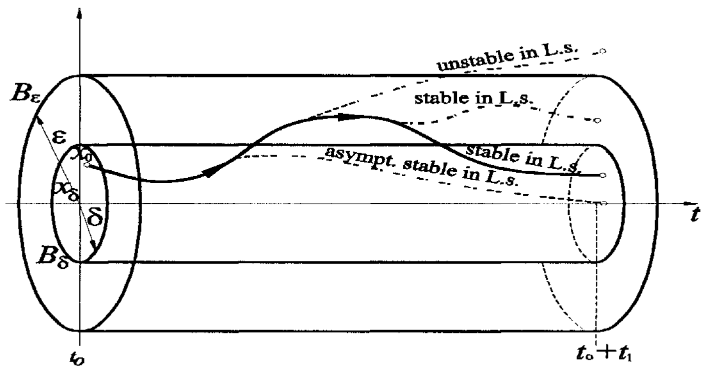


Figure 2.7: Behavior of stable and unstable state trajectory with time.

that the trajectories from B_δ are expanding, but it is not a rule with unstable equilibrium states, as is the case with the Van der Pol equation (Example 2.1) when $\mu = 1$. In Example 2.1 $x_e = [0\ 0]^T$. Trajectories which start with $\mathbf{x}(t_0) = \mathbf{x}_0 \neq \mathbf{x}_e$, regardless of how near they are to the equilibrium state, will finish in a limit cycle (self-oscillations), as can be seen in Fig. 2.3 for Van der Pol equation. If Definition 2.1 is proved, we can see that for sufficiently small $\varepsilon > 0$ can be ascertained that the larger sphere is completely within the limit cycle. Therefore all the trajectories which start from $\mathbf{x}_0 \neq \mathbf{x}_e$ inside the sphere B_ε will leave this sphere, and it will be impossible to find such $\delta > 0$ which will satisfy (2.9). This means that the equilibrium state (origin) is unstable. It can be noticed that all the trajectories of the system are bounded and none is expanding, i.e. $\|\mathbf{x}(t)\| \to \infty$ when $t \to \infty$ is not valid. Essentially, stability in the Lyapunov sense implies that all state trajectories can remain close to the origin (equilibrium state), if their motion starts sufficiently near the origin. That guarantees only that the trajectory will finish at the equilibrium state. If the system is subject to disturbance, it is normal to expect that the system returns to the original equilibrium state, and not to remain in a new equilibrium state. Such stability is called asymptotic stability in the Lyapunov sense. Fig. 2.7 gives the possible state trajectories with their behavior. The difference between stable and unstable equilibrium states is clearly demonstrated.

DEFINITION 2.4 (EXPONENTIAL STABILITY)
Equilibrium state $\mathbf{x}_e = 0$ is exponentially stable if there exist positive constants δ, α and β so that the following expression holds:

$$\|\mathbf{s}(t_0, t_0 + t, \mathbf{x}_0)\| \leq \alpha\ \|\mathbf{x}_0\| \exp(-\beta\ t);\ \forall t, t_0 \geq 0,\ \forall x_0 \in B_\delta \tag{2.17}$$

Exponential stability is more strict then uniform asymptotic stability. If the sphere B_δ has a small radius δ, then local exponential stability is meant. Otherwise

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

when $\delta \to \infty$, global exponential stability is obtained. Exponential stability always implies uniform asymptotic stability.

EXAMPLE 2.4 (LINEAR TIME-VARYING FIRST ORDER SYSTEM)
A linear first order system is considered, given with:

$$\dot{x}(t) = -a(t)x(t)$$

The zero-input response is given by:

$$x(t) = x(t_0)e^{-\int_{t_0}^{t} a(\tau)d\tau}$$

The system will be stable if $a(t) > 0$, $\forall t \geq t_0$. *The system will be asymptotically stable if* $\int_0^\infty a(\tau)d\tau = +\infty$. *The system will be exponentially stable if there exists some positive number T such that* $\forall t \geq 0$, *the expression* $\int_t^{t+T} a(\tau)d\tau \geq \gamma$ *is valid, where* γ *is a positive constant.*

For $a(t) = 1/(1+t)^2$ *the system is stable, but not asymptotically stable.*

For $a(t) = 1/(1+t)$ *the system is asymptotically stable.*

For $a(t)=t$ *the system is exponentially stable.*

In order for an equilibrium state to be globally uniformly asymptotically stable or globally exponentially stable, it is necessary that this is the only possible equilibrium state, so that no other equilibrium states exist. This condition is the chief argument while with the linear systems we speak only of the global asymptotic stability when we discuss either stability of the equilibrium state or we say that the linear system is stable. With nonlinear systems with possibly several equilibrium states, it makes no sense to speak of the system stability. Only with one equilibrium state does the concept of system stability have meaning.

2.2.2 Lyapunov Direct Method

At the end of the 19th century A.M. Lyapunov (1892) developed a method for stability analysis, which is known as the *direct method* or *Lyapunov second method.* The concept of energy, which is much used in engineering practice, can be used for stability analysis. In the 18th century J.L. Lagrange demonstrated how the quantitative behavior of a conservative mechanical system about the equilibrium state can be analyzed by means of potential energy. According to Lagrange, if a function of potential energy of a conservative mechanical system is at a local minimum, the equilibrium state is stable — if it is at a local maximum, the equilibrium state is unstable. Lyapunov's greatness is that he has generalized such reasoning by introducing a function, with certain properties which are similar to those of a function of energy of the equilibrium state. This function is known as the Lyapunov function in his honor. Moreover, the Lyapunov function may not be a function of the energy (e.g. of a mechanical system) but it can represent a generalized function of energy of a mechanical system and as such is applicable to the stability analysis of the equilibrium states for any differential equation or systems which can be described by differential equation.

If an unforced system is considered, its dynamics can be described by equation (1.17). If the system possesses more equilibrium states, one of which is the equilibrium state at the origin ($\mathbf{x}_e = \mathbf{0}$), and if it is possible to define the total energy of the system by a function depending on the state variables, which is everywhere positive except at the origin, then the behavior of the system can be analyzed indirectly by following this function in time. If the analyzed system which is at the equilibrium state $x_e = \mathbf{0}$ is disturbed at the moment $t = t_0$ to a new initial state which is not at the origin ($\mathbf{x}(t_0) \neq \mathbf{x}_e = \mathbf{0}$), the energy of the system should be at any moment t positive. For $t > t_0$, the state of the system can behave in several ways. If the dynamics are such that the energy does not increase with time, then the energy level of the system will not increase above the initial positive value at the moment t_0. Depending on the nature of the function of energy, this is enough to conclude that the equilibrium state $\mathbf{x}_e = \mathbf{0}$ is stable. However, if the energy of the system is monotonically decreasing with the time, even coming to zero, the right conclusion is that the equilibrium state $\mathbf{x}_e = \mathbf{0}$ is asymptotically stable. The third possibility is that after the disturbance the energy increases above the initial value, which means that the equilibrium state $\mathbf{x}_e = \mathbf{0}$ is unstable. Such considerations led Lyapunov to a function which has the characteristics of a function of energy, but is more general than the later one. By checking the properties of this function (Lyapunov function) we can draw conclusions about the stability of the equilibrium state. This can be illustrated in two examples of second-order systems, a linear electrical system and a nonlinear mechanical system.

EXAMPLE 2.5 (UNEXCITED LC OSCILLATOR)
The electrical circuit of an unexcited LC oscillator is shown in Fig. 2.8. The voltage $v(t)$ at the contacts of the capacitor and the current $i(t)$ through the coil are possible state variables. The mathematical model of this circuit is:

$$\dot{x}_1(t) = -\frac{1}{L}x_2(t)$$
$$\dot{x}_2(t) = \frac{1}{C}x_1(t)$$

where $\mathbf{x} = [x_1\ x_2]^T = [i(t)\ v(t)]^T$. The total energy (the sum of electrical and magnetic) at any moment is given by:

$$V(t,\mathbf{x}) = \frac{1}{2}L(t)i^2(t) + \frac{1}{2}C(t)v^2(t) \tag{2.18}$$

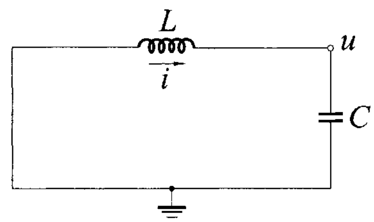


Figure 2.8: Unexcited LC oscillator.

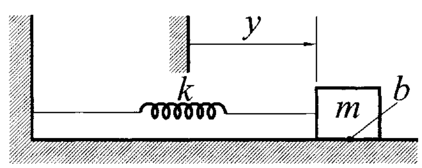


Figure 2.9: Mechanical oscillator.

EXAMPLE 2.6 (MECHANICAL OSCILLATOR)
The mechanical oscillator consists of the mass and the spring (Fig. 2.9). In this example the possible state variables are position of the mass $y(t)$ measured from the end of the relaxed spring and the velocity of the mass dy/dt. Coulomb friction force between the mass and the base is denoted by $b(t)$. The mathematical model is:

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = \frac{-kx_1(t) + b(t)}{m}$$

where $x = [x_1\ x_2]^T = [y(t)\ dy(t)/dt]^T$. The total energy (the sum of kinetic and potential) of this system at any moment is:

$$V(t, \mathbf{x}) = \frac{1}{2}m(t)\dot{y}^2(t) + \frac{1}{2}k(t)y^2(t)$$

From both examples can be concluded that the total energy is a quadratic function of the state variables, which in the case of a time-invariant and unforced system has the form:

$$V(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x} \tag{2.19}$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

where **P** is a symmetric and positive definite matrix. $P = 0.5\begin{bmatrix} L & 0 \\ 0 & C \end{bmatrix}$ for the LC oscillator, and $P = 0.5\begin{bmatrix} k & 0 \\ 0 & m \end{bmatrix}$ for the mechanical oscillator.

The total energy is positive, i.e. $V(\mathbf{x}) > 0, \forall x \neq 0$; also $V(x) = 0$ if and only if $x = 0$. If the energy decreases with time, the energy gradient is always negative:

$$\frac{dV(\mathbf{x})}{dt} = \dot{V}(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial V(\mathbf{x})}{\partial x_i} \cdot \dot{x}_i < 0 \tag{2.20}$$

except for $\mathbf{x} = \mathbf{0}$. The conclusion is that $V(\mathbf{x})$ will continually decrease and eventually approach zero. Thanks to the nature of the energy function, $V(\mathbf{x}) = 0$ implies that $\mathbf{x} = \mathbf{0}$. Therefore, if $dV(\mathbf{x})/dt < 0$; $\forall t$, except for $\mathbf{x} = \mathbf{0}$, it can be said that $\mathbf{x}(t) \to \mathbf{0}$ for a sufficiently large t. The equilibrium state is asymptotically stable in the Lyapunov sense. The same is true if the gradient of the function $V(\mathbf{x})$ is never positive, i.e. $\dot{V}(\mathbf{x}) \leq 0$ — it means that the energy will never increase. This may not signify that the energy will be consumed, i.e. to reach zero. In other words, the energy $V(\mathbf{x})$ and the state $\mathbf{x}$ remain in a certain sense limited, and the equilibrium state is stable in the Lyapunov sense.

The total energy for both oscillators is:

$$V(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x} = 0.5(p_{11}x_1^2 + p_{22}x_2^2) \tag{2.21}$$

In case the coefficients p_{11} and p_{22} are constants independent of time (the systems are time-invariant), the gradient is:

$$\dot{V}(\mathbf{x}) = \sum_{i=1}^{2} \frac{\partial V(\mathbf{x})}{\partial x_i} \cdot \dot{x}_i = p_{11}x_1\dot{x}_1 + p_{22}x_2\dot{x}_2 \tag{2.22}$$

The knowledge of the mathematical model or of equations $x_1 = f_1(x_1)$ and $x_2 = f_2(x_2)$, respectively, enables $V(\mathbf{x})$ and $\dot{V}(\mathbf{x})$ to be expressed as functions of the state variables.

Therefore, if we use this method, the knowledge of the solution of the differential equation is not necessary to conclude about stability. Therefore, it is termed the *Lyapunov direct method.*

The energy gradient for the LC oscillator (Example 2.5) will be:

$$\dot{V}(\mathbf{x}) = Lx_1(\frac{-x_2}{L}) + Cx_2(\frac{x_1}{C}) = 0, \; \forall t \tag{2.23}$$

The LC oscillator is therefore stable in the Lyapunov sense, but is not asymptotically stable. If a resistor is included in the circuit, the system will become dissipative as the energy will be spent in the resistor, $\dot{V}(\mathbf{x})$ will be always negative except for $\mathbf{x} = \mathbf{0}$, and the system will be asymptotically stable in the Lyapunov sense.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

With the mechanical oscillator, the energy gradient is:

$$\dot{V}(\mathbf{x}) = kx_1\dot{x}_1 + mx_2\dot{x}_2 = kx_1x_2 - kx_1x_2 + x_2b = x_2b \tag{2.24}$$

It is obvious that the sign of $\dot{V}(\mathbf{x})$ depends on the signs of the friction force $b(t)$ and of the velocity of the mass $x_2 = \dot{y}(t)$. $\dot{V}(\mathbf{x})$ is not positive if the friction force is always opposite to the motion of the mass — the equilibrium state is stable in the Lyapunov sense. Moreover, it is asymptotically stable in the Lyapunov sense, as can be shown in the example. However, if $b(t)$ and $x_2(t)$ have the same sign, then the gradient $\dot{V}(\mathbf{x})$ is positive, $\dot{V}(\mathbf{x}) > 0$. This means that $V(\mathbf{x})$ can increase. In a specific case, since the spring is moving the mass left and right, the stability problem cannot be judged without additional analysis.

Stability of Time-Invariant Systems

A.M. Lyapunov introduced the scalar function $V(\mathbf{x})$ which can be conceived as a generalized energy function. The fact is that the energy function is often used as a possible Lyapunov function. On the other hand, if the system is described by a mathematical model, it may not be always clear what "energy" means. The conditions which the function $V(\mathbf{x})$ must fulfil in order to be a Lyapunov function are based on mathematical and not physical considerations.

DEFINITION 2.5
(POSITIVE DEFINITE AND POSITIVE SEMIDEFINITE FUNCTION)
A scalar continuous function $V(\mathbf{x})$ which has continuous partial derivatives is positive definite in a region Ω around the origin of state space if

1. $V(\mathbf{0}) = 0$
2. $V(\mathbf{x}) > 0,\ \forall \mathbf{x} \neq \mathbf{0},\ \mathbf{x} \in \Omega$

The function is positive semidefinite if condition 2 weakens, and instead of 2:

3. $V(\mathbf{x}) \geq 0,\ \forall \mathbf{x} \in \Omega.$

Definitions for negative definite and negative semidefinite functions are the same as above, only in 2 greater than zero is replaced by smaller than zero, and in 3 ≥ 0 is replaced by ≤ 0.

The Lyapunov theorems are given without proof, which can be found in Vidyasagar (1993).

THEOREM 2.1
(GLOBAL STABILITY IN LYAPUNOV SENSE–LYAPUNOV DIRECT METHOD)
If it is possible to find a continuous scalar function $V(\mathbf{x})$ which has continuous first derivatives and which satisfies:

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

1. $V(\mathbf{x}) > \mathbf{0}, \forall \mathbf{x} \neq 0$ — $V(\mathbf{x})$ *is positive definite*

2. $\dot{V}(\mathbf{x}) \leq 0$ — $\dot{V}(\mathbf{x})$ *is negative semidefinite*

3. $V(\mathbf{x}) \to \infty$ *as* $\|\mathbf{x}\| \to \infty$ — $V(\mathbf{x})$ *is radially unbounded*

then the equilibrium state $\mathbf{x}_e$ *which satisfies* $\mathbf{f}(\mathbf{x}_e, \mathbf{0}) = \mathbf{0}$ *is globally stable in the Lyapunov sense.*

Conditions 1 and 2 imply that the equilibrium state is stable in the Lyapunov sense—these are the conditions of local stability in the vicinity of the origin. In order that the equilibrium state be globally stable, condition 3 is also necessary.

THEOREM 2.2
(ASYMPTOTIC STABILITY IN LYAPUNOV SENSE–LYAPUNOV DIRECT METHOD)
If it is possible to find a continuous scalar function $V(\mathbf{x})$ *which has continuous first derivatives and which satisfies:*

1. $V(\mathbf{x}) > 0, \forall \mathbf{x} \neq \mathbf{0}$ — $V(\mathbf{x})$ *is positive definite*

2. $\dot{V}(\mathbf{x}) < 0$ — $\dot{V}(\mathbf{x})$ *is negative definite*

3. $V(\mathbf{x}) \to \infty$ *as* $\|\mathbf{x}\| \to \infty$ — $V(\mathbf{x})$ *is radially unbounded*

then the equilibrium state $\mathbf{x}_e$ *which satisfies* $\mathbf{f}(\mathbf{x}_e, \mathbf{0}) = \mathbf{0}$ *is globally asymptotically stable in the Lyapunov sense.*

A scalar function which satisfies conditions 1, 2 and/or 3 is called a *Lyapunov function*. For the stable equilibrium state at the origin of coordinates, and for two-dimensional systems, one possible form of Lyapunov function is given in Fig. 2.10. The graph shows a globally asymptotic equilibrium state, since for every initial condition the trajectory is converging to the origin (equilibrium state). The graphical display which explains different kinds of stability, defined by the Theorems 2.1 and 2.2, and their differences are shown in Fig. 2.11. Here are presented the curves (isohypses) $V(x_1, x_2) = C_i$ where C_i are positive constants. Such a display for $V(\mathbf{x}) = C_i$, for $C_i > 0$ in the plane with the state variables as coordinates, is called *Lyapunov plane* ($n = 2$) or *space* ($n > 2$).

The curves (isohypses) in the Lyapunov plane never intersect since the Lyapunov function is unique. The curves are continuous because the Lyapunov function is continuous as well as all its first partial derivatives. In the example in Fig. 2.11 it is seen that condition 3 is not satisfied. Namely, it is possible that $\|\mathbf{x}\| \to \infty$, while $V(\mathbf{x})$ remains bounded, as is the case with the curve $V(\mathbf{x}) = C_4$ in the figure. It may happen that $V(\mathbf{x})$ continually decreases — the state trajectory which starts, e.g. from initial condition $\mathbf{x}_N(t_0)$ on C_6, over C_5 approaches C_4, while $\|\mathbf{x}\| \to \infty$ the condition of global stability 3 must be valid if such behavior of the trajectory is to be excluded. If the trajectory starts from the initial condition $\mathbf{x}_S(t_0)$ on C_3

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

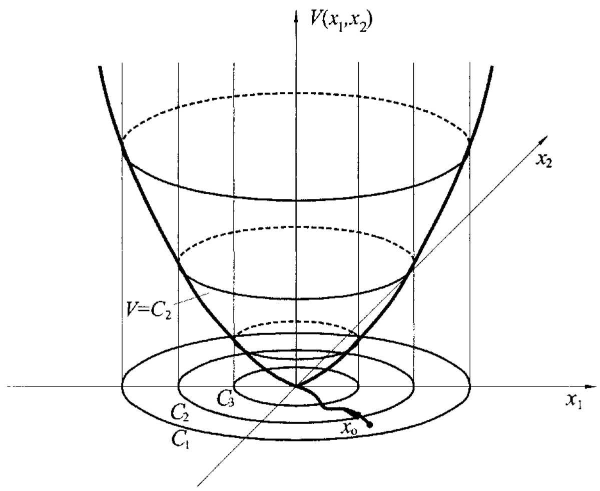


Figure 2.10: Possible form of Lyapunov function for a second-order system.

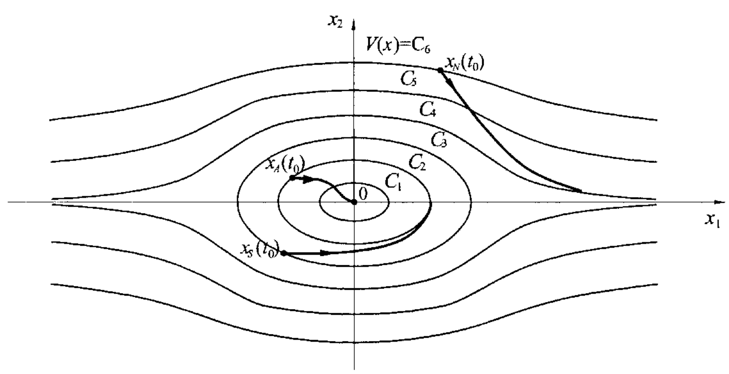


Figure 2.11: Example of a possible Lyapunov plane for a second-order system.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

and ends at $[x_1\ x_2]^T = [C_2\ 0]^T$, then the equilibrium state $\mathbf{x}_e = \mathbf{0}$ is locally stable in the Lyapunov sense because of $\dot{V}(\mathbf{x}) \leq 0$. In this example local stability allowed only those initial states in the vicinity of the origin which are inside the region bounded by the curve C_3. As can be seen, in this case the new equilibrium state is bounded, but the trajectory may not end at the origin. On the other hand, the trajectory which starts from $\mathbf{x}_A(t_0)$ on C_2 and ends at the origin suggests that the origin is the equilibrium state, which is locally asymptotically stable in the Lyapunov sense since a strict condition 2 of the Theorem 2.2, i.e. $\dot{V}(\mathbf{x}) < 0, \forall \mathbf{x} \neq \mathbf{0}$, must be respected.

If we only know that $\dot{V}(\mathbf{x}) \leq 0$, it cannot be concluded with confidence that the trajectory will approach to the origin — nevertheless, we can conclude that the origin is stable in the Lyapunov sense, as the trajectory remains inside the sphere B_ε (the vicinity of origin) and the initial conditions $\mathbf{x}(t_0)$ were in Lyapunov space which includes the vicinity of origin. When $\dot{V}(\mathbf{x}) \leq 0$, the asymptotic stability of the origin is presumed only if we can prove that no solution (trajectory) can forever remain in the set $\{\dot{V}(\mathbf{x}) = 0\}$, except for the trivial case $\mathbf{x}(t) = \mathbf{0}$. Under this condition $V(\mathbf{x})$ must decrease towards zero — therefore $\mathbf{x}(t) \to \mathbf{0}$ when $t \to \infty$. Such extension of the basic stability Theorem 2.1 is called the principle of invariance.

THEOREM 2.3 (LA SALLE PRINCIPLE OF INVARIANCE)
If there exists a positive definite scalar continuous function $V : \Re_+ \times \Re^n \to \Re_+$, whose gradient is also a continuous but negative semidefinite scalar function, and supposing the origin is the only equilibrium state of the system $\mathbf{f}(t, \mathbf{0}) = \mathbf{0}$, $\forall t$, and if supposing that the

$$\Omega = \left\{\mathbf{x} \in \Re^n;\ \dot{V}(\mathbf{x}) = 0\right\} \tag{2.25}$$

does not include the trivial solution (trajectory) of the system, then the origin is globally asymptotically stable.

The problem with the Lyapunov direct method is that there is not just one Lyapunov function. It is possible to find several Lyapunov functions — the consequence is that if a Lyapunov function was not found, that doesn't mean that it is nonexistent, or that the equilibrium state under investigation is unstable. In order to make a clear difference between the chosen function $V(\mathbf{x})$ which is a possible choice for the Lyapunov function and Lyapunov function itself which satisfies the conditions 1, 2 and/or 3, every function which is a candidate for Lyapunov function will be called a *possible* Lyapunov function or Lyapunov function candidate.

Only when the possible Lyapunov function satisfies conditions 1, 2 and/or 3 it becomes a Lyapunov function. Lyapunov theorems give no indication how to find a Lyapunov function. As there are several Lyapunov functions for a specific system, obviously some will be better than others. It can happen that for a particular system, Lyapunov function $V_1(\mathbf{x})$ which indicates local asymptotic stability in the Lyapunov sense is found for initial states in the vicinity of the equilibrium

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

state (origin), some other Lyapunov function $V_2(\mathbf{x})$ may indicate local stability in the Lyapunov sense and a third Lyapunov function indicates global asymptotic stability in the Lyapunov sense.

We can say that if the equilibrium state is stable, there surely exists a Lyapunov function, and the question is if we can find it.

There doesn't exist a universally accepted method to seek a Lyapunov function. The form of a possible Lyapunov function is usually assumed, either by a mere presumption or by knowing the physical system or by analyzing the system. After an assumed form of $V(\mathbf{x})$, the gradient of the function is tested by means of a mathematical model of the system $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x})$.

EXAMPLE 2.7 (PENDULUM FROM EXAMPLE 2.2)
For a pendulum described by the equation (2.13), a possible Lyapunov function can be sought in the form of a function of a total system's energy (sum of the potential and the kinetic one), which can be expressed as:

$$V(x_1,x_2) = \frac{g}{l}(1-\cos x_1) + \frac{1}{2}x_2^2$$

where the first term represents potential, and the second kinetic energy. The chosen scalar function $V(x_1,x_2)$ is continuous with continuous first partial derivatives, and is positive definite. Therefore, it is a possible Lyapunov function. The gradient of $V(\mathbf{x})$ is:

$$\dot{V}(x_1,x_2) = \frac{g}{l}\sin x_1 \cdot \dot{x}_1 + x_2 \cdot \dot{x}_2 = \frac{g}{l}\sin x_1(x_2) + x_2(-\frac{g}{l}\sin x_1) = 0$$

The gradient of the possible Lyapunov function is a negative semidefinite function. It can be concluded that $V(x_1,x_2)$ is a Lyapunov function as it satisfies the necessary conditions. Equilibrium state $\mathbf{x}_e = \mathbf{0}$ is locally stable as the conditions 1 and 2 of the Theorem 2.1 are satisfied, while the condition 3 is not.

It is also possible to anticipate in the first place the form of the gradient of a possible Lyapunov function, either $\dot{V}(x)$ or $\nabla_x V(\mathbf{x})$. The form of $V(\mathbf{x})$ can be obtained by integration. Thereafter, the conditions for a Lyapunov function are tested.

For linear systems it is easier to determine a Lyapunov function, since it must belong to the class of quadratic functions $V(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x}$. For nonlinear systems various methods to find Lyapunov function are developed. All of them assume a certain form of the mathematical model. Here are mentioned the methods of A.I. Lur'e (1957), N.N. Krasovskii (1963), D.G. Schultz and J.E. Gibson (1962), V.I. Zubov (1961), V.M. Popov (1973), etc.

Forming of Possible Lyapunov Functions

To find a Lyapunov function is quite a demanding task, especially with nonlinear systems. Here are given some procedures which make it easier. Once again it

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

must be emphasized that no general procedure exists to find a simple Lyapunov function for a nonlinear system.

1. Method — Linear System

For a linear unforced system described by $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$, $\mathbf{x} \in \Re^n$ a possible Lyapunov function can be sought in the form $V(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x}$, where $\mathbf{P}$ is a positive definite symmetric matrix. As a positive definite scalar continuous function is in question, which has continuous first partial derivatives, it is necessary to see the condition for this system in order to be stable in the Lyapunov sense. The gradient of a possible Lyapunov function is:

$$\dot{V}(\mathbf{x}) = \dot{\mathbf{x}}^T\mathbf{P}\mathbf{x} + \mathbf{x}^T\mathbf{P}\dot{\mathbf{x}} = \mathbf{x}^T\mathbf{A}^T\mathbf{P}\mathbf{x} + \mathbf{x}^T\mathbf{P}\mathbf{A}\mathbf{x} = \mathbf{x}^T(\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A})\mathbf{x} \tag{2.26}$$

If it must be a negative definite function, then it can be said that:

$$\dot{V}(\mathbf{x}) = \mathbf{x}^T(\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A})\mathbf{x} = -\mathbf{x}^T\mathbf{Q}\mathbf{x} < 0 \tag{2.27}$$

$\dot{V}(\mathbf{x}) < 0$ if and only if Q is a positive definite symmetric matrix. Then a possible Lyapunov function becomes a true one. Therefore for a linear unforced system the equilibrium state (origin) $\mathbf{x}_e = \mathbf{0}$ will be globally asymptotically stable in the Lyapunov sense if:

$$\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q} \tag{2.28}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are positive definite symmetric matrices. Equation (2.28) is known as the Lyapunov equation.

The stability analysis of the linear system by the Lyapunov direct method unfolds in the following manner: first a symmetric matrix $\mathbf{Q}$ is chosen, which must be positive definite. Usually is $\mathbf{Q}$ an identity matrix. Secondly, from (2.28) a symmetric matrix $\mathbf{P}$ is calculated. If it is positive definite, the system is globally asymptotically stable in the Lyapunov sense.

DEFINITION 2.6 (POSITIVE DEFINITE MATRIX)
Necessary and sufficient conditions for a real ($\mathbf{P}$) and symmetric ($\mathbf{P} = \mathbf{P}^T$) matrix to be positive definite are:

1. $\mathbf{x}^T\mathbf{P}\mathbf{x} > 0, \forall \mathbf{x} \neq \mathbf{0}$,
2. *All eigenvalues of the matrix* $\mathbf{P}$ *satisfy* $\lambda_i(\mathbf{P}) > 0$,
3. *All upper left minors (submatrices)* $\mathbf{P}_k$ *have positive determinants,*
4. *A matrix* $\mathbf{W}$ *exists, such that* $\mathbf{P} = \mathbf{W}^T\mathbf{W}$ *is valid.*

A positive definite matrix is both symmetric and strictly positive.

DEFINITION 2.7 (POSITIVE SEMIDEFINITE MATRIX)
A real symmetric matrix is positive semidefinite if the scalar (quadratic form) satisfies $\alpha = \mathbf{x}^T\mathbf{P}\mathbf{x} \geq 0,\ \forall \mathbf{x} \in \Re^n$.

DEFINITION 2.8 (STRICTLY POSITIVE MATRIX)
An asymmetric real matrix $\mathbf{A} \neq \mathbf{A}^T$ *is strictly positive if* $\frac{1}{2}(\mathbf{A}+\mathbf{A}^T)$ *is a positive definite matrix. The matrix* $\mathbf{A}$ *can be expressed in the following way:*

$$\mathbf{A} = \frac{1}{2}(\mathbf{A}+\mathbf{A}^T) + \frac{1}{2}(\mathbf{A}-\mathbf{A}^T) \tag{2.29}$$

where the first addend $(0.5\mathbf{A}+0.5\mathbf{A})^T$ *is a symmetric matrix, and the second one* $(0.5\mathbf{A}-0.5\mathbf{A})^T$ *is skew symmetric matrix (*$\mathbf{A}^T = -\mathbf{A}$, *or* $a_{ij} = -a_{ji},\ \forall i,j$ *with all the elements on the main diagonal equal to zero). Quadratic form* $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0.5\mathbf{x}^T(\mathbf{A}+\mathbf{A}^T)\mathbf{x}$ *must be a positive scalar, for all* $\mathbf{x} \neq \mathbf{0}$, *in order that the matrix* $\mathbf{A}$ *should be strictly positive.*

The importance of a symmetric strictly positive matrix can be illustrated on the example of the stability analysis of a linear time-varying system.

EXAMPLE 2.8 (STABILITY OF TIME-VARYING SYSTEM)
Examined is a linear time-varying system:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) \tag{2.30}$$

given by:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} -1 & e^{2t} \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

It is well known that the linear time-invariant systems are asymptotically stable if the system matrix $\mathbf{A}$ *has eigenvalues in the open left half plane of the complex plane, or* $\mathrm{Re}\{\lambda_i\} < 0$. *Such properties of the linear time-invariant system lead to the conclusion that the linear time-varying system will be stable when for every* $t \geq 0$ *the eigenvalues of the matrix* $\mathbf{A}(t)$ *have negative real values. This suggests that the stability analysis of the linear time-varying systems is very simple. Alas, this is not true as this example demonstrates. Namely, for all* $t \geq 0$ *the eigenvalues of the matrix* $\mathbf{A}(t)$ *are the same, i.e.* $\lambda_{1,2} = -1$. *However, when the differential equations of the system are solved for the state* $x_2(t)$:

$$\dot{x}_2(t) = -x_2(t) \rightarrow x_2(t) = x_2(0)e^{-t} \tag{2.31}$$

and for the state $x_1(t)$:

$$\dot{x}_1(t) + x_1(t) = e^{2t}x_2(t) \tag{2.32}$$

By substitution of (2.31) in (2.32) follows:

$$\dot{x}_1(t) + x_1(t) = x_2(0)e^{t}$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

which gives unstable behavior of $x_1(t)$. The function $x_1(t)$ can be conceived as the output of a first-order filter whose input is unbounded, i.e. equal to $x_2(0)e^t$. The example shows that the logic that was valid for the time-invariant systems cannot be applied to the time-varying ones.

A linear time-varying system given by (2.30) will be asymptotically stable if the eigenvalues of the symmetric matrix $\mathbf{A}(t)+\mathbf{A}^T(t)$ which are all real remain strictly in the left half of the complex plane. There exist such $\lambda_i > 0$, $\forall i, \forall t \geq 0$, so that:

$$\lambda_i \left[\mathbf{A}(t)+\mathbf{A}^T(t)\right] \leq -\lambda$$

This can be demonstrated by means of the Lyapunov function in the quadratic form $V(\mathbf{x}) = \mathbf{x}^T\mathbf{x}$, because:

$$\dot{V}(\mathbf{x}) = \mathbf{x}^T\dot{\mathbf{x}} + \dot{\mathbf{x}}^\mathbf{T}\mathbf{x} = \mathbf{x}^T\left[\mathbf{A}(t)+\mathbf{A}^T(t)\right]\mathbf{x} \leq -\lambda\mathbf{x}^T\mathbf{x} = -\lambda V(\mathbf{x})$$

and then:

$$\forall t \geq 0, \quad 0 \leq \mathbf{x}^T\mathbf{x} = V(t) \leq V(0)e^{-\lambda t}$$

As $\mathbf{x}$ decays exponentially to zero the system is asymptotically stable.

EXAMPLE 2.9
(STABILITY OF A LINEAR SYSTEM BY THE LYAPUNOV DIRECT METHOD)
Investigated is a linear system given by:

$$\begin{aligned}\dot{x}_1(t) &= -x_1(t) - 2x_2(t)\\ \dot{x}_2(t) &= x_1(t) - 4x_2(t)\end{aligned}$$

The equilibrium state of the system at the origin is $\mathbf{x}_e = \mathbf{0}$. If $\mathbf{Q} = \mathbf{I}$ is chosen, the Lyapunov equation will be:

$$\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{I}$$

Under assumption [11] *of $p_{12} = p_{22}$ follows:*

$$\begin{aligned}-2p_{11} + 2p_{12} &= -1\\ -2p_{11} - 5p_{12} + p_{22} &= 0\\ -4p_{12} - 8p_{22} &= -1\end{aligned}$$

Therefore the matrix $\mathbf{P}$ is:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} \frac{23}{60} & \frac{-7}{60} \\ \frac{-7}{60} & \frac{11}{60} \end{bmatrix}$$

Positive definiteness of the matrix $\mathbf{P}$ can be examined in several ways (see Definition 2.6). If the left subdeterminants of the matrix (Sylvester method — see 3 in

[11] $\mathbf{P}$ matrix is symmetric.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

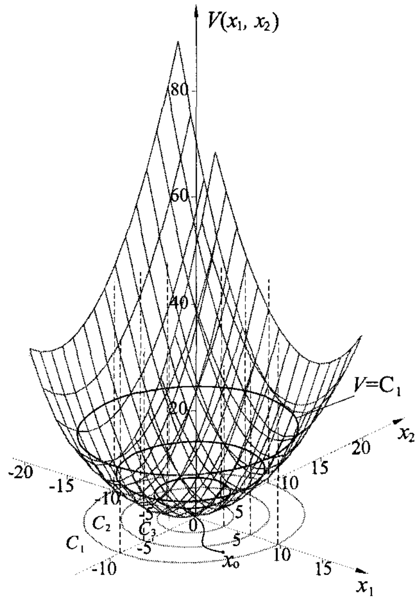


Figure 2.12: Display of Lyapunov function $V(\mathbf{x})$ for Example 2.9.

Definition 2.6) are examined, which is not a recommended procedure[12]*, the result is that* $\det(\mathbf{P}_1) = p_n = 23/60 > 0$ *and also* $\det(\mathbf{P}_2) = \det(\mathbf{P}) = 302/60 > 0$. *It can be concluded that* $\mathbf{P}$ *is a positive definite matrix, or the equilibrium state* $\mathbf{x}_e = \mathbf{0}$ *is asymptotically stable in the Lyapunov sense. The Lyapunov function is:*

$$V(\mathbf{x}) = [x_1\ x_2] \cdot \begin{bmatrix} \frac{23}{60} & \frac{-7}{60} \\ \frac{-7}{60} & \frac{11}{60} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{60}(23x_1^2 - 14x_1x_2 + 11x_2^2)$$

while the gradient $\dot{V}(\mathbf{x})$ *is:*

$$\dot{V}(\mathbf{x}) = -\|x\|^2 = -(x_1^2 + x_2^2) \tag{2.33}$$

A graphical display of $V(\mathbf{x})$ *is given in Fig. 2.12.*

[12]For numerical reasons, due to calculation of subdeterminants.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

2. Method — Krasovskii procedure (Krasovskii, 1963)

In an attempt to generalize the procedure for a linear system, Krasovskii examines an unforced, time-varying nonlinear system described by:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}),\ \mathbf{x} \in \Re^n$$

where the function $\mathbf{f}$ is differentiable and for which $\mathbf{f}(\mathbf{0}) = \mathbf{0}$. The Jacobian matrix of the system is:

$$\begin{aligned} \mathbf{J}(\mathbf{x}) = (\frac{\partial f_i}{\partial x_j})_{\substack{1\le i\le n \\ 1\le j\le n}} &= (\nabla f_i)_{1\le i\le n} \\ &= \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \end{aligned}$$

where $\nabla f_i = \left[\frac{\partial f_i}{\partial x_1} \cdots \frac{\partial f_i}{\partial x_n}\right]$ is a row vector. Assuming that $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ (the origin may not be the only equilibrium state), and if the Jacobian is symmetrically arranged, i.e. $\hat{\mathbf{J}}(\mathbf{x}) = \mathbf{J}^T(\mathbf{x}) + \mathbf{J}(\mathbf{x})$, then the origin will be asymptotically stable in the Lyapunov sense if $\hat{\mathbf{J}}(\mathbf{x})$ is negative definite matrix for every $\mathbf{x}$. The Lyapunov function is:

$$V(\mathbf{x}) = \mathbf{f}^T(\mathbf{x})\mathbf{P}\mathbf{f}(\mathbf{x}) \tag{2.34}$$

where $\mathbf{P}$ is a symmetric positive definite matrix. Now:

$$\dot{V}(\mathbf{x}) = \frac{dV(\mathbf{x})}{dt} = \dot{\mathbf{f}}^T(\mathbf{x})\mathbf{P}\mathbf{f}(\mathbf{x}) + \mathbf{f}^T(\mathbf{x})\mathbf{P}\dot{\mathbf{f}}(\mathbf{x}) \tag{2.35}$$

where:

$$\dot{\mathbf{f}}(\mathbf{x}) = \frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial t} = \mathbf{J}(\mathbf{x})\mathbf{f}(\mathbf{x}) \tag{2.36}$$

It follows:

$$\begin{aligned} \dot{V}(\mathbf{x}) = \frac{dV(\mathbf{x})}{dt} &= \mathbf{f}^T(\mathbf{x})\mathbf{J}^T(\mathbf{x})\mathbf{P}\mathbf{f}(\mathbf{x}) + \mathbf{f}^T(\mathbf{x})\mathbf{P}\mathbf{J}(\mathbf{x})\mathbf{f}(\mathbf{x}) \\ &= \mathbf{f}^T(\mathbf{x})(\mathbf{J}^T\mathbf{P} + \mathbf{P}\mathbf{J})\mathbf{f}(\mathbf{x}) \end{aligned} \tag{2.37}$$

Let it be $\mathbf{Q} = \mathbf{J}^T\mathbf{P} + \mathbf{P}\mathbf{J}$. As $\mathbf{P}$ is a symmetric positive definite matrix, it follows that $\mathbf{Q}$ must be a negative definite one in order that the equilibrium state be asymptotically stable in the Lyapunov sense. Moreover, if $V(\mathbf{x}) = \mathbf{f}^T(\mathbf{x})\mathbf{P}\mathbf{f}(\mathbf{x}) \to \infty$ when $\|\mathbf{x}\| \to \infty$, then the origin is globally asymptotically stable.

3. Method — method of varying gradient (Schultz and Gibson, 1962)

This procedure is based on the assumption that the gradient of a possible Lyapunov function has a strictly determined form given by (2.43). There appear the functions denoted by α_{ij} $(i,j = 1,2,...,n)$, which must be chosen in such a way that the negative definiteness or semidefiniteness of the scalar function $\dot{V}(\mathbf{x})$ is ensured — thereafter a possible Lyapunov function can be determined by integration.

Consider an unforced time-invariant nonlinear system given with $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x})$, $\mathbf{x} \in \Re^n$, for which $\mathbf{f}(\mathbf{0}) = \mathbf{0}$, and assume that $V(\mathbf{x})$ is a possible Lyapunov function for this system. If $\mathbf{x}(t) : \Re^+ \rightarrow \Re^n$ is any differentiable function with $\mathbf{x}(\mathbf{0}) = \mathbf{0}$, then for any possible Lyapunov function:

$$\frac{dV(\mathbf{x})}{dt} = \dot{V}(\mathbf{x}) = (\nabla_\mathbf{x} V)^T \dot{\mathbf{x}} = (\nabla_\mathbf{x} V)^T \mathbf{f}(\mathbf{x}) \tag{2.38}$$

The posible Lyapunov functions could be obtained by integration, respectively:

$$V[\mathbf{x}(t)] = \int_0^t (\nabla_\mathbf{x} V)^T \frac{d\mathbf{x}}{d\tau} d\tau = \int_0^{\mathbf{x}(t)} (\nabla_\mathbf{x} V)^T d\mathbf{x} \tag{2.39}$$

where $\int_0^{\mathbf{x}(t)} (...) \, d\mathbf{x}$ denotes a linear integral along the line $L : t \rightarrow \mathbf{x}(t)$. The line integral is independent of the path of integration in $\Re^n$ between 0 and $\mathbf{x}(t)$ if the symmetry condition is valid:

$$\frac{\partial \nabla V_i}{\partial x_j} = \frac{\partial \nabla V_j}{\partial x_i} \quad (\text{for } i,j = 1,2,\ldots,n) \tag{2.40}$$

where the i-th component $\nabla_\mathbf{x} V$ is denoted by ∇V_i. The condition (2.40) can be also expressed by the symmetry condition of other partial derivatives of a possible Lyapunov function:

$$\nabla_\mathbf{x}(\nabla_\mathbf{x} V)^T = \left[\frac{\partial^2 V}{\partial x_i \partial x_j}\right] = \left[\frac{\partial^2 V}{\partial x_j \partial x_i}\right] \quad (\text{for } i,j = 1,2,\ldots,n) \tag{2.41}$$

The symmetry conditions enable us to uniquely define a possible Lyapunov function as a sum of n scalar integrals:

$$\begin{aligned} V(\mathbf{x}) &= \int_L (\nabla_\mathbf{x} V)^T d\mathbf{x}, \; x \in \Re^n \\ V(\mathbf{x}) &= \int_0^{x_1} (\nabla V)_1 (x_1', 0, \ldots, 0) dx_1' \\ &+ \int_0^{x_2} (\nabla V)_2 (x_1, x_2', 0, \ldots, 0) dx_2' \\ &+ \ldots + \int_0^{x_n} (\nabla V)_n (x_1, x_2, ..., x_n') dx_n' \end{aligned} \tag{2.42}$$

for any path which links 0 and $\mathbf{x}(t)$, where $(\nabla V)_i = \partial V/\partial x_i$.

It is expected that $V(\mathbf{x})$ has a quadratic form, and $\nabla_{\mathbf{x}} V(\mathbf{x})$ can be written as:

$$\nabla_{\mathbf{x}} V(\mathbf{x}) = \begin{bmatrix} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n \\ \vdots \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \ldots + \alpha_{nn}x_n \end{bmatrix} \tag{2.43}$$

where α_{ij} may not be constants, but the functions of the components of the state vector $\mathbf{x}(t)$.

If the freedom of choice of the function $V(\mathbf{x})$ is to be increased, it must be allowed that d_{ij} of the coefficients depends on $x_1, x_2, \ldots, x_{n-1}$. As $\dot{V}(\mathbf{x}) = (\nabla_{\mathbf{x}} V)^T \dot{\mathbf{x}} = (\nabla_{\mathbf{x}} V)^T \mathbf{f}(x)$, $\dot{V}(\mathbf{x})$ can be expressed through a_{ij} and $\mathbf{f}(\mathbf{x})$. The procedure to find a possible Lyapunov function consists of the following: first to assume $\nabla_{\mathbf{x}} V$ as in (2.43), where some of the elements a_{ij} are chosen in such a way that the gradient of $V(\mathbf{x})$ is negative definite or semidefinite scalar function, while the remaining a_{ij} are chosen so that the expression (2.40) is satisfied. Finally, the function $V(\mathbf{x})$ can be evaluated from (2.42), and its positive definiteness determined.

It is possible to equate the elements a_{ij} to zero. Generally α_{nn} is assumed a constant number, the most often $\alpha_{nn} = 1$ or $\alpha_{nn} = 2$, in order that $V(\mathbf{x})$ will be a quadratic scalar function in regard to x_n.

EXAMPLE 2.10
(LYAPUNOV FUNCTION BY METHOD OF VARYING PARAMETER)
The system is given by the block diagram in Fig. 2.13, where the nonlinear element has a cubic characteristic $y_n = F(x) = x^3$. The process is described by the transfer function $G_L(s) = \frac{1}{s(s+1)}$ and a possible mathematical model in the state variables $x_1(t) = y(t); x_2(t) = dy(t)/dt$ is:

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= -x_1^3(t) - x_2(t) \end{aligned} \tag{2.44}$$

Gradient $V(\mathbf{x})$ is assumed to have the form given by (2.43), or:

$$\nabla_{\mathbf{x}} V(\mathbf{x}) = \begin{bmatrix} \alpha_{11}x_1 + \alpha_{12}x_2 \\ \alpha_{21}x_1 + 2x_2 \end{bmatrix} \tag{2.45}$$

In order to assure the quadratic form of a possible Lyapunov function with regard to the state variable x_2, the value $\alpha_{22} = 2$ is taken. The condition of symmetry (2.40) requires that:

$$\frac{\partial \nabla V_1}{\partial x_2} = \frac{\partial \nabla V_2}{\partial x_1}$$

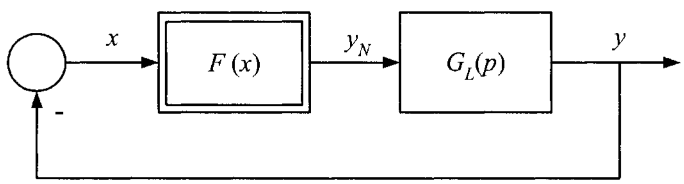


Figure 2.13: Block diagram of a nonlinear system.

or:

$$x_1\frac{\partial \alpha_{11}}{\partial x_2}+\alpha_{12}+x_2\frac{\partial \alpha_{12}}{\partial x_2}=x_1\frac{\partial \alpha_{21}}{\partial x_1}+\alpha_{21}+2x_2$$

Using the assumed form of the gradient (2.45) as well as the equation of the system's dynamics (2.44), it follows that:

$$\begin{aligned}\dot{V}(\mathbf{x}) &= (\nabla_{\mathbf{x}}V)\dot{x}=(\nabla_{\mathbf{x}}V)^T\begin{bmatrix} x_2 \\ -x_1^3-x_2 \end{bmatrix} \\ &= [\alpha_{11}x_1+\alpha_{12}x_2 \quad \alpha_{21}x_1+2x_2]\begin{bmatrix} x_2 \\ -x_1^3-x_2 \end{bmatrix} \\ &= (\alpha_{11}-2x_1^2-\alpha_{21})x_1x_2+(\alpha_{12}-2)x_2^2-\alpha_{21}x_1^4 \end{aligned} \tag{2.46}$$

With the proper choice of α_{ij} (for $i,j=1,2$), the scalar function (2.46) must be rendered negative definite or semidefinite. If the first addend is to be eliminated (term with x_1x_2), $\alpha_{11}=\alpha_{21}+2x_1^2$ has to be chosen. Obviously, the second addend (term with x_2^2) will be negative definite only if $0<\alpha_{12}<2$, while the third addend (term with x_1^4) will be negative semidefinite if $\alpha_{21}>0$. With the choice of $\alpha_{12}=1$, (2.46) will be:

$$\dot{V}(\mathbf{x})=-x_2^2-\alpha_{21}x_1^4$$

choosing further α_{ij}, $(i,j=1,2)$, the expression (2.45) is converted to:

$$\nabla_{\mathbf{x}}V=\begin{bmatrix} \alpha_{21}x_1+2x_1^3+x_2 \\ \alpha_{21}x_1+2x_2 \end{bmatrix}$$

The condition of symmetry is:

$$\frac{\partial^2 V}{\partial x_2 \partial x_1}=1$$

$$\frac{\partial^2 V}{\partial x_1 \partial x_2}=\alpha_{21}+x_1\frac{\partial \alpha_{21}}{\partial x_1}$$

The conditions (2.41) will be satisfied when $\alpha_{21} = 1$, *and* $\nabla_{\mathbf{x}} V(\mathbf{x})$ *will be:*

$$\nabla_x V = \begin{bmatrix} x_1 + 2x_1^3 + x_2 \\ x_1 + 2x_2 \end{bmatrix}$$

The possible Lyapunov function is:

$$\begin{aligned} V(\mathbf{x}) &= \int_0^{x_1} (x_1' + 2x_1'^3)\, dx_1' + \int_0^{x_2} (x_1 + 2x_2')\, dx_2' \\ &= \frac{x_1^2}{2} + \frac{x_1^4}{2} + x_1 x_2 + x_2^2 \\ &= \frac{(x_1 + x_2)^2}{2} + \frac{x_1^4}{2} + \frac{x_2^2}{2} \end{aligned}$$

This function satisfies the conditions 1 and 2 of the Theorem 2.1; therefore, it is a Lyapunov function.

EXAMPLE 2.11
(POSSIBLE LYAPUNOV FUNCTION OF VAN DER POL EQUATION)
A possible Lyapunov function for the Van der Pol equation (see Example 2.1) can be obtained using the method of varying gradient. For equation (2.8) it is assumed that the form of the varying gradient is like (2.43), or:

$$\nabla_{\mathbf{x}} V = \begin{bmatrix} \alpha_{11} x_1 + \alpha_{12} x_2 \\ \alpha_{21} x_1 + \alpha_{22} x_2 \end{bmatrix}$$

Here $\alpha_{22} = 1$ *is chosen in order to get the quadratic form of a possible Lyapunov function with regard to* $x_2(t)$. *The condition of symmetry is now:*

$$x_1 \frac{\partial \alpha_{11}}{\partial x_2} + \alpha_{12} + x_2 \frac{\partial \alpha_{12}}{\partial x_2} = x_1 \frac{\partial \alpha_{21}}{\partial x_1} + \alpha_{21}$$

The gradient of a possible Lyapunov function is:

$$\begin{aligned} &\dot{V}(\mathbf{x}) = (\nabla_{\mathbf{x}} V)^T \dot{x}(t) = \\ &\qquad (\alpha_{11} - 1 + \alpha_{21}\mu) x_1 x_2 + (\alpha_{12} + \mu) x_2^2 - \alpha_{21} x_1^2 - \mu x_1^2 x_2^2 - \alpha_{21} \mu x_1^3 x_2 \end{aligned}$$

With the choice of $\alpha_{21} = 0$, $\alpha_{12} = 0$ *and* $\alpha_{11} = 1$, *the condition of symmetry is satisfied, and the gradient is:*

$$\dot{V}(\mathbf{x}) = \mu x_2^2 - \mu x_1^2 x_2^2 = \mu (1 - x_1^2) x_2^2$$

By integrating $\dot{V}(\mathbf{x})$, *a possible Lyapunov function is obtained:*

$$V(\mathbf{x}) = \int_0^{x_1} x_1\, dx_1 + \int_0^{x_2} x_2\, dx_2 = \frac{x_1^2}{2} + \frac{x_2^2}{2}$$

A possible Lyapunov function is positive definite, and its gradient $\dot{V}(\mathbf{x})$ will be for $\mu > 0$ a positive definite scalar function along any of the solutions of equation (2.8), if $x_1^2(t) < 1$ is valid. The conclusion is that the equilibrium state (origin) is unstable in the Lyapunov sense. However this doesn't mean an unbounded rise of $\|\mathbf{x}(t)\|$; in Example 2.1 it is shown that $\mathbf{x}(t)$ will approach to a limit cycle, i.e. $\|\mathbf{x}(t)\|$ will be limited.

Stability of Time-Varying Systems

In the preceding subsection the stability of time-invariant systems was discussed with the aid of Lyapunov stability theory. In practice time-varying systems are often met. Examples could be mobile objects like rockets, airplanes, ships, underwater vehicles, etc. In such cases the parameters which determine their dynamic behavior are changing with time under various circumstances, for example change of mass because of the fuel consumption (rocket), change of the velocity (airplane, ship), height of the flight (rocket, airplane), depth of the dive (underwater vehicle), etc. Therefore, the stability analysis of time-varying systems is very important and will be discussed further.

Stability Concepts of Time-Varying Systems

The stability concepts of time-varying systems are equivalent to those of time-invariant systems, which were dealt with in Section 2.1. As time-varying systems are dependent upon the initial moment t_0, the definitions of stability must explicitly contain t_0 — this is true for some definitions in Subsection 2.2.1. For time-varying systems the concept of uniformity (Definition 2.3) is of exceptional importance. It eliminates all the systems which are "more and more unstable" as t_0 increases. Definitions of stability in the Lyapunov sense (Definition 2.1), of asymptotic stability in the Lyapunov sense (Definition 2.2), of uniform stability (Definition 2.3) and of exponential stability (Definition 2.4) are valid both for time-invariant, and for time-varying systems, as t_0 is explicitly present in them.

Lyapunov Direct Method for Time-Varying Systems

The basic idea of the Lyapunov direct method is applicable for time-varying systems, too — the difference is that in such systems the Lyapunov function is also time-varying. This brings additional complications when a possible Lyapunov function is sought. However, the greatest drawback in contrast to the stability analysis of time-invariant systems is that La Salle principle of invariance (Theorem 2.3) is not valid here. This drawback circumvents the hypothesis of Barbalat (1959).

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Time-Varying Positive Definite and Decreasing Functions

For the stability analysis of the equilibrium state of a time-varying system by the Lyapunov direct method, instead of a scalar function $V(\mathbf{x})$ it is necessary to use time-varying scalar function $V(t,mvx)$. It is important to define it as positive definite and as decreasing.

DEFINITION 2.9 (LOCALLY POSITIVE DEFINITENESS OF FUNCTION)
A scalar time-varying function $V(t,\mathbf{x})$ is locally positive definite if $V(t,\mathbf{0}) = 0$, and if the time-invariant positive definite scalar function $V_0(\mathbf{x})$ exists, such that:

$$\forall t \geq t_0;\ \ V(t,\mathbf{x}) \geq V_0(\mathbf{x}) \tag{2.47}$$

In other words, the time-varying scalar function $V(t,\mathbf{x})$ is locally positive definite if it "dominates" the time-invariant locally positive definite scalar function $V_0(\mathbf{x})$. Global positive definiteness of the scalar functions generalizes the domain of consideration to the whole state space.

DEFINITION 2.10 (DECREASING TIME-VARYING SCALAR FUNCTION)
Scalar function $V(t,\mathbf{x})$ is decreasing if $V(t,\mathbf{0}) = 0$, and if there exists a time-invariant positive definite scalar function $V_1(\mathbf{x})$, such that:

$$\forall t \geq t_0;\ \ V(t,\mathbf{x}) \leq V_1(\mathbf{x}) \tag{2.48}$$

Therefore the time-varying scalar functions $V(t,\mathbf{x})$ is decreasing if it is "dominated" by a time-invariant positive definite scalar function $V_1(\mathbf{x})$.

EXAMPLE 2.12
A time-varying positive definite scalar function is given by:

$$V(t,\mathbf{x}) = (1+\sin^2 t)(x_1^2+x_2^2) \tag{2.49}$$

$V(t,\mathbf{x})$ is positive definite scalar function since it dominates the time-invariant positive definite scalar function $V_0(\mathbf{x}) = x_1^2+x_2^2$. At the same time, $V(t,\mathbf{x})$ is also a decreasing scalar function, since it is dominated by the positive definite scalar function $V_1(\mathbf{x}) = 2(x_1^2+x_2^2)$.

Stability analysis of time-varying systems considers the stability of the equilibrium states of the forced system which tracks a time-varying reference trajectory or of the systems which are inherently time-varying. It is well known that the problem of the stability of the equilibrium states of a forced system is considered through the dynamics of the system's error, and not through the state of the system. Even if the system is time-invariant, tracking the time-varying trajectory implies that the equivalent system will be time-varying. Stability analysis of the equilibrium states of the time-variant systems in the Lyapunov sense is based on the following theorems:

THEOREM 2.4
(LOCAL STABILITY IN THE LYAPUNOV SENSE—TIME-VARYING SYSTEM)
Stability: *If in the vicinity of the equilibrium state ($\mathbf{x}_e = \mathbf{0}$) a scalar function $V(t,\mathbf{x})$ with continuous partial derivatives exists and if the following conditions are met:*

1. *$V(t,\mathbf{x})$ is a positive definite scalar function,*
2. *$\dot{V}(t,\mathbf{x})$ is a negative semidefinite scalar function,*

then the equilibrium state $\mathbf{x}_e = \mathbf{0}$ is locally stable in the Lyapunov sense.
Local Uniform Stability: *If besides 1 and 2 the following is true:*

3. *$V(t,\mathbf{x})$ is a decreasing function,*

then the equilibrium state (origin) is locally uniformly stable.
Local Uniform Asymptotic Stability: *If 1 and 3 are valid, and 2 is made more strict, i.e. if*

4. *$\dot{V}(t,\mathbf{x})$ is a negative definite scalar function,*

then the equilibrium state is locally uniformly asymptotically stable.

THEOREM 2.5
(GLOBAL ASYMPTOTIC STABILITY OF A TIME-VARYING SYSTEM)
If there exists an unique continuous scalar function $V(t,\mathbf{x})$, defined in the whole state space, which has continuous first partial derivatives, and for which the following are true:

1. $V(t,\mathbf{0}) = 0;\ \forall t$
2. $V(t,\mathbf{x}) > 0;\ \forall t \neq 0$ — *$V(t,\mathbf{x})$ is positive definite*
3. $\dot{V}(t,\mathbf{x}) < 0$ — *$\dot{V}(t,\mathbf{x})$ is negative definite*
4. $V(t,\mathbf{x}) \leq V_0(\mathbf{x});\ \forall t \geq 0$, *where* $V_0(\mathbf{x}) > 0$ — *$V(t,\mathbf{x})$ is decreasing function*
5. $V(t,\mathbf{x}) \to \infty$ *as* $||x|| \to \infty$ — *$V(t,\mathbf{x})$ is radially unbounded*

then the equilibrium state $\mathbf{x}_e$ which satisfies $\mathbf{f}(t,\mathbf{x}_e) = \mathbf{0}$ is globally asymptotically stable in the Lyapunov sense.

For proof of the theorem see Slotine and Li (1991). As distinguished from the time-invariant systems, stability in the Lyapunov sense for time-varying systems requires an additional condition that $V(t,\mathbf{x})$ is a decreasing function. The reason for this is that the Lyapunov function depends on time, so the first partial derivative of $V(t,\mathbf{x})$ is evaluated:

$$\dot{V}(t,\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial V(t,\mathbf{x})}{\partial x_i} \cdot \dot{x}_i + \frac{\partial V(t,\mathbf{x})}{\partial t} \tag{2.50}$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Conditions 1, 2, 3 and 4 of the Theorem 2.5 are necessary for the asymptotic stability of the equilibrium state of a time-varying system, while condition 5 is necessary for a global view. In engineering practice it is not easy to satisfy these conditions. The problem can be avoided by taking the Popov hypothesis (Popov, 1973) which is based on the researsh results of Romanian mathematician Barbalat (Barbalat, 1959). This hypothesis guarantees only the convergence of the state trajectory of the system towards the origin. Alas, the convergence doesn't imply stability, as it may happen that the trajectories first move away from the origin before they approach the origin. Therefore, the origin is not stable in the Lyapunov sense although the state is converging. However, the hypothesis is often applied in engineering practice.

The definitions of the positive definiteness and of the decreasing function, which are of paramount importance for the stability analysis of a time-varying system, can be given indirectly through the so-called class-K functions (definitions of the classes of nonlinear functions are given in the Section 1.8). Here will be given some basic definitions.

DEFINITION 2.11 (CLASS-K FUNCTION)
Continuous function $\alpha : \Re_+ \to \Re_+$ *is of class K if the following is valid:*

1. $\alpha(0) = 0$,
2. $\alpha(q) > 0;\ \forall q > 0$,
3. α *does not decrease.*

A relation exists between class-K and positive definite functions and/or the decreasing functions that is given by following hypothesis:

LEMMA 2.1
(RELATION OF CLASS-K FUNCTION WITH POSITIVE DEFINITE SCALAR LYAPUNOV FUNCTION)
Positive Definite Function: *The function* $V(t,\mathbf{x})$ *is locally (or globally) positive definite if and only if there exists a function* α *of the class K so that* $V(t,\mathbf{0}) = 0$ *is valid, and also:*

$$V(t,\mathbf{x}) \geq \alpha(\|\mathbf{x}\|) \tag{2.51}$$

This equation is valid $\forall t \geq 0$ *and* $\forall \mathbf{x} \in B_\varepsilon$ *(or the whole state space).*
Decreasing Function: *Function* $V(t,\mathbf{x})$ *is locally (or globally) decreasing if and only if there exists function* β *of class K so that* $V(t,\mathbf{0}) = 0$ *is valid, and also:*

$$V(t,\mathbf{x}) \leq \beta(\|\mathbf{x}\|) \tag{2.52}$$

This equation is valid $\forall t \geq 0$ *and* $\forall \mathbf{x} \in B_\varepsilon$ *(or the whole state space).*

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

By making use of Lemma 2.1 the Theorems 2.4 and 2.5 can be expressed as follows:

THEOREM 2.6 (STABILITY IN LYAPUNOV SENSE — TIME-VARYING SYSTEM)
Local Stability: *It is supposed that in the vicinity of the equilibrium state* ($\mathbf{x}_e = \mathbf{0}$) *exists a scalar function* $V(t,\mathbf{x})$ *with continuous first derivatives, and a scalar function* α *of the class K, so that with* $\forall \mathbf{x} \neq \mathbf{0}$*, we have:*

1. $V(t,\mathbf{x}) \geq \alpha(\|\mathbf{x}\|) > 0$

2a. $\dot{V}(t,\mathbf{x}) \leq \mathbf{0}$

Then the equilibrium state (origin) is locally stable in the Lyapunov sense.
Local Uniform Stability: *Besides, if there exists a scalar function* β *of class K, so that:*

3. $V(t,\mathbf{x}) \leq \beta(\|x\|)$

then the equilibrium state ($\mathbf{x}_e = \mathbf{0}$) *is locally uniformly stable.*
Local Uniform Asymptotic Stability: *If the conditions 1 and 3 are met, and 2a is made more strict, so that instead 2a comes:*

2b. $\dot{V}(t,\mathbf{x}) \leq \gamma(\|\mathbf{x}\|)$

where γ *is another scalar function of the class K, then the equilibrium state* ($\mathbf{x}_e = \mathbf{0}$) *is locally uniformly asymptotically stable.*
Global Uniform Asymptotic Stability: *If the conditions 1, 2b and 3 are valid in the whole state space, and if:*

$$\lim_{x \to \infty} \alpha(\|x\|) \to \infty$$

then the equilibrium state $\mathbf{x}_e = \mathbf{0}$ *is globally uniformly asymptotically stable.*[13]

For proof see Vidyasagar (1993).

Local stability in the Lyapunov sense can be interpreted in yet another way. If it can be shown that for a given $\varepsilon > 0$ there exists $\delta > 0$, such that (2.9) is satisfied, because of the conditions 1 and 2a of Theorem 2.6 follows:

$$\alpha(\|\mathbf{x}\|) \leq V[t,\mathbf{x}(t)] \leq V[t_0,\mathbf{x}(t_0)]\,; \forall t \geq t_0 \tag{2.53}$$

As $V(t,\mathbf{x})$ is a continuous scalar function in regard to $\mathbf{x}(t)$, and also $V(t_0,\mathbf{0}) = 0$, a $\delta > 0$ can be found that satisfies the expression:

$$\|\mathbf{x}(t_0)\| < \delta \Rightarrow V[t_0,\mathbf{x}(t_0)] < \alpha(\varepsilon)$$

[13] Also known as La Salle-Yoshizawa theorem (La Salle and Lefschetz, 1961; Yoshizawa, 1968).

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

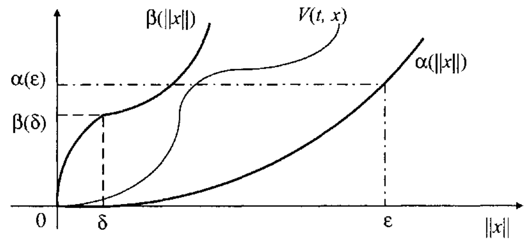


Figure 2.14: Positive definite and decreasing function $V(t,\mathbf{x})$.

Therefore if $\|\mathbf{x}(t_0)\| < \delta$, then $\alpha(\|\mathbf{x}(t)\|) < \alpha(\varepsilon)$ and it follows that $\|\mathbf{x}(t)\| < \varepsilon;\ \forall t \geq t_0$.

Local uniform stability as well as local uniform asymptotic stability may be interpreted according to Theorem 2.6 as follows: from the conditions 1 and 3 of Theorem 2.6 comes:

$$\alpha(\|\mathbf{x}(t)\|) \ \leq\ V[t,\mathbf{x}(t)] \ \leq\ \beta(\|\mathbf{x}(t)\|)$$

For any $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ so that $\beta(\delta) < \alpha(\varepsilon)$, as in Fig. 2.14. If the initial condition is chosen so that $\|\mathbf{x}(t_0)\| < \delta$, then:

$$\alpha(\varepsilon) \ >\ \beta(\delta) \ \geq\ V[t_0,\mathbf{x}(t_0)] \ >\ V[t,\mathbf{x}(t)] \ \geq\ \alpha(\|\mathbf{x}(t)\|)$$

It follows:

$$\forall t \geq t_0;\ \|\mathbf{x}(t_0)\| < \delta \tag{2.54}$$

Local uniform stability is understood, as δ is independent of t_0. On the other hand, the basic idea for local uniform asymptotic stability is that if $\mathbf{x}(t)$ doesn't converge towards the origin, then a positive number a exists, so that:

$$-\dot{V}[t,\mathbf{x}(t)] \geq a > 0 \tag{2.55}$$

The above expression implies:

$$V[t,\mathbf{x}(t)] - V[t_0,\mathbf{x}(t_0)] = \int_{t_0}^{t} \dot{V}[\tau,\mathbf{x}(\tau)]\,d\tau \tag{2.56}$$

and the result is:

$$0 \ \leq\ V[t,\mathbf{x}(t)] \leq V[t_0,\mathbf{x}(t_0)] - (t-t_0)a \tag{2.57}$$

which leads to a contradiction when t is large. Thus the proof of local uniform asymptotic stability is possible with the following arguments. Let $\|\mathbf{x}(t_0)\| < \delta$, where δ is determined as before. Further, μ is any positive constant:

$$0 < \mu < \|\mathbf{x}(t_0)\| \tag{2.58}$$

Another positive constant $r(\mu)$ can be found, so that $\beta(r) < \alpha(\mu)$. Now $R = \gamma(r)$ is defined, and the expression is written as:

$$T = T(\mu, \delta) = \frac{\beta}{R}$$

If $\|\mathbf{x}(t_0)\| > \mu$ for every t in the interval $t_0 \le t \le t_1 \equiv t_0 + T$, we can write:

$$0 < \alpha(\mu) \le V[t_1, \mathbf{x}(t_1)] \le V[t_0, \mathbf{x}(t_0)] - \int_{t_0}^{t_1} \gamma(\|\mathbf{x}(\tau)\|)d\tau$$

$$\le V[t_0, \mathbf{x}(t_0)] - \int_{t_0}^{t_1} \gamma(r)d\tau \le$$

$$V[t_0, \mathbf{x}(t_0)] - (t_1 - t_0)R \le \beta(\delta) - TR = 0$$

It is obvious that here is a contradiction, so there must exist $t_2 \in [t_0, t_1]$ such that $\|\mathbf{x}(t_2)\| \le r$. Therefore, for $t \ge t_2$ is:

$$\alpha(\|\mathbf{x}(t)\|) \le V[t, \mathbf{x}(t)] \le V[t_2, \mathbf{x}(t_2)] \le \beta(r) < \alpha(\mu)$$

and the result is:

$$\|x(t)\| < \mu; \ \forall t \ge t_0 + T \ge t_2$$

which is a proof of local uniform asymptotic stability. For global uniform asymptotic stability it is enough that δ increases arbitrarily. Namely, as α is a radially unbounded function, it is possible to find ε for which is valid $\beta(\delta) < \alpha(\varepsilon)$ for any δ.

EXAMPLE 2.13 (GLOBAL ASYMPTOTIC STABILITY)
Analysis is to be made of the stability of the equilibrium state at the origin ($\mathbf{x}_e = \mathbf{0}$) for the system described by:

$$\dot{x}_1(t) = -x_1(t) - e^{-2t}x_2(t)$$
$$\dot{x}_2(t) = x_1(t) - x_2(t)$$

A possible Lyapunov function in this example may be:

$$V(t, \mathbf{x}) = x_1^2 + (1 + e^{-2t})x_2^2$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

This is a positive definite scalar function which dominates the time-invariant positive definite scalar function $V_0(\mathbf{x}) = x_1^2 + x_2^2$. At the same time it is also a decreasing function as it is dominated by the time-invariant positive definite scalar function $V_1(\mathbf{x}) = x_1^2 + 2x_2^2$. The gradient of the time-variant possible Lyapunov function is:

$$\dot{V}[t,\mathbf{x}(t)] = -2\left[x_1^2 - x_1x_2 + x_2^2(1+2e^{-2t})\right]$$

where $\dot{V}(t,\mathbf{x})$ is a negative definite function, as can be seen from:

$$\dot{V}[t,\mathbf{x}(t)] \leq -2(x_1^2 - x_1x_2 + x_2^2) = -(x_1 - x_2)^2 - x_1^2 - x_2^2$$

The equilibrium state will be globally asymptoticaly stable.

Stability Analysis Using Barbalat's Lemma

Stability analysis of the equilibrium state of a time-varying system is more difficult than that of a time-invariant system, as it is more difficult to find a corresponding Lyapunov function. The conditions of Theorem 2.5 are not simple to satisfy so in engineering practice Barbalat's lemma is therefore much more often used. Barbalat's lemma treats the mathematical conditions for the asymptotical properties of the function and its derivatives. With proper application of this lemma, many problems in the stability analysis of time-varying systems can be avoided.

Asymptotic properties of the function and its derivatives. For a differentiable function $f(t)$, the following facts should be remembered:

- $\dot{f} \longrightarrow 0$ does not imply convergence of the function, neither does it imply that $f(t)$ is bounded when $t \longrightarrow \infty$. Geometric interpretation shows that decreasing the $\dot{f}(t)$ means only that the slope of the tangent to the function $f(t)$ becomes smaller, but it doesn't mean the convergence of the function. As an example, the function $f(t) = \sin(\log t)$ is taken. While $\dot{f}(t) = \frac{\cos(\log t)}{t} \longrightarrow 0$ as $t \longrightarrow \infty$, the function $f(t)$ oscillates more and more slowly as t increases. Another example $f(t) = \sqrt{t}\sin(\log t)$ indicates that the function is unbounded.

- The convergence of $f(t)$ doesn't imply that $\dot{f} \longrightarrow 0$; the fact that $f(t)$ converges when $t \longrightarrow \infty$ doesn't mean that $\dot{f} \longrightarrow 0$. An additional example is $f(t) = e^{-t}\sin(e^{2t})$ which tends to zero, while $\dot{f}(t)$ is unbounded for $t \longrightarrow \infty$.

- If $f(t)$ is bounded from below and decreasing ($\dot{f}(t) \leq 0$), then it converges to some boundary. The question arises if the function tends to a final boundary, what additional requirements assure that the derivatives of the function tend to zero? Barbalat's lemma points to the fact that the derivatives of the function must have smoothness.

LEMMA 2.2 (BARBALAT'S LEMMA)
If a function $f(t)$ has a final boundary when $t \longrightarrow \infty$, and if $\dot{f}(t) = \frac{df(t)}{dt}$ is a differentiable and uniformly continuous function, then $\dot{f}(t) \longrightarrow 0$ when $t \longrightarrow \infty$.

Uniform continuity of a function is often difficult to determine directly. In order that the differentiable function be uniformly continuous, a sufficient condition is that the function differential be bounded. A corollary to Barbalat's lemma is if the differentiable function $f(t)$ has a final boundary when $t \longrightarrow \infty$ and for which a bounded second derivative $\ddot{f}(t)$ exists, then $\dot{f}(t) \longrightarrow 0$ when $t \longrightarrow \infty$.

EXAMPLE 2.14 (UNIFORM CONTINUITY OF A LINEAR SYSTEM)
An asymptotically stable linear system with a bounded excitation $\mathbf{u}(t)$ will have a uniformly continuous response $\mathbf{y}(t)$. Namely, if the linear system is described by the mathematical model with state variables:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$

then, because of the bounded excitation and of asymptotic stability of the system, the state will be bounded, too. The boundedness of the state implies also that $\dot{\mathbf{x}}(t)$ is bounded. From the second (algebraic) equation it follows that $\dot{\mathbf{y}}(t) = C\dot{\mathbf{x}}(t)$ is also bounded. Therefore the system response $\mathbf{y}(t)$ is uniformly continuous.

The application of Barbalat's lemma in the analysis is used by the following lemma on the convergence (Slotine and Li, 1991):

LEMMA 2.3 (LEMMA ON CONVERGENCE — "LYAPUNOV-LIKE LEMMA")
Supposing that a scalar function $V(t,\mathbf{x})$ exists which satisfies:

1. *$V(t,\mathbf{x})$ is bounded below,*
2. *$\dot{V}(t,\mathbf{x})$ is negative semidefinite,*
3. *$\dot{V}(t,\mathbf{x})$ is uniformly continuous with time,*

then $\dot{V}(t,\mathbf{x}) \longrightarrow 0$ when $t \longrightarrow \infty$.

Boundedness below: A sufficient condition for the scalar function $V(t,\mathbf{x})$ to be bounded below is that it is positive semidefinite, i.e. $V(t,\mathbf{x}) \geq 0$; $\forall t > t_0$.
Uniform continuity: A sufficient condition for the differentiable function $\dot{V}(t,\mathbf{x})$ to be uniformly continuous is that $\ddot{V}(t,\mathbf{x})$ is bounded for $\forall t \geq t_0$.

EXAMPLE 2.15
(APPLICATION OF LEMMA ON CONVERGENCE—LYAPUNOV-LIKE ANALYSIS)

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

The error dynamics of the closed-loop adaptive control system for the first order process with one unknown parameter are:

$$\begin{aligned}\dot{e}(t) &= -e(t) + \Theta(t)w(t)\\ \dot{\Theta}(t) &= -e(t)w(t)\end{aligned}$$

where $e(t)$ is the tracking error, $\Theta(t)$ is the error in the parameter and $w(t)$ is a bounded continuous function.

Both the error of tracking and error in parameter can be conceived as the state variables of the closed-loop dynamics. The analysis of the asymptotic properties of the system can be carried out with the lemma of convergence (Lemma 2.3). If a possible Lyapunov function is chosen:

$$V(t,e,\Theta) = e^2 + \Theta^2$$

its gradient will be:

$$\dot{V}(t,e,\Theta) = 2e(-e+\Theta w) + 2\Theta(-ew) = -2e^2 \leq 0 \tag{2.59}$$

This implies that $V(t,\mathbf{x}) \leq V(0)$, therefore $e(t)$ and $\Theta(t)$ are bounded. As the dynamics of this system are time-varying, the lemma on convergence must be applied (Lemma 2.3). In order to introduce Barbalat's lemma, we must test the uniform continuity of the derivative of a possible Lyapunov function (condition 3) of the Lemma 2.3). The derivative of $\dot{V}(t,\mathbf{x})$ is:

$$\ddot{V}(t,\mathbf{x}) = -4e(-e+\Theta w) \tag{2.60}$$

The second derivative of a possible Lyapunov function is a bounded function, as $w(t)$ is bounded (the initial supposition), while $e(t)$ and $\Theta(t)$ have previously been shown to be bounded. Therefore, $\dot{V}(t,\mathbf{x})$ is a uniformly continuous function. The application of Barbalat's lemma implies that $e(t) \to 0$ when $t \to \infty$. It may be observed that although $e(t)$ converges to zero, the system is not asymptotically stable—for $\Theta(t)$ is only guaranteed to be bounded.

EXAMPLE 2.16 (STABILITY OF TIME-VARYING NONLINEAR SYSTEM)
Conduct the stability analysis of the equilibrium state of the system:

$$\ddot{x}(t) + a\dot{x}(t) + g(t,x)x(t) = 0$$

or:

$$\begin{aligned}\dot{x}_1(t) &= x_2(t)\\ \dot{x}_2(t) &= -ax_2(t) - g(t,x_1)x_1(t)\end{aligned} \tag{2.61}$$

If it is supposed that:

$$\nabla_x V = \begin{bmatrix} \alpha_{11}x_1 + \alpha_{12}x_2 \\ \alpha_{21}x_1 + x_2 \end{bmatrix}$$

then it follows that:

$$\dot{V}(t,\mathbf{x}) = (\nabla_{\mathbf{x}} V)^T \dot{\mathbf{x}}(t) + \frac{\partial V(t,\mathbf{x})}{\partial t}$$

or:

$$\begin{aligned} V(t,\mathbf{x}) &= \alpha_{11}x_1x_2 + \alpha_{12}x_2^2 - a\alpha_{21}x_1x_2 \\ &\quad - g(t,x_1)x_1^2\alpha_{21} - ax_2^2 - g(t,x_1)x_1x_2 + \frac{\partial V(t,\mathbf{x})}{\partial t} \end{aligned}$$

If the terms with x_1x_2 should be eliminated, then the expression $\alpha_{11} = a\alpha_{21} + g(t,x_1)$ must be used. The conditions of symmetry will be satisfied if $\alpha_{12} = \alpha_{21} = const$. Then:

$$V(t,\mathbf{x}) = \int_0^{x_1} [a\alpha_{21} + g(t,x_1) + \alpha_{12}x_2]x_1dx_1|_{x_2=0} + \int_0^{x_2} [\alpha_{21}x_1 + x_2)dx_2|_{x_1=const.}$$

and after integrating:

$$V(t,\mathbf{x}) = a\alpha_{21}\frac{x_1^2}{2} + \alpha_{21}x_1x_2 + \frac{x_2^2}{2} + \int_0^{x_1} g(t,x_1)x_1dx_1 \tag{2.62}$$

If $\int_0^{x_1} g(t,x_1)x_1dx_1 > 0$ for all $x_1(t)$ and t, from (2.62) it follows:

$$\begin{aligned} V(t,\mathbf{x}) &> a\alpha_{21}\frac{x_1^2}{2} + \alpha_{21}x_1x_2 + \frac{x_2^2}{2} \pm \frac{\alpha_{21}^2x_1^2}{2} \\ &= \frac{1}{2}(x_2 + \alpha_{21}x_1)^2 + \frac{1}{2}(a\alpha_{21} - \alpha_{21}^2)x_1^2 \end{aligned}$$

$V(t,\mathbf{x})$ will be a positive definite scalar function if:

$$a > 0$$
$$a\alpha_{21} > \alpha_{21}^2$$

These conditions can be assured with the choice $\alpha_{21} = a - \varepsilon$, where ε is a small positive number. The possible Lyapunov function derivation is:

$$\begin{aligned} \dot{V}(t,\mathbf{x}) &= a\alpha_{21}x_1\dot{x}_1 + \alpha_{21}\dot{x}_1x_2 + \alpha_{21}x_1\dot{x}_2 + x_2\dot{x}_2 \\ &\quad + g(t,x_1)x_1\dot{x}_1 + \int_0^{x_1} \frac{\partial g(t,x_1)}{\partial t}x_1dx_1 \end{aligned} \tag{2.63}$$

Substitution of (2.61) and of $\alpha_{21} = a - \varepsilon$ *in the expression (2.63) gives:*

$$\dot{V}(t,\mathbf{x}) = -\varepsilon x_2^2 - (a-\varepsilon)g(t,x_1)x_1^2 + \int_0^{x_1} \frac{\partial g(t,x_1)}{\partial t} x_1 dx_1 \qquad (2.64)$$

Negative semidefiniteness of (2.64) is assured if $g(t,x_1) > 0$ *for all* x_1 *and* t *and if:*

$$\int_0^{x_1} \frac{\partial g(t,x_1)}{\partial t} x_1 dx_1 < a g(t,x_1) x_1^2 \qquad (2.65)$$

The expression (2.65) is valid for:

$$\max_{x_1,t} \left[\frac{\partial g(t,x_1)}{\partial t} \right] < 2ag(t,x_1);\ \forall x_1, t \qquad (2.66)$$

and if $g(t,x_1)$ *is bounded for every* x_1 *and* t. *In this case* $V(t,\mathbf{x})$ *can be bounded as is required by condition 4 of Theorem 2.5. In such a way, the system will be globally asymptotically stable.*

From the above discussion, two essential differences in the stability analysis of time-varying and time-invariant nonlinear systems can be observed. The first difference is that with the time-varying system, $V(t,\mathbf{x})$ has to be bounded below, instead of being only a positive definite function, as was the case with the time-invariant system. The second difference is that $\dot{V}(t,\mathbf{x})$ must be tested for uniform continuity, in addition to being a negative definite or semidefinite function. Usually $\ddot{V}(t,\mathbf{x})$ is used for this purpose, while testing its boundedness. As was the case with time-invariant systems, with time-varying systems the primary difficulty remains in a proper choice of a possible Lyapunov function.

Positive Linear System

The analysis of nonlinear systems often starts with the decomposition of system to linear and nonlinear subsystems. If the transfer function (or matrix) of a linear subsystem is *positive real*, then it has important properties that facilitate the choice of a possible Lyapunov function for the whole system. A linear system with a positive real transfer function is called a *positive linear system*. These systems play therefore an important role in the analysis and the synthesis of many nonlinear control systems, e.g. the adaptive control.

DEFINITION 2.12
(POSITIVE REAL AND STRICTLY POSITIVE REAL TRANSFER FUNCTION)
A transfer function given by:

$$G(s) = \frac{b_m s^m + b_{m-1}s^{m-1} + \ldots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0};\ \text{with } n \geq m$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

and real coefficients a_i, $(i = 0, 1, \ldots, n-1)$ *and* b_j $(j = 0, 1, \ldots, m)$ *is positive real if:*

$$Re\{G(s)\} \geq 0 \textit{ for all } Re\{s\} \geq 0 \tag{2.67}$$

and it is strictly positive real if:

$$G(s - \varepsilon) \textit{ is positive real for an } \varepsilon > 0$$

The condition of positivity realness (2.67) requires in fact that $G(s)$ has always positive (or zero) real parts when s has positive (or zero) real parts. The linear scalar system is *passive* if and only if it is positive real. The importance of the passivity of a linear subsystem lies in the fact that parallel and feedback combinations of positive real transfer functions does not jeopardize the stability of the nonlinear system. Geometrical interpretation of positivity realness means that rational function $G(s)$ with the complex variable $s = \sigma \pm j\omega$ maps every point from the closed right half-plane of the complex s plane (including the imaginary axis) into the closed right half-plane of the $G(s)$ plane. For transfer functions of higher order it is not easy with the above definition to determine the condition of positivity realness, as it includes testing of the positivity condition over the entire right half-plane. For this purpose the next theorem simplifies the testing of positivity realness.

THEOREM 2.7 (STRICTLY POSITIVE REAL TRANSFER FUNCTION)
Transfer function $G(s)$ is strictly positive real (SPR) if and only if:

1. *$G(s)$ is a strictly stable transfer function,*
2. *The real part of $G(s)$ is strictly positive along the $j\omega$ axis, i.e.* $\forall \omega > 0, Re\{G(j\omega)\} > 0$.

The above condition is also the condition of passivity if the sign '> 0' is replaced by '≥ 0'. According to Theorem 2.7, necessary conditions for the transfer function to be SPR are:

- $G(s)$ is strictly stable (denominator is a Hurwitz polynomial),
- For the frequency characteristic $G(j\omega)$, the polar (Nyquist) plot lies completely in the right half-plane—that means the phase shift always is less than $90°$,
- $G(s)$ has a relative degree $n - m = 0$ or $n - m = 1$,
- $G(s)$ is strictly phase-minimum (i.e. all its zeros are strictly in the left half-plane of the complex plane).

If the transfer function of the linear subsystem of a nonlinear system is SPR, then the linear part of the system has an important property, expressed by the famous Kalman-Jakubovich lemma, also known as the positive real lemma.

LEMMA 2.4 (KALMAN-JAKUBOVICH LEMMA)
Consider a completely controllable linear time-invariant system:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$$
$$y(t) = \mathbf{c}^T\mathbf{x}(t)$$

The transfer function $G(s) = \mathbf{c}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}$ *is SPR if and only if there exist positive definite matrices P and Q such that*

$$\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}$$
$$\mathbf{P}\mathbf{b} = \mathbf{c}$$

The significance of this lemma is that it is applicable to transfer functions that contain a pure integrator, a situation which quite often occurs in adaptive systems. This lemma requires that the system be asymptotically stable and completely controllable. A modified version of the Kalman-Jakubovich lemma which doesn't require complete controllability follows.

LEMMA 2.5 (MEYER-KALMAN-JAKUBOVICH LEMMA)
Let the scalar γ *be nonnegative, vectors* **b** *and* **c** *are of corresponding dimensions, matrix* **A** *of the linear system is asymptotically stable, and a symmetric positive definite matrix* **L** *with dimensions* $n \times n$ *exists. If the transfer function:*

$$H(s) = \frac{\gamma}{2} + \mathbf{c}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}$$

is strictly positive real, then there exists a scalar $\varepsilon > 0$*, vector* **q***, and a symmetric positive definite matrix* **P** *such that:*

$$\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{q}\mathbf{q}^T - \varepsilon\mathbf{L}$$
$$\mathbf{P}\mathbf{b} = \mathbf{c} + \sqrt{\gamma}\,\mathbf{q}$$

This lemma differs from the Kalman-Yakubovich lemma in two aspects. First, the output equation is now:

$$y(t) = \mathbf{c}^T\mathbf{x}(t) + \frac{\gamma}{2}u(t)$$

and second, the system is not required to be controllable, but only "stabilizable" or the uncontrollable part of the system must be stable. Kalman-Yakubovich lemma is largely applied in the design of adaptive systems, and has to be further elaborated with the design of nonlinear control systems.

Lyapunov Function in Other Applications Besides Stability Analysis

Besides the use of the Lyapunov direct method or the Lyapunov function in stability analysis, it is also useful with automatic control systems for the estimation of a dominant time constant of the system, a fast rejection of disturbances, evaluation of the region of atractiveness, system stabilization, synthesis of adaptive control systems, etc.

Assessment of Dominant Time Constant

The Lyapunov function can be used for estimation of the velocity with which the state of the system returns to the equilibrium after a disturbance. If the origin is the equilibrium state, the velocity of return to the origin, following a disturbance, is evaluated. If the system is globally asymptotically stable in the Lyapunov sense, and if $V(t,\mathbf{x})$ is a Lyapunov function, the quasi constant can be defined:

$$\mu = \min_{x} \left[\frac{-\dot{V}(t,\mathbf{x})}{V(t,\mathbf{x})} \right] \tag{2.68}$$

where minimization is accomplished along all state vector components $\mathbf{x} \neq \mathbf{0}$ in the region where the equilibrium state is asymptotically stable. From (2.68) it follows:

$$\dot{V}(t,\mathbf{x}) \leq \mu V(t,\mathbf{x}) \tag{2.69}$$

By integrating the left and the right sides, we obtain:

$$\int_{t_0}^{t} \mu dt \geq - \int_{V[t_0,\mathbf{x}(t_0)]}^{V[t,\mathbf{x}(t)]} \frac{dV(t,\mathbf{x})}{V(t,\mathbf{x})}$$

and further:

$$\ln \left[\frac{V[t,\mathbf{x}(t)]}{V[t_0,\mathbf{x}(t_0)]} \right] \leq - \int_{t_0}^{t} \mu dt = -\mu(t-t_0) \tag{2.70}$$

or:

$$V[t,\mathbf{x}(t)] \leq V[t_0,\mathbf{x}(t_0)] e^{-\mu(t-t_0)} \tag{2.71}$$

Since for an asymptotically stable equilibrium state $\mathbf{x}(t) \to \mathbf{0}$ and $V[t,\mathbf{x}(t)] \to 0$, then $1/\mu$ may be interpreted as the upper bound for the time constant of the system. When several Lyapunov functions are available, the choice will be for the function which yields the greatest μ or which has the greatest ratio $\frac{-\dot{V}(t,\mathbf{x})}{V(t,\mathbf{x})}$. For a known Lyapunov function, the procedure to find μ (expression (2.71)) is achieved numerically, since the problem is difficult to calculate. For linear systems a simplified procedure can be used. Namely, as the Lyapunov function has the form:

$$V(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x}$$

and its gradient is:

$$\dot{V}(\mathbf{x}) = -\mathbf{x}^T \mathbf{Q} \mathbf{x}$$

where $\mathbf{P}$ is a symmetric positive definite matrix and $\mathbf{Q}$ is a symmetric positive definite matrix for which the Lyapunov equation is valid:

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} = -\mathbf{Q}$$

The coefficient μ can be obtained in the following way:

$$\mu = \min_x \left[\frac{-\mathbf{x}^T\mathbf{Q}\mathbf{x}}{\mathbf{x}^T\mathbf{P}\mathbf{x}} \right] \tag{2.72}$$

At the minimum the following relation must be satisfied:

$$\frac{\partial \mu}{\partial \mathbf{x}} = \frac{-2\mathbf{Q}\mathbf{x}(\mathbf{x}^T\mathbf{P}\mathbf{x}) + (\mathbf{x}^T\mathbf{Q}\mathbf{x})2\mathbf{P}\mathbf{x}}{(\mathbf{x}^T\mathbf{P}\mathbf{x})^2} = 0 \tag{2.73}$$

Since there are many 'points' $\mathbf{x}$ which fulfil the above condition, usually a typical point $\gamma = \frac{\mathbf{x}^T\mathbf{Q}\mathbf{x}}{\mathbf{x}^T\mathbf{P}\mathbf{x}}$ is taken. Then from (2.72) we have:

$$(\mathbf{Q} - \gamma\mathbf{P})\mathbf{x} = \mathbf{0}$$

or:

$$(\mathbf{P}^{-1}\mathbf{Q} - \gamma\mathbf{I})\mathbf{x} = \mathbf{0}$$

which shows that γ must be an eigenvalue of the matrix $\mathbf{P}^{-1}\mathbf{Q}$—actually μ is the minimum eigenvalue of the matrix $\mathbf{P}^{-1}\mathbf{Q}$.

EXAMPLE 2.17
(EVALUATION OF DOMINANT TIME CONSTANT OF LINEAR SYSTEM)
An unforced system (see Example 2.9) is described by:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t); \qquad \mathbf{A} = \begin{bmatrix} -1 & -2 \\ 1 & -4 \end{bmatrix}$$

The matrices $\mathbf{P}$ *and* $\mathbf{Q}$ *follow from Example 2.9.*

$$\mathbf{P} = \begin{bmatrix} \frac{23}{60} & \frac{-7}{60} \\ \frac{-7}{60} & \frac{11}{60} \end{bmatrix}; \qquad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

As $\mathbf{Q} = \mathbf{I}$, μ *will be the smallest eigenvalue of the matrix* $\mathbf{P}^{-1}$, *as* $\mathbf{P}^{-1}\mathbf{Q} = \mathbf{P}^{-1}$. *To find the eigenvalue of the matrix* $\mathbf{P}^{-1}$, *the solution is equivalent to that of* $|\mathbf{I} - \mathbf{P}\lambda| = 0$, *respectively:*

$$\begin{vmatrix} 1 - \frac{23}{60}\lambda & \frac{7}{60}\lambda \\ \frac{7}{60}\lambda & 1 - \frac{11}{60}\lambda \end{vmatrix} = 0$$

which yields $\lambda_1 = 2.288$, $\lambda_2 = 7.71$ *and* $\mu = 2.288$. *The equation (2.71) is now:*

$$V(\mathbf{x}) \leq V(\mathbf{x}_0)e^{-2.288(t-t_0)}$$

If the upper time bound has to be determined in order to bring the state of the system from the initial condition $x(0) = [1\,0]^T$ *to the vicinity of the equilibrium*

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

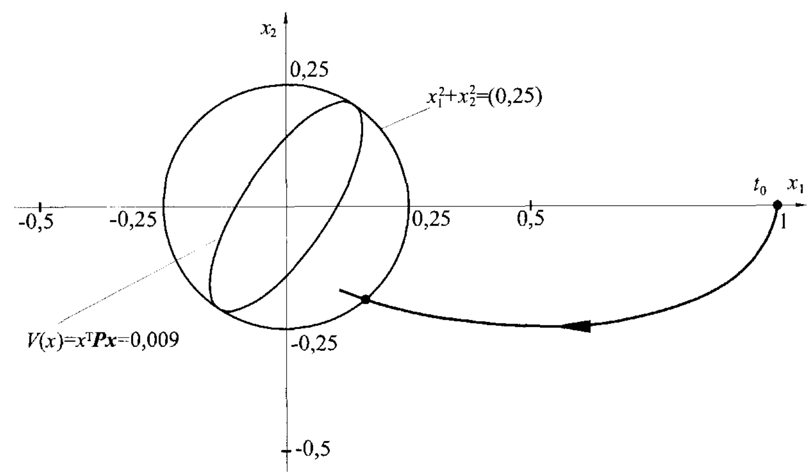


Figure 2.15: State trajectory of a linearized system (Example 2.17).

state defined by the circle with the radius 0.25 *round the origin, then the necessary time can be calculated as follows. First, the maximum value of C has to be calculated so that* Lyapunov plane $V(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x} = C$ *lies completely inside or on the circle* $x_1^2 + x_2^2 = 0.25^2$. *The value of C which satisfies this condition is* 0.009. *The situation is illustrated in Fig. 2.15. If the settling time* $t = t_s$ *is necessary for the system to approach the vicinity of the origin (equilibrium state), defined by the equation of the circle with the radius* 0.25, *then (2.70) yields:*

$$t_s \leq -\frac{1}{\mu}\left[\ln\frac{C}{V(\mathbf{x}_0)}\right]$$

As:

$$V(\mathbf{x}_0) = \mathbf{x}_0^T\mathbf{P}\mathbf{x}_0 = [1\,0]\begin{bmatrix} \frac{23}{60} & \frac{-7}{60} \\ \frac{-7}{60} & \frac{11}{60} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{23}{60}$$

it follows that:

$$t_s \leq -\frac{1}{2.288}\left[\ln\frac{0.09\cdot 60}{23}\right] = 1.64$$

Fast Rejection of Disturbances on a System

If an automatic control system is to be designed which rejects the influence of disturbances, the procedure is as follows. First the Lyapunov function is defined—

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

appropriate for the open-loop control system. When discussing the appropriate input signal $\mathbf{u}(t)$, the Lyapunov function remains unchanged since it doesn't depend on $\dot{\mathbf{x}}(t)$ (or $\mathbf{u}(t)$). As $\dot{V}(t,\mathbf{x})$ depends indirectly on $\mathbf{u}(t)$ through $\dot{\mathbf{x}}(t)$, such a control signal is chosen so that the gradient of $V(t,\mathbf{x})$ is maximum. This assures that $V(t,\mathbf{x})$ is approaching as fast as possible to the origin. An approximately optimal control can be realized in this way.

EXAMPLE 2.18 (APPROXIMATE OPTIMAL CONTROL OF LINEAR SYSTEM)
A linear system is described by:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$$

Assuming the control vector is bounded for every t, we can write:

$$U_{\max} = \mathbf{u}^T(t)\mathbf{u}(t) = [u_1(t)\, u_2(t)] \cdot \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} = u_1^2(t) + u_2^2(t) \leq 1$$

It is necessary to determine the control algorithm as a funciton of the state variables, so that the state of the system returns as fast as possible to the equilibrium state ($\mathbf{x}_e = \mathbf{0}$) after being disturbed from it. To solve this problem the Lyapunov function must be determined as well as a state control vector which gives a higher gradient of the Lyapunov function and simultaneously ensures that $U_{max} \leq 1$ is satisfied.

A Lyapunov function is given with $V(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x}$. The Lyapunov equation is:

$$\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}$$

If $\mathbf{Q} = \mathbf{I}$, and knowing that $\mathbf{P}$ must be a symmetric positive definite matrix in order that the system be asymptotically stable, the Lyapunov equation takes the form:

$$\begin{bmatrix} -3 & -1 \\ 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \cdot \begin{bmatrix} -3 & 2 \\ -1 & -1 \end{bmatrix} = - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

or:

$$\begin{bmatrix} -6p_{11} - 2p_{12} & -4p_{12} - p_{22} + 2p_{11} \\ -4p_{12} - p_{22} + 2p_{11} & 4p_{12} - 2p_{22} \end{bmatrix} = - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

It follows:

$$\begin{bmatrix} -6 & -2 & 0 \\ 0 & 4 & -2 \\ 2 & -4 & -1 \end{bmatrix} \cdot \begin{bmatrix} p_{11} \\ p_{12} \\ p_{22} \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$$

and:

$$\mathbf{P} = \begin{bmatrix} \frac{7}{40} & \frac{-1}{40} \\ \frac{-1}{40} & \frac{18}{40} \end{bmatrix}, \text{ or } |\mathbf{P}| = \frac{125}{1600} > 0$$

The determinants of the left minors (submatrices) of the matrix $\mathbf{P}$ *are* $\det(\mathbf{P}_1) = 7/40 > 0$, *and* $\det(\mathbf{P}_2) = |\mathbf{P}| > 0$, *the matrix* $\mathbf{P}$ *is positive definite and the system is asymptotically stable. With an assumed* $V(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x}$ *its gradient is:*

$$\begin{aligned}\dot{V}(\mathbf{x}) &= \dot{\mathbf{x}}^T\mathbf{P}\mathbf{x} + \mathbf{x}^T\mathbf{P}\dot{\mathbf{x}} = (\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{u})^T\mathbf{P}\mathbf{x} + \mathbf{x}^T\mathbf{P}(\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{u})\\ &= \mathbf{x}^T\mathbf{A}^T\mathbf{P}\mathbf{x} + \mathbf{u}^T\mathbf{B}^T\mathbf{P}\mathbf{x} + \mathbf{x}^T\mathbf{P}\mathbf{A}\mathbf{x} + \mathbf{x}^T\mathbf{P}\mathbf{B}\mathbf{u}\end{aligned}$$

where $\mathbf{x}^T(\mathbf{A}^T\mathbf{P}+\mathbf{P}\mathbf{A})\mathbf{x} = x^T(-\mathbf{I})\mathbf{x} = -\mathbf{x}^T\mathbf{x}$ *and* $\mathbf{x}^T\mathbf{P}\mathbf{B}\mathbf{u} = \mathbf{u}^T\mathbf{B}^T\mathbf{P}\mathbf{x}$. *Thus we obtain:*

$$\dot{V}(\mathbf{x}) = -\mathbf{x}^T\mathbf{x} + 2\mathbf{u}^T\mathbf{B}^T\mathbf{P}\mathbf{x}$$

The control vector must be chosen so as to minimize $\mathbf{u}^T\mathbf{B}^T\mathbf{P}\mathbf{x}$ *with the constraint* $U_{max} \leq 1$, *since the gradient of* $V(\mathbf{x})$ *will be at maximum. It is obvious that the control vector* $\mathbf{u}(t)$ *must be parallel to the vector* $\mathbf{B}^T\mathbf{P}\mathbf{x}$, *but of opposite sign and of maximally allowed amplitude. This can be expressed thus:*

$$\mathbf{u}(t) = -\frac{\mathbf{B}^T\mathbf{P}\mathbf{x}(t)}{\|\mathbf{B}^T\mathbf{P}\mathbf{x}(t)\|}$$

By inserting the matrices $\mathbf{B}$ *and* $\mathbf{P}$ *in the above expression, the desired control vector is obtained:*

$$\mathbf{u}(t) = \frac{\begin{bmatrix} -7x_1(t)+x_2(t) \\ 8x_1(t)-19x_2(t) \end{bmatrix}}{\sqrt{113x_1^2 - 318x_1x_2 + 362x_2^2}}$$

When the region of attraction of the asymptotically stable equilibrium state (origin) is to be evaluated, i.e. when the sets of initial conditions contained in the region of attraction are to be determined, then first the Lyapunov function which satisfies the conditions of asymptotic stability in a region Ω must be found. If $D\{V(t,x) \leq c\}$ is a bounded set and encompassed in Ω, then every trajectory which starts in D will approach the origin when $t \to \infty$, and will remain in the region D. Therefore, D represents the evaluation of the region of attraction.

2.3 Absolute Stability

The first published papers about the absolute stability of nonlinear control systems appeared in 1944. Lur'e and Postnikov (1944) had researched nonlinear systems with a continuous single-valued nonlinear characteristic which passes through the first and third quadrants. M.A. Aizerman (1946) formulated the problem of absolute stability when the nonlinear characteristic is within a sector—in 1947 he established a hypothesis whereby the stability of nonlinear systems may be analyzed with linear procedures. The Romanian mathematican V.M. Popov in 1959

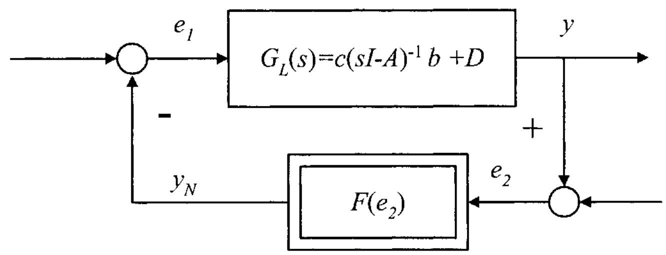


Figure 2.16: Structure of nonlinear systems in the problems of analysis of absolute stability.

proposed a fundamentally new approach to the problems of absolute stability—he established necessary conditions which the amplitude-frequency characteristic of the linear part of the system must fulfil, so that the nonlinear system will be absolutely stable. This frequency approach immediately found the approval of engineers to whom the frequency approach was familiar. The work of Popov was later published in French (Popov, 1960), in Russian (Popov, 1961), and in English (Popov, 1973). This problem is elaborated in detail by Naumov and Tsypkin (1964), Sandberg (1964), Vavilov (1970), etc. The concept of absolute stability means a global asymptotical stability of the equilibrium states in the Lyapunov sense.

The structure of the nonlinear system in problems of absolute stability is given in Fig. 2.16. In the direct branch is the linear time-invariant system, while in the feedback branch is a single-valued nonlinearity (nonlinear element without memory), which means that the feedback performs nonlinear static mapping of the signal e_2 to the signal y_N. The signal r_1 may represent a reference signal, while the signal r_2 may represent an error signal, for example measurement noise. In the case that the transfer function of the linear part is strictly proper ($D = 0$), and $r_1 = r_2 = 0$, the system in Fig. 2.16 can be mathematically described by:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} - \mathbf{b}F(y) \\ y &= \mathbf{c}^T\mathbf{x} \end{aligned} \tag{2.74}$$

where $y = e_2$ and $e_1 = -y_N = -F(y)$. Many automatic control systems can be represented with this structure. If the system acts in a stabilization mode (so-called regulator problem), the structure from Fig. 2.16 can be reduced to the structure in Fig. 2.17. If the nonlinear function $F(y)$ belongs to the sector[14] $[k_1, k_2]$ and if the linear part of the system is stable[15], the question arises: what

[14]See Definition 1.14 in Section 1.8.

[15]Matrix $\mathbf{A}$ is a Hurwitz matrix, respectively a polynomial $A(s)$ is a Hurwitz polynomial.

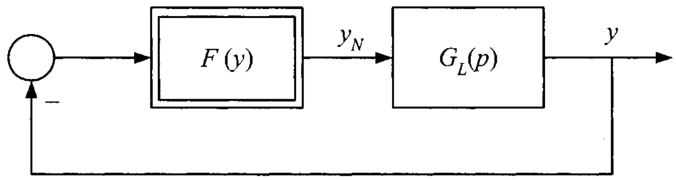


Figure 2.17: Structure of unforced nonlinear system.

additional constraint is necessary for the nonlinear element (static characteristic $F(y)$) in the feedback loop so that a closed-loop system is stable? As the nonlinear characteristic is in the sector $[k_1,k_2]$, that means it is bounded by two straight lines which intersect at the origin—this corresponds to feedback with constant gain—it is normal to suppose that the stability of the nonlinear system will have certain similarities with the stability of the system which is stabilized by the gain in the feedback loop. Contrary to the stability analysis discussed so far, in this situation our interest is not in a specific system, but in the whole family of systems, as $F(y)$ can be any nonlinear function inside the sector $[k_1,k_2]$—see Fig. 1.35a. This is the reason why this is called the problem of absolute stability in the sense that if the system is absolutely stable, it is stable for the whole family of nonlinearities in the sector $[k_1,k_2]$.

M.A. Aizerman (1949) has considered this problem and established the following hypothesis. If the system shown described by (2.76) is globally asymptotically stable, for all linear mappings given by:

$$F(t,y) = ky, \ \forall t,y,k \in [k_1,k_2] \tag{2.75}$$

where the gain of the linear feedback k is inside the interval $[k_1,k_2]$, then the same is true for all time-invariant nonlinear systems with a single-valued static characteristic $F(y)$ inside the sector $[k_1,k_2]$. In other words, if the nonlinear feedback is replaced by a linear proportional feedback and if we obtain a closed-loop system (matrix $[\mathbf{A}-\mathbf{bc}^T k]$), globally asymptotically stable for all values of the linear gain inside $[k_1,k_2]$, then the nonlinear system which possesses a nonlinear feedback is globally asymptotically stable if the static characteristic of the nonlinear element is inside the sector $[k_1,k_2]$. This attractive hypothesis is, alas, not valid in the general case, as Willems (1971) has proved. R.E. Kalman (1957) has proposed a similar hypothesis which assumes that $F(y)$ belongs to the incremental[16] sector $[k_1,k_2]$. The set of nonlinearities in the incremental sector is smaller than the

[16] If the function belongs to incremental sector $[k_1,k_2]$, then it belongs to the sector $[k_1,k_2]$, too. The reverse is not true. The function belongs to the incremental sector if $F(0)=0$, and if $k_1 \leq F(y) \leq k_2, \forall y \in \Re$, is true.

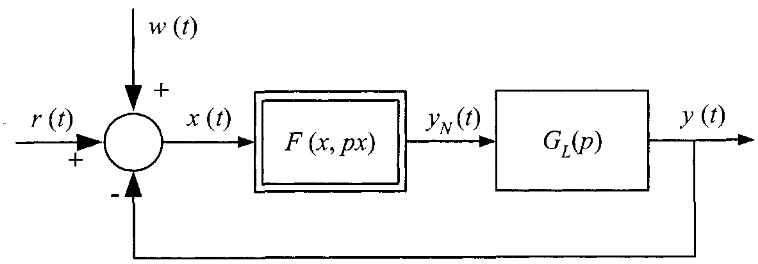


Figure 2.18: Basic structure of a nonlinear system for analysis of absolute stability with external signals $r(t)$ and $w(t)$.

set of nonlinearities treated by Aizerman so the Kalman hypothesis has greater probability to be correct.

Further considerations will be given to nonlinear systems with the structure as in Fig. 2.18. The linear part of the system can be stable, unstable or neutrally stable, with the transfer function:

$$G_L(s) = \frac{B(s)}{A(s)} = \frac{K_L(b_m s^m + b_{m-1}s^{m-1} + \ldots + 1)}{a_n s^n + a_{n-1}s^{n-1} + \ldots + 1}; \; m \leq n-1 \tag{2.76}$$

The nonlinear part can be:

1. Single-valued time-invariant nonlinear element $y_N(t) = F[x(t)]$,
2. Single-valued time-varying nonlinear element $y_N(t) = F[t, x(t)]$,
3. Double valued time-invariant nonlinear element $y_N(t) = F[x(t), \dot{x}(t)]$.

Besides the above-mentioned nonlinear elements, there appear:

4. Linear time-varying element $y_N = k(t)x(t)$,
5. Linear time-invariant element $y_N = kx(t)$.

All these elements must have static characteristic inside the sector[17] $[k_1, k_2]$ for all $t > 0$, where $0 \leq k_1 < k_2 < \infty$.

On the closed-loop nonlinear system act the reference signal $r(t)$ and disturbance signal $w(t)$. The total input signal to the system is therefore $f(t) = r(t) + w(t)$. The dynamics of the closed-loop nonlinear system in Fig. 2.18 can be described by the differential equation:

$$x(t) = f(t) - G_L(p)F(x) \tag{2.77}$$

[17] See Definition 1.14.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

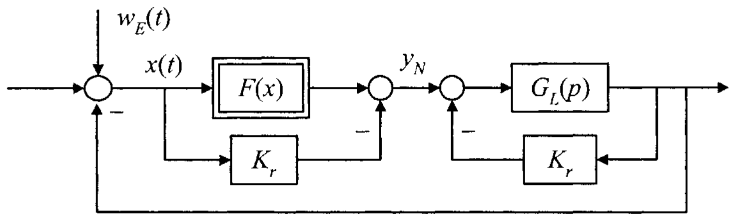


Figure 2.19: Equivalent structure of nonlinear system in the case of unstable linear part.

or by the integral equation:

$$x(t) = f(t) - \int_0^t g(t-\tau)F(x)d\tau \tag{2.78}$$

where $g(t)$ is the weighting function of the linear part of the system, $f(t)$ is the total external signal acting on the system, $F(x)$ is a mathematical description of nonlinear part of the system and $p = d/dt$ is a derivative operator.

Absolute stability of the equilibrium states of the system in Fig. 2.18 is best treated by the frequency criterion of V.M. Popov. In order to apply this criterion, the system must be time-invariant and, moreover, the linear part of the system $G_L(s)$ stable. In the case when the linear part is unstable, the system's structure is replaced by equivalent structure in Fig. 2.19. As is evident from the block diagram, the unstable linear part of the system is stabilized with the linear operator K_r in the negative feedback loop. The same linear operator is placed in parallel with the nonlinear element, so that the influence of the stabilized feedback on the dynamics of the closed-loop system is eliminated. Equivalent external actions (reference and disturbance), equivalent nonlinear element and equivalent linear part of the system are given by:

$$f_E(t) = r_E(t) + w_E(t) = \frac{R(p)}{1+K_rG_L(p)} + \frac{W(p)}{1+K_rG_L(p)} \tag{2.79}$$

$$G_E(p) = \frac{G_L(p)}{1+K_rG_L(p)} \tag{2.80}$$

$$F_E(x) = F(x) - K_rx \tag{2.81}$$

where K_r is a stable linear operator which stabilizes the otherwise unstable linear part of the system $G_L(s)$.

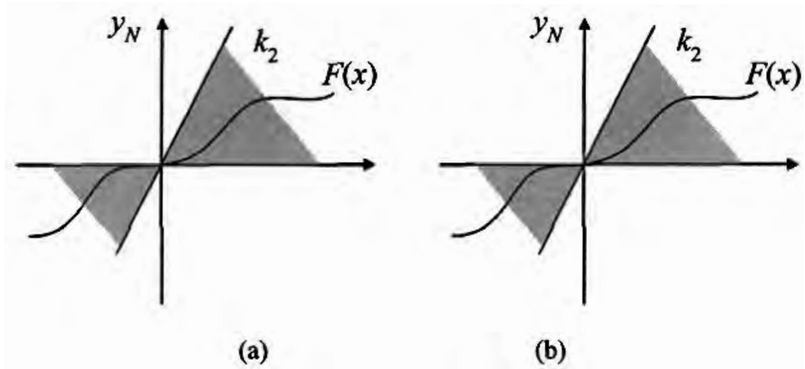


Figure 2.20: Nonlinear characteristic $F(x)$ of (a) class $[0, k_2]$ and (b) class $[K_r, k_2]$.

For the system in Fig. 2.19 the condition that $F_E(x)$ falls in the sector $[0, k_2]$ boils down to the condition that the original nonlinear characteristic of the nonlinear part of the system $F(x)$ will be in the reduced sector $[K_r, k_2]$—in other words it means a reduced region of the static characteristic (Fig. 2.20b). In the following, systems with a stable linear part (Fig. 2.18) and systems with equivalent structure (Fig. 2.19) will be discussed. The external actions on the system $f(t) = r(t) + w(t)$ can be divided in two groups:

1. Bounded external actions $f_1(t)$ described by the equations:

$$|f_1(t)| < M_1; \ t \geq 0 \tag{2.82}$$

which can represent reference and other disturbance inputs.

2. Vanishing (time-decreasing) external actions $f_2(t)$ described by the expressions:

$$\int_0^\infty |f_2(t)|dt < M_2; \ t \geq 0, \tag{2.83}$$

$$\lim_{t \to \infty} f_2(t) = 0 \tag{2.84}$$

which represent initial conditions different from zero.

Dynamics of the system (Fig. 2.18) which was at rest up to the moment $t = 0$, when the external signal $f(t) = f_1(t) + f_2(t)$ was applied, are described by the

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

integral equation:

$$x(t) = f_1(t) + f_2(t) - \int_0^t g(t-\tau)F(x)d\tau \tag{2.85}$$

In order to find the absolute stability of the solution (2.85), it is appropriate to look separately at the equilibrium states of the forced system ($f_1(t) \neq 0$) and those of the unforced one ($f_1(t) = 0$ and $f_2(t) \neq 0$).

In the case when only a vanishing external action $f_2(t)$ acts on the nonlinear system, the absolute stability (global asymptotic stability) of the equilibrium states[18] x_e of the unforced system is considered. The solution (a zero-input response) $x_{zi}(t)$ will be asymptotically stable if:

$$\begin{aligned} x_e &= \lim_{t\to\infty} x_{zi}(t) = M_x = const. \quad \text{or} \\ x_e &= \lim_{t\to\infty} x_{zi}(t) = 0 \end{aligned} \tag{2.86}$$

It must be stressed here that in the case when a stable nonlinear system has the nonlinearity of the type dead zone (Fig. 2.21), its equilibrium states $x_e = x_{zi}(\infty) = M_x \leq |x_a|$ may belong to any part of the dead zone (part of the stability), i.e. the nonlinear system can possess an infinite number of equilibrium states, so the condition of asymptotic stability (2.86) cannot be applied. Therefore it is more appropriate to consider the equilibrium state as stable if the following condition is met:

$$\lim_{t\to\infty} |x_{zi}(t) - x_e| = 0 \tag{2.87}$$

where $x_e = x_{zi}(\infty) = M_x$ is any value inside the dead zone $-x_a < M_x < x_a$. In accordance with the definition of asymptotic stability (2.87), we distinguish the local asymptotic stability—when condition (2.87) is satisfied for small deviations $f_2(t)$ from the equilibrium state, and global asymptotic stability—when the condition (2.87) is satisfied for large deviations $f_2(t)$ from the equilibrium state.

Contrary to linear systems where local asymptotic stability assures global asymptotic stability, in nonlinear systems local asymptotic stability may exist, but not a global one.

Generally, two approaches to the problem of stability are possible. The first one is to find the solution of the differential (2.77) or integral (2.78) equation, which is in practice not applicable because of well-known difficulties. The second one is determining the stability conditions without the inevitable quest for the solution of the dynamic equations of the system. This approach is necessary because of the fact that quite often the nonlinear characteristic $y_N = F(x)$ cannot be determined. Namely, the dynamics of the nonlinear system are changing with

[18] The notion "state" is used here for the scalar signal $x(t)$ which represents the input signal to the nonlinear element, but if $x(t) \to 0$ the equilibrium state in the true sense (vector) will tend to zero.

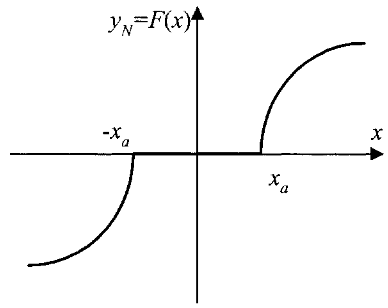


Figure 2.21: Nonlinear static characteristic of dead zone.

change of operating conditions. For example, change of the load or of the supply energy of the control mechanism results in deformity of the static characteristic, which greatly complicates the exact determination of values of the parameters of a differential/integral equation of the system.

The static characteristics of many actuators of modern control systems can be categorized as nonlinear functions with the following properties:

$$\begin{aligned} x \cdot F(x) &> 0, \; x \neq 0 \\ F(0) &= 0 \end{aligned} \tag{2.88}$$

where $F(x)$ is a continuous function:

$$\int_0^{\pm\infty} F(x)dx = \pm\infty \tag{2.89}$$

Nonlinear functions with the properties of (2.88) and (2.89) can have very different graphical presentations. For nonlinear systems given in Fig. 2.18 the interesting nonlinearity classes are those discussed in Section 1.8, where the characteristic of the nonlinear element is situated within the sector bounded by straight lines k_1x and k_2x (Fig. 1.35a) (Voronov, 1979; Mutter, 1973; Aizerman, 1949; Nelepin, 1971):

$$k_1 < \frac{F(x)}{x} < k_2; \; x \neq 0 \tag{2.90}$$

If the nonlinear function $F(x)$ is situated within the sector $[k_1, k_2]$ and fulfils the conditions (2.88) and (2.89), then global asymptotic stability of the system with the function $y_N = F(x)$ is called absolute stability (Voronov, 1979). Very often the nonlinear functions can be of the class $[0, k_2]$ and $[0, \infty]$ which result from (2.88) for $k_1 = 0$ (Fig. 2.20a) and for $k_1 = 0$, $k_2 = \infty$.

In this chapter the frequency criteria for absolute stability of the equilibrium states of the forced nonlinear systems will be discussed—the circle Naumov-Tsypkin frequency criterion[19] (Naumov and Tsypkin, 1964), as well as the Popov frequency criterion for unforced nonlinear systems (Popov, 1973).

2.4 Absolute Stability of Equilibrium States of an Unforced System (Popov Criterion)

The Romanian mathematician V.M. Popov formulated in 1959 the frequency criterion of the absolute stability of a time-invariant unforced nonlinear system which has the structure as in Fig. 2.18. With $f(t) = w(t)$; $r(t) = 0$ a system is described, and a vanishing external quantity $w(t) = f_2(t)$ (initial condition) which satisfies conditions (2.83) and (2.84) is applied. The time-invariant linear part of the system has a stable equilibrium state, while the nonlinear part is a time-invariant single-valued function[20] $F(x)$ which belongs to nonlinear functions of the class $[0, k_2]$ (see Fig. 2.20a) that satisfy conditions (2.88) and (2.89). It is:

$$\begin{gathered} F(0) = 0 \\ xF(x) > 0;\ x \neq 0 \\ \int_0^{\pm\infty} F(x)dx = \pm\infty \\ 0 < \frac{F(x)}{x} < k_2;\ \text{for } x \neq 0 \end{gathered} \tag{2.91}$$

The Popov criterion of absolute stability for an unforced nonlinear system which has only the vanishing external quantity $f_2(t)$, or the initial condition which differs from zero, and shown in the block diagram in Fig. 2.18, is formulated as follows:

THEOREM 2.8 (POPOV CRITERION OF ABSOLUTE STABILITY—G_L STABLE)
The equilibrium state of an unforced nonlinear control system of the structure as in Fig. 2.18 will be globally asymptotically stable–absolutely stable if the following is true:

1. *Linear part of the system is time-invariant, stable and completely controllable.*

[19]In the literature several criteria are found with the name circle criterion. See Sandberg (1964); Zames (1966a, 1966b); with generalization to multivariable systems Rosenbrock (1972); and for infinite dimensional systems Freedman et al. (1969); and Banks (1981).

[20]In later papers this is expanded to the types of linear and nonlinear elements, as is given after (2.75).

2. *Nonlinear function $F(x)$ is of class $[0,k_2]$ with $0 < k_2 < \infty$ and satisfies conditions (2.91).*

3. *There exist two strictly positive real numbers $q > 0$ and an arbitrarily small number $\delta > 0$, such that for all $\omega \geq 0$ the following inequality is true*[21] *(Popov, 1973):*

$$Re\{(1+jq\omega)G_L(j\omega)\} + \frac{1}{k_2} \geq \delta > 0 \tag{2.92}$$

or:

$$Re\{(1+jq\omega)G_L(j\omega)\} + \frac{1}{k_2} > 0 \tag{2.93}$$

where:

$$k_2 < \infty; \; \lim_{\omega\to\infty} G_L(j\omega) = 0 \tag{2.94}$$

The Popov criterion enables the relatively simple determination of the stability of the nonlinear system, based on knowledge of the sector where eventually the nonlinear static characteristic lies and on knowledge of the frequency characteristic of the linear part of the system. There are special cases which allow the linear part to have one or two poles at the origin. In such cases—besides (2.92) and (2.93)—the following must be used:

1. When $G_L(s)$ has one pole at the origin:

$$\lim_{\omega\to+0}\{Im[G_L(j\omega)]\} \to -\infty \tag{2.95}$$

2. When $G_L(s)$ has two poles at the origin:

$$\begin{aligned} &\lim_{\omega\to+0}\{Re[G_L(j\omega)]\} \to -\infty \\ &Im\{G_L(j\omega)\} < 0 \text{ for small } \omega \end{aligned} \tag{2.96}$$

The inequality (2.92) or (2.93) is called the Popov inequality. The criterion is proved by construction of a possible Lyapunov function with the aid of Kalman-Yakubovich lemma. It must be emphasized here that the Popov criterion gives only the sufficient condition for stability. Its importance lies in the fact that the stability of the nonlinear system can be evaluated on the basis of the frequency characteristic of the linear part of the system, without the need for seeking a Lyapunov function. The criterion is constrained by the requirement that the nonlinear static characteristic must be single-valued and that it lies in the first and third quadrants ($k_1 > 0$)—it must pass through the origin.

[21]The condition $\delta > 0$ in (2.92) is necessary for proving the theorem. For practical purposes (2.92) can be written in the form (2.93).

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Later extensions have allowed the absolute stability to be determined by using the Popov criterion (2.93), where q is set depending on the type of nonlinearity. For the case when the nonlinear element $y_N(t) = F[t, x(t)]$ is single-valued and time-varying, it is necessary to put $q = 0$ into (2.93). For double-valued time-invariant nonlinear elements $y_N(t) = F[x(t), \dot{x}(t)]$, the value $q = 0$ in (2.93) is also used.

In the analysis and the synthesis of nonlinear control systems of the proposed structure (Fig. 2.18), the most appropriate procedure is the geometric interpretation of the criterion of absolute stability as it enables the treatment of nonlinear systems by applying frequency methods which were developed in the theory of linear control systems.

2.4.1 Geometrical Interpretation of Popov Criterion

The analytical condition (2.93) can be satisfied for various values of q and for various frequency characteristics of the linear part of the system. In order to determine q which satisfies criterion (2.93), V.M. Popov has proposed a geometrical interpretation of the analytic condition, so that instead of the frequency characteristic of the linear part of the system $G_L(j\omega)$, a modified frequency characteristic $G_P(j\omega)$—the *Popov characteristic* or *Popov plot*—is used (Popov, 1973):

$$G_P(j\omega) = Re\{G_L(j\omega)\} + j\omega Im\{G_L(j\omega)\} = U(\omega) + jV_P(\omega) \qquad (2.97)$$

where $V_p(j\omega) = \omega V(\omega)$ is the imaginary part of the Popov characteristic.

Combining equations (2.93) and (2.97) gives the criterion of absolute stability which include the Popov plot $G_P(j\omega)$:

$$Re\{G_P(j\omega)\} - qIm\{G_P(j\omega)\} + \frac{1}{k_2} > 0 \qquad (2.98)$$

or:

$$U(\omega) - q\omega V(\omega) + \frac{1}{k_2} > 0 \qquad (2.99)$$

The boundary value (2.99) is the equation of the straight line—*Popov line*:

$$U(\omega) = q\omega V(\omega) - \frac{1}{k_2} \qquad (2.100)$$

The Popov line in the $G_P(s)$ plane passes the point $(-1/k_2, j0)$ with the slope $1/q$. The condition of absolute stability (2.93) is satisfied if the position of the $G_P(j\omega)$ plot is to the right of the Popov line, i.e. if the Popov line doesn't intersect the $G_P(j\omega)$ plot (Fig. 2.22). Comparing the Popov characteristics $G_P(j\omega)$ and frequency characteristics $G_L(j\omega)$ of the linear part of the system, the following features can be observed:

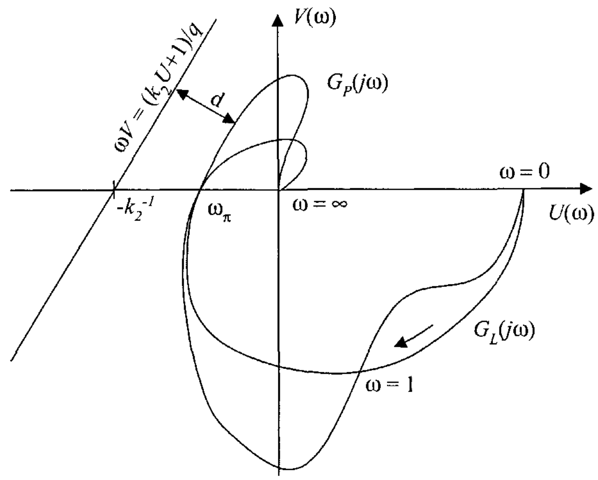


Figure 2.22: Popov line doesn't intersect Popov plot—graphical condition of absolute stability.

1. $G_P(j\omega)$ and $G_L(j\omega)$ intersect the real axis at the same point $\omega = \omega_\pi$.
2. $Im\{G_P(j\omega)\} = \omega V(\omega) = V_p(\omega)$ is an even function of frequency ω, while $Im\{G_L(j\omega)\} = V(\omega)$ is an odd function of frequency ω; the $G_P(j\omega)$ plot is not symmetric with respect to the real axis, while $G_L(j\omega)$ plot is symmetric for $\omega = -\omega$.
3. $G_P(j\omega)$ plot starts for $\omega = 0$ always from the real axis of the complex plane, while $G_L(j\omega)$ plot can have the starting point on the imaginary axis.
4. If $\lim_{\omega\to\infty} G_L(j\omega) = 0$; $\lim_{\omega\to\infty} G_P(j\omega)$ can be equal either to zero or to some other final value.

For example, if $m < n-1$ (see (2.76)), the final point of both the Popov plot and of the amplitude-phase frequency characteristic of the linear part of the system will be identical, i.e. $G_P(j\infty) \equiv G_L(j\infty)$, while if $m = n-1$, the final point of the Popov plot will not be at the origin but on the imaginary axis at the point determined by b_m/a_n, i.e. $G_P(j\infty) \neq G_L(j\infty)$ (see Fig. 2.23c).

In Fig. 2.23 $G_P(j\omega)$ plots are constructed for absolute stable (Fig. 2.23a and b) and unstable (Fig. 2.23c and d) equilibrium states of the nonlinear system. For convex forms of the $G_P(j\omega)$ plot it is possible to determine the maximal value $k_{2max} = K_{crit}$ by drawing the Popov line through the intersection point of the Popov plot and the real axis (Fig. 2.24).

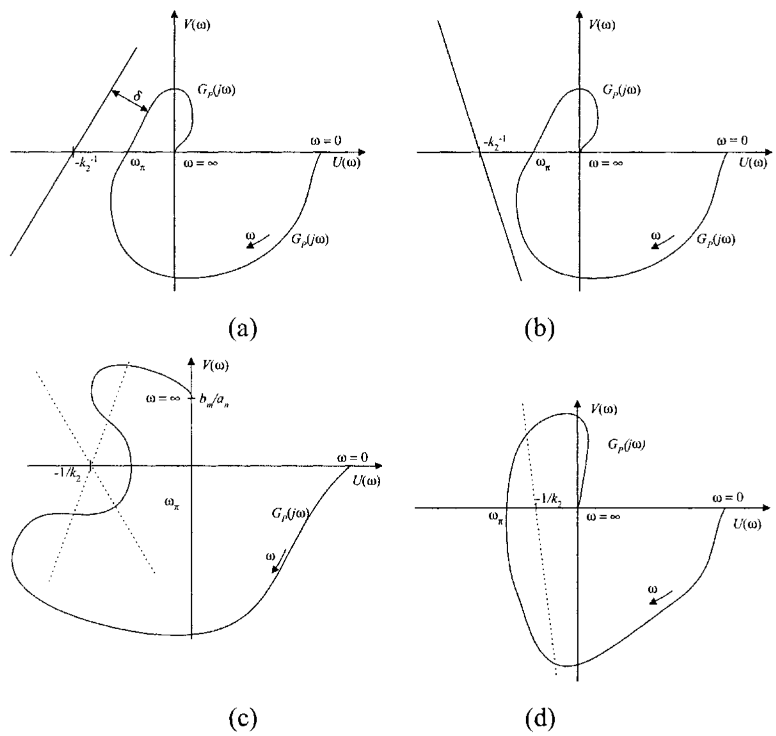


Figure 2.23: Absolute stable ((a) and (b)) and unstable ((c) and (d)) equilibrium states.

As both the Popov plot $G_P(j\omega)$ and $G_L(j\omega)$ plot intersect the real axis at the same point $(-1/k_2, j0)$, and a tangent can be drawn on the convex form of the Popov plot, many authors have used the Nyquist criterion by linearizing the nonlinear characteristics $y_N = F(x)$ from the sector $[0, k_2]$ with the straight line $y_N = k_2 x$, i.e. the Aizerman hypothesis was used. Indeed, in this case the condition that the locus $k_2 G_L(j\omega)$ does not enclose the point $(-1, j0)$ coincides with the condition for absolute stability. Such nonlinear systems are termed stable in the Hurwitz angle, the latter being understood as the angle between the straight line $y_N = k_{2max} x$ and the horizontal axis. As was already mentioned, the Aizerman hypothesis has been rejected (Willems, 1971). It means that in the general case, such linearization is not allowed.

In cases when the Popov plot has a non-convex form, i.e. when it is of much more complex form, the criterion of absolute stability is much more strict. K_{crit} of a convex plot can be much greater than k_{2max} of a non-convex plot, which could mean that the gain of the nonlinear system with a convex Popov plot can be greater than that of the nonlinear system which has a non-convex Popov plot.

Convex forms for Popov plots represent nonlinear systems where the linear part of the system contains cascaded inertial and oscillatory terms and no more than one integral component, with the condition that the damping ratio of the oscillatory terms is $\xi \geq \sqrt{2}/2$. Convex plots $G_P(j\omega)$ also represent nonlinear systems with the following linear parts (Mutter, 1973; Vavilov, 1970):

$$G_L(s) = K \cdot e^{-s\gamma}$$

$$G_L(s) = K \cdot s^{-1} e^{-s\gamma}$$

$$G_L(s) = K \prod_{i=1}^{i=n} (T_i s + 1)^{-1}; \; n \leq 6$$

$$G_L(s) = K s^{-1} \prod_{i=1}^{i=n} (T_i s + 1)^{-1}; \; n \leq 5$$

$$G_L(s) = K \prod_{i=1}^{i=n} (T_i s + 1)^{-1} (\tau_i s^2 + 2\xi \tau_i s + 1)^{-1}; \; n \leq 4; \; \xi \geq \sqrt{2}/2$$

Instead of the inequality (2.93) which contains the variable quantities q and ω, the Popov criterion can be expressed by one variable quantity only — ω. Then it is more appropriate to determine k_2 analytically (Netushil, 1983):

$$\Theta = \alpha_{\max} - \alpha_{\min} < \pi \tag{2.101}$$

where:

$$\alpha(\omega) = \arg \left[G_P(j\omega) + \frac{1}{k_2} \right]$$

α_{max} and α_{min} are maximal and minimal values of $\alpha(\omega)$ in the region $0 < \omega < \infty$ (Fig. 2.26). $\alpha(\omega)$ represents the argument of a complex number $[G_P(j\omega) + 1/k_2]$

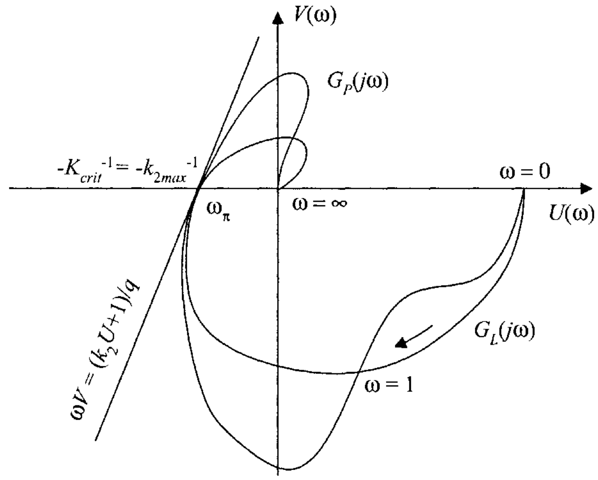


Figure 2.24: Determining of maximum allowable gain.

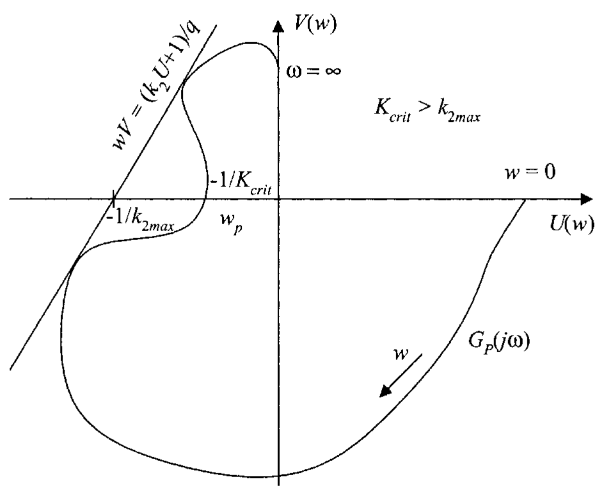


Figure 2.25: Example of a maximally allowable k_{2max} being smaller of K_{crit} which was obtained for convex Popov plot.

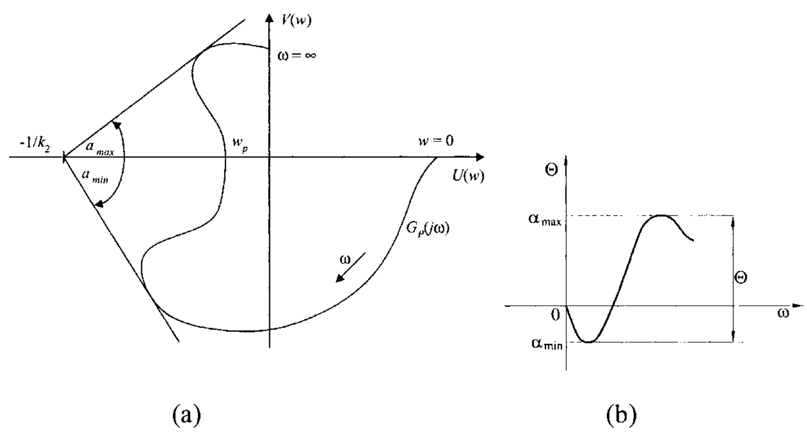


Figure 2.26: Graphical interpretation α_{max} and α_{min} in the region $0 < \omega < \infty$.

when ω changes between $0 < \omega < \infty$. For some ω_1, $\alpha(\omega_1)$ will be the angle of a phasor with the real axis, starting at the point $-\frac{1}{k_2}$ and with the peak at $G_P(j\omega_1)$. From (2.101) it is obvious that the absolute stability of the nonlinear system can be determined without knowing the exact value of the parameter q—it is enough to draw the Popov line through the point $(-1/k_2, j0)$ for at least one slope q with the condition that $0 < q < \infty$.

2.4.2 Absolute Stability with Unstable Linear Part

When the linear part of the system[22] is unstable, it is necessary to accomplish its stabilization with linear feedback (Fig. 2.19). The equivalent transfer function of the linear part of the system $G_E(s)$ is then given by (2.80):

$$G_E(j\omega) = \frac{U(1+K_rU)+K_rV^2}{(1+K_rU)^2+(K_rV)^2} + j\frac{V}{(1+K_rU)^2+(K_rV)^2} = U_E(j\omega) + jV_E(j\omega) \tag{2.102}$$

[22]Some poles of $G_L(s)$ are in the right half-plane of the s-plane.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

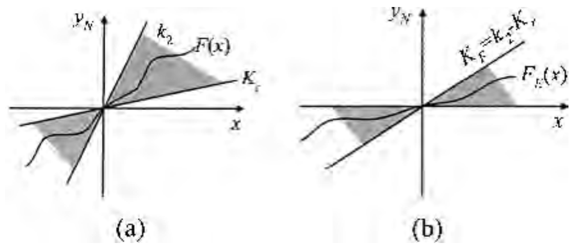


Figure 2.27: (a) Reduced allowable sector $[K_r, k_2]$ for nonlinear characteristic $F(x)$, (b) Corresponding allowed sector $[0, K_F]$ for equivalent nonlinear characteristic $F_E(x)$.

Now the equivalent single-valued time-invariant nonlinear function must satisfy the following conditions:

$$
\begin{aligned}
&F_E(0) = 0 \\
&xF_E(x) > 0;\ x \neq 0 \\
&\int_0^{\pm\infty} F_E(x)dx = \pm\infty \\
&0 < \frac{F_E(x)}{x} < K_F;\ x \neq 0
\end{aligned}
\tag{2.103}
$$

where $K_F = k_2 - K_r$ is the value of the new and smaller slope of the sector $[0, K_F]$ inside which the equivalent nonlinear function $F_E(x)$ may be situated. K_r is the stabilizing linear operator in the feedback of the unstable linear part of the system.

Equivalent nonlinear characteristic $F_E(x)$ given by (2.81) must be inside the reduced sector $[0, K_F]$ in Fig. 2.27b, while the original single-valued nonlinear function $F(x)$ must be inside the reduced sector $[K_r, k_2]$ in Fig. 2.27a.

For this case, absolute stability of the nonlinear system is expressed by the following theorem:

THEOREM 2.9
(POPOV CRITERION OF ABSOLUTE STABILITY—G_L STABILIZED)
The equilibrium state of an unforced nonlinear control system with the structure as in Fig. 2.19 will be globally asymptotically stable–absolutely stable if the following is true:

1. *The linear part of the system is time-invariant, stabilized (equivalent linear part of the system is stable) and completely controllable.*

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

2. *Time-invariant single-valued nonlinear function $F_E(x)$ is of the class $[0, K_F]$.*

3. *There exist two strictly positive numbers $q > 0$ and arbitrarily small number $\delta > 0$ such that for all $\omega \geq 0$ the following inequality is valid:*

$$Re\{(1+jq\omega)G_E(j\omega)\} + \frac{1}{K_F} \geq \delta > 0 \tag{2.104}$$

or:

$$Re\{(1+jq\omega)G_E(j\omega)\} + \frac{1}{K_F} > 0 \tag{2.105}$$

where $0 < K_F < \infty$; $K_F = k_2 - K_r < \infty$; $\lim_{\omega\to\infty} G_E(j\omega) = 0$

$$G_E(j\omega) = [G_E(s)]_{s=j\omega} = U_E(\omega) + jV_E(\omega)$$

Graphical Interpretation of Theorem 2.9

By inserting (2.102) in the inequality—Popov condition (2.105)—and rearranging, the conditions for absolute stability are obtained, and graphical interpretation is possible. From (2.105) and (2.102) it follows:

$$U(1+K_rU) + K_rV^2 - q\omega V + \frac{(1+K_rU)^2 + (K_rV)^2}{K_F} \geq 0 \tag{2.106}$$

or:

$$\begin{aligned} U^2 + \frac{K_F + 2K_r}{K_r(K_r+K_F)}U + V^2 - \frac{qK_F}{K_r(K_r+K_F)}\omega V \\ + \frac{1}{K_r(K_r+K_F)} \geq 0 \end{aligned} \tag{2.107}$$

For $\omega > 0$ and $V^2 > 0$, the inequality (2.107) can be quite well approximated by:

$$\begin{aligned} V_P(\omega) < \frac{K_r(K_r+K_F)}{qK_F}U^2(\omega) \\ + \frac{2K_r+K_F}{qK_F}U(\omega) + \frac{1}{qK_F} \end{aligned} \tag{2.108}$$

where $K_r > 0$; $K_F = k_2 - K_r$; $V_P(\omega) = \omega V(\omega) = Im\{G_P(j\omega)\}$. The inequality (2.108) can be graphically interpreted in the following manner: In order that the nonlinear system is absolutely stable, the Popov plot must be outside of the parabola with peak at the point S with coordinates:

$$\left[-\frac{(K_F+2K_r)}{2K_r(K_r+K_F)}, -\frac{0.25K_F}{qK_r(K_r+K_F)}\right]$$

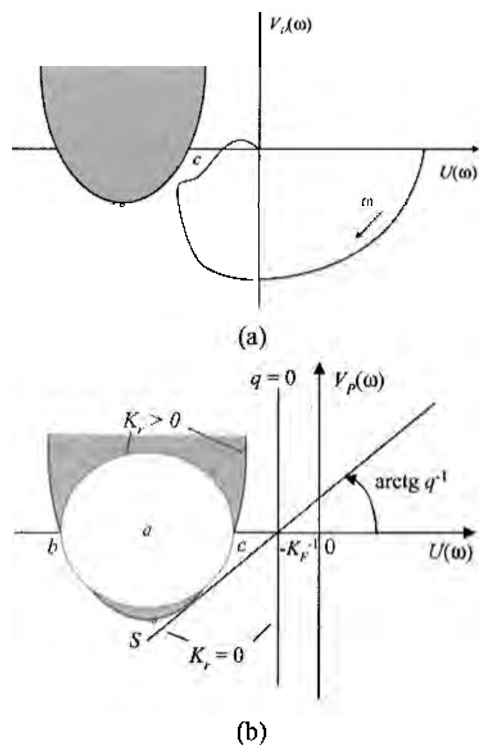


Figure 2.28: Prohibited region for modified Popov plot.

The parabola intersects the real axis at the points $b = -1/K_r$ and $c = -1/(K_r + K_F)$, as illustrated in Fig. 2.28a. When $K_r = 0$[23], inequality (2.108) is transformed into Popov criterion (2.93), while the parabola becomes the Popov line. The line crosses the point $(-1/K_F, j\omega)$ with the slope q^{-1} (Fig. 2.28b), namely $K_F = k_2 - K_r$, or with $K_r = 0$ is $K_F = k_2$.

2.5 Examples of Determining Absolute Stability by Using Popov Plot

The absolute stability of the equilibrium states of nonlinear systems by applying the Popov criterion is based on the modified amplitude-phase characteristic of the linear part of the system—Popov plot $G_P(j\omega)$. The modulus and argument of the function $G_P(j\omega)$ are used to determine analytically the points on the plot:

$$M_P(\omega) = |G_P(j\omega)| = \frac{\mathrm{mod}\,B_P(j\omega)}{\mathrm{mod}\,A_P(j\omega)} \tag{2.109}$$

$$\varphi_P(\omega) = \arg G_P(j\omega) = \arg B_P(j\omega) - \arg A_P(j\omega) \tag{2.110}$$

From (2.109) and (2.110) the real and imaginary parts of $G_P(j\omega)$ are found:

$$Re\{G_P(j\omega)\} = Re\{G_L(j\omega)\} = M_P(\omega)\cos\varphi_P(\omega) = U(\omega) \tag{2.111}$$

$$Im\{G_P(j\omega)\} = \omega \cdot Im\{G_L(j\omega)\} = M_P(\omega)\sin\varphi_P(\omega) = \omega V(\omega) = V_P(\omega) \tag{2.112}$$

When the plot of the linear part of the system $G_L(j\omega)$ is determined analytically or experimentally, the points of the plot $G_P(j\omega_i)$ are obtained by multiplying the ordinate $G_L(j\omega_i)$ with the natural frequency ω_i:

$$G_P(j\omega_i) = Re\{G_L(j\omega_i)\} + j\omega_i Im\{G_L(j\omega_i)\} = U(\omega_i) + jV_P(\omega_i) \tag{2.113}$$

EXAMPLE 2.19

Find the absolute stability of the equilibrium states of the nonlinear system, with single-valued nonlinearities which satisfy conditions (2.91) and whose transfer function for the linear part is[24]*:*

$$G_L(s) = G(s) = \frac{K(1+\tau_1 s)}{(1+T_1 s)(1+T_2 s)(1+2\xi T s + T^2 s^2)} \tag{2.114}$$

where $K = 10$; $T_1 = 1[s]$; $T_2 = 0.01[s]$; $T = 0.1[s]$; $\tau_1 = 0.05[s]$; $\xi = 0.1$.

[23] There is no need to stabilize the linear part of the system, as it is stable.

[24] In the examples the notation K_F instead of k_2 and $G(s)$ instead of $G_L(s)$ is used.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Table 2.2: Values for $G_P(j\omega) = U(\omega) + j\omega V(\omega)$ in Example 2.19.

ω	$U(\omega) = Re\{G(j\omega)\}$	$V(\omega) = Im\{G(j\omega)\}$	$\omega V(\omega) = \omega\{ImG(j\omega)\}$
0	10	0	0
0.5	8.06164	-3.9309	-1.9654
1	5.15443	-4.9547	-4.9547
2	2.24901	-4.1004	-8.2009
4	0.88635	-2.7916	-11.166
8.0347	0	-3.0321	-27.274
10	-4.9505	-4.3235	-24.752
10,5277	-4.491	0	0
12	-1.618	1.10636	12.3716
14	-0.6337	0.58333	8.16665

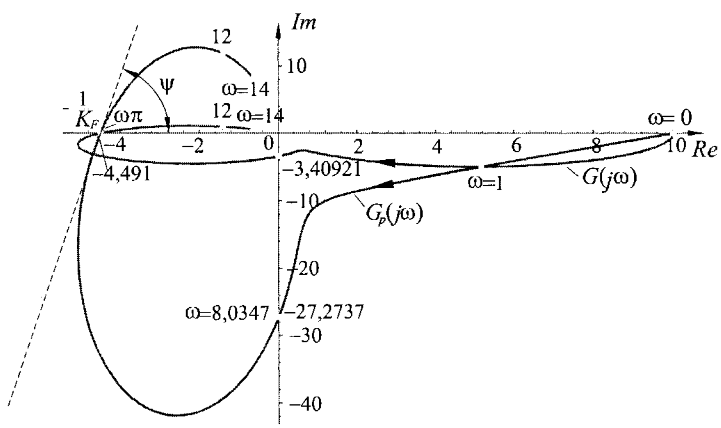


Figure 2.29: Graphical solution of Example 2.19.

Solution. *By substituting $s = j\omega$ in (2.114) and by applying (2.109) to (2.112), the real and imaginary parts of the Popov plot $G_P(j\omega)$ can be determined. See table 2.2 and plots $G(j\omega)$ and $G_P(j\omega)$ (Fig. 2.29).*

Popov plot $G_P(j\omega)$ and the plot $G(j\omega)$ intersect the real axis at the point $(-4.491, j0)$ for the frequency $\omega_\pi = 10.5277[s^{-1}]$. The plot $G_P(j\omega)$ is a convex function, i.e. through the point $(-1/K_F, j0) = (-4.491, j0)$ can be drawn a tangent—Popov line. The slope of the tangent with respect to the real axis is $q^{-1} = \tan\psi = 20$, or $\psi = 87$[deg].

The maximal value of the coefficient K_F for which the nonlinear system is

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

absolute stable is $K_F = 1/4.491 = 0.22$.

The analyzed nonlinear system has absolutely stable equilibrium states for all single-valued nonlinearities of the class $[0, K_F]$ *with the condition that* $K_F < 0.22$.

■

EXAMPLE 2.20

Determine the boundary value of the gain for the absolute stability of the nonlinear system whose structure is as in Fig. 2.17. The transfer function of the linear part of the nonlinear system is:

$$\begin{aligned} G_L(s) = G(s) &= \frac{K}{(1+T_1s)(1+T_2s)(1+T_3s)} \\ &= \frac{25}{(1+0,5s)(1+0,02s)(1+0,01s)} \end{aligned}$$

while the nonlinear characteristic of the class $[0, K_F]$ *is:*

$$0 \leq \frac{F(x)}{x} \leq K_F; \text{ for } x \neq 0 \text{ and } K_F = 2$$

Solution. *To determine the absolute stability of the system, it is often convenient to use a normalized Popov plot* $G_{0P}(j\omega)$*:*

$$G_{0P}(j\omega) = G_P(j\omega) \cdot K_F$$

The normalized Popov plot $G_{0P}(j\omega)$ *is found by means of the normalized plot of the linear part of the system:*

$$G_0(j\omega) = G(j\omega) \cdot K_F$$

$$G_{0P}(j\omega) = Re\{K_F G(j\omega)\} + j\omega Im\{K_F G(j\omega)\}$$

$$\begin{aligned} G_{0P}(j\omega) = &\frac{\left[1-\omega^2(T_1T_2+T_1T_3+T_2T_3)\right]KK_F}{[1+(\omega T_1)^2][1+(\omega T_2)^2][1+(\omega T_3)^2]} \\ &-j\frac{\omega(T_1+T_2+T_3-\omega^2T_1T_2T_3)KK_F}{[1+(\omega T_1)^2][1+(\omega T_2)^2][1+(\omega T_3)^2]} \end{aligned} \tag{2.115}$$

From (2.115) comes Table 2.3 and plots $G_0(j\omega)$ *and* $G_{0P}(j\omega)$ *in Fig. 2.30.*

When $G_{0P}(j\omega) = K_F GP(j\omega)$ *is put into (2.93), the condition of absolute stability with the normalized Popov plot follows:*

$$\begin{aligned} &Re\{G_P(j\omega)\} - qIm\{G_P(j\omega)\} + \frac{1}{K_F} \\ &\qquad = Re\{G_{0P}(j\omega)\} - qIm\{G_{0P}(j\omega)\} + 1 > 0 \end{aligned}$$

Table 2.3: Values for the normalized Popov plot in Example 2.20.

ω	$Re\{G_0(j\omega)\}$	$Im\{G_0(j\omega)\}$	$Im\{G_{0P}(j\omega)\}$
1	39.3723	-21.1854	-21.1854
2	23.4331	-26.4271	-52.8543
4	7.50786	-20.968	-83.8722
8.0347	0	-11.4587	-95.3898
16	-1.96706	-5.49079	-87.8526
32	-1.8235	-1.71313	-54.8201
64	-0.80349	-0.10107	-6.46839
72.894	-0.624989	0	0
128	-0.15189	0.086879	11.1205

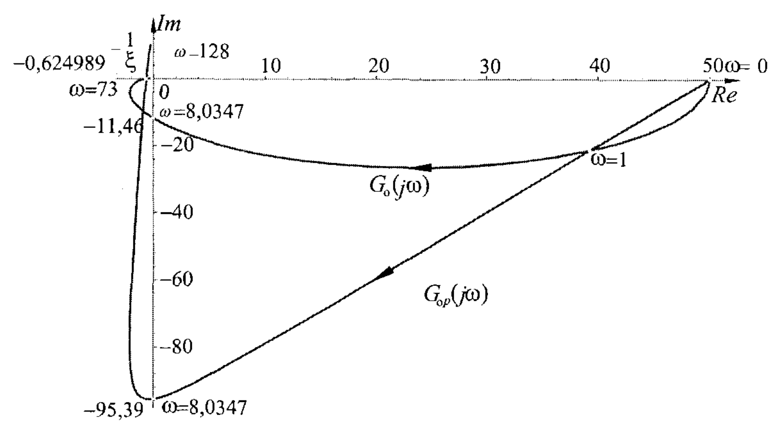


Figure 2.30: Normalized Popov plot—solution of the Example 2.20.

The boundary gain of the nonlinear system comes from (2.115):

$$K_{Fcrit} = \frac{1}{qIm\{G_P(j\omega)\} - Re\{G_P(j\omega)\}}$$

By analogy, the normalized Popov plot $G_{0P}(j\omega)$ and Popov line are passing through the point $(-1/\xi, j\omega)$, so the boundary gain for the normalized Popov plot will be:

$$\xi = \frac{1}{qIm\{G_{0P}(j\omega_i)\} - Re\{G_{0P}(j\omega_i)\}} = \frac{K_{crit}}{KK_F} \tag{2.116}$$

Table 2.4: Values for the normalized Popov plot in Example 2.21.

ω	$Re\{G_0(j\omega)\}$	$Im\{G_0(j\omega)\}$	$Im\{G_{0P}(j\omega)\}$
1	-8.47416	-15.7489	-15.7489
2	-5.28543	-4.68662	-9.37325
4	-2.0968	-0.75079	-3.00314
7.5102	0.51879	0	0
8	-0.59680	-0.0039	-0.03101
16	-0.13727	0.04918	0.78682
32	-0.0214	0.02279	0.7294

For $\omega = \omega_i = 64$ the common point of the Popov line and the normalized Popov plot is:

$$Re\{G_{oP}(j64)\} = -0,75;\ Im\{G_{oP}(j64)\} = -6 \tag{2.117}$$

The slope of the Popov line is $q^{-1} = 50$. From (2.116) follows the boundary value of the gain of the system:

$$K_{crit} = \frac{KK_F}{qIm\{G_{oP}(j\omega_i)\} - Re\{G_{oP}(j\omega_i)\}} = \frac{25 \cdot 2}{-\frac{6}{50} + 0,75} = 79,4$$

■

EXAMPLE 2.21

Determine the boundary gain in the system with a nonlinear characteristic of the class $[0, K_F]$ with the normalized transfer function of the linear part:

$$G_L(s) = G(s) = \frac{K_0}{s(1+T_1 s)(1+T_2 s)(1+T_3 s)}$$

where $K_0 = KK_F = 20[s^{-1}]$; $T_1 = 0.5[s]$; $T_2 = 0.02[s]$; $T_3 = 0.01[s]$.

Solution. *Normalized plots $G_o(j\omega)$ and $G_{0P}(j\omega)$, as well as Table 2.4 are illustrated by Fig. 2.31, together with the Popov line $q^{-1} = 2$.*

In Type 1 control systems [25] there always exists the boundary stability, i.e.:

$$\lim_{\omega \to 0} Im\{G(j\omega)\} = -\infty$$

The boundary gain of the system is according to (2.116):

$$K_{crit} = \frac{KK_F}{qIm\{G_{oP}(j\omega_i)\} - Re\{G_{oP}(j\omega_i)\}} \tag{2.118}$$

[25] Type 1 system has one integrator. System with astatism of first order.

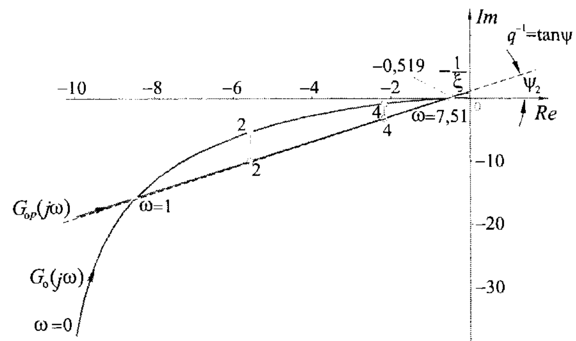


Figure 2.31: Graphical solution of Example 2.21.

Inserting in (2.118) $q^{-1} = \tan\psi = 2$, $\omega_i = 4$ *(common point of Popov line and Popov plot) and the result is:*

$$K_{crit} = \frac{20}{-1,50157+2,0968} = 44.4[s^{-1}]$$

■

2.6 Absolute Stability of an Unforced System with Time-Varying Nonlinear Characteristic

Up to now we have pursued the nonlinear systems under the following conditions:

1. Stable linear part of the system (structure as in Fig. 2.18),

2. Stabilized linear part of the system (structure as in Fig. 2.19).

In both cases the nonlinear part of the system was of the class $[0, k_2]$ for the first case, or of the class $[K_r, k_2]$ for the second case. It was time-invariant without the parallel branch (first case) or with a necessary parallel branch (second case).

In the subsequent text, the main interest lies with the absolute stability of an unforced nonlinear system when the nonlinear part of the system is time-varying. The sufficient conditions for absolute stability of the equilibrium states of the nonlinear system with the structure in Fig. 2.32 are in accordance with the Popov frequency criterion (2.93). Additional discussion for some other properties of

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

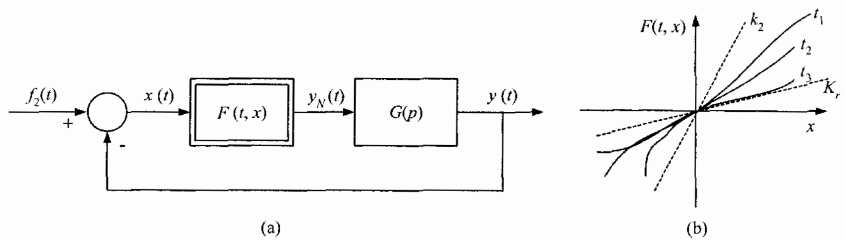


Figure 2.32: Structure of nonlinear system and static characteristics of nonlinear part of system which vary with time, but never come out of sector $[K_r, k_2]$.

the functions $F(t,x)$ and $G_E(s)$ can be found in Tsypkin (1977); Nelepin (1971); Voronov (1979); and Mutter (1973).

For time-dependent nonlinear elements, the coefficient q in the Popov criterion (2.93) is equal to zero. The functions $F(t,x)$ and $G_E(s)$ must satisfy the following properties:

1. Function $F(t,x)$ satisfies the condition of belonging to the sector $[K_r, k_2]$ (Fig. 2.32b) for any t, i.e. it is always true that:

$$\begin{gathered} F(t,0) = 0 \\ K_r < \frac{F(t,x)}{x} < k_2;\ x \neq 0 \end{gathered} \tag{2.119}$$

2. Function $G_E(s)$ is an equivalent transfer function of the linear part of the system, determined by (2.80) and (2.102):

$$G_E(s) = \frac{G_L(s)}{1 + K_r G_L(s)} = \frac{b_m s^m + b_{m-1}s^{m-1} + \ldots + b_o}{a_n s^n + a_{n-1}s^{n-1} + \ldots + a_o};\ n \geq m+2 \tag{2.120}$$

THEOREM 2.10 (ABSOLUTE STABILITY OF TIME-VARYING SYSTEM)
Equilibrium states of an unforced nonlinear system with the structure given in Fig. 2.32 will be absolutely stable if:

$$Re\{G_E(j\omega)\} + \frac{1}{k_2 - K_r} > 0 \tag{2.121}$$

This condition results from the frequency stability criterion (2.105) for $q = 0$, $K_F = k_2 - K_r$.

The criterion (2.121) can also be used for determining the absolute stability of the equilibrium states of unforced time-invariant systems with multi-valued[26]

[26]Sometimes the multi-valued nonlinear characteristic can be treated as a time-varying single-valued one, for example, the three-position relay with hysteresis.

nonlinearities $F(x,\dot{x})$ (Naumov, 1972; Netushil, 1983). Here the nonlinear function must satisfy the following conditions:

$$\begin{gathered} F(0,0)=0 \\ F(x,\dot{x})x>0;\ x\neq 0 \\ 0<\frac{F(x,\dot{x})}{x}<K_F;\ x\neq 0 \end{gathered} \tag{2.122}$$

while the function $G_E(s)$ satisfies conditions (2.120).

Graphical Interpretation of Theorem 2.10

When $q=0$, the condition for absolute stability has the following graphical interpretation. If $q=0$ is inserted in the expression (2.107), the condition of the absolute stability has the form:

$$U^2+\frac{K_F+2K_r}{K_r(K_r+K_F)}U+V^2+\frac{1}{K_r(K_r+K_F)}\geq 0 \tag{2.123}$$

If the first two terms in (2.123) are complemented to the full quadratic term with the additional expression:

$$\frac{(K_F+2K_r)^2}{4K_r^2(K_r+K_F)^2}$$

and after rearranging (2.123):

$$\left[U+\frac{K_F+2K_r}{2K_r(K_r+K_F)}\right]^2+V^2\geq\left[\frac{K_F}{2K_r(K_r+K_F)}\right]^2 \tag{2.124}$$

The expression (2.124) with the replacement of the sign '$\geq$' with the sign '$=$' represents the equation of a circle with the center on the negative[27] real axis in the $G(s)$-plane which has its center at the point a (Fig. 2.28b) with the coordinates:

$$(x_C,y_C)=\left[-\frac{K_F+2K_r}{2K_r(K_r+K_F)}\ ,\ 0\right]$$

and radius:

$$r_C=\frac{K_F}{2K_r(K_r+K_F)}$$

This circle intersects the real axis at the same points as the parabola (see Fig. 2.28b) and lies completely within the parabola which is drawn for $q>0.5$—see expression (2.108). The condition for absolute stability of the time-varying nonlinear systems is graphically interpreted so that the stable linear frequency characteristic may not enter inside the circle. When $K_r=0$, the circle deforms into the

[27]Additional conditions are $K_r>0$ and $K_F>0$.

Popov line which passes the point $-K_F^{-1}$, parallel to the imaginary axis. When $q = 0$ (Fig. 2.28b), the condition (2.93) follows from (2.121).

In the case that the nonlinear element has a static characteristic inside the sector $[0, \infty]$ or $[K_r, \infty]$, and the linear part of the system is stable, then:

$$Re\{(1 + jq\omega)G_L(j\omega)\} > 0 \tag{2.125}$$

or:

$$U(\omega) - qV_P(\omega) > 0 \tag{2.126}$$

and:

$$U(1 + K_rU) + K_rV^2 - q\omega V \geq 0 \tag{2.127}$$

If we substitute $\omega V = V_P(\omega)$ the latter term can be written in the following form:

$$\begin{aligned} V_P(\omega) &\leq \frac{K_r}{q}\left[U^2 + \frac{1}{K_r}U + V^2\right] \\ &= \frac{K_r}{q}\left[\left(U + \frac{1}{2K_r}\right)^2 + V^2 - \left(\frac{1}{2K_r}\right)^2\right] \end{aligned} \tag{2.128}$$

or:

$$V_P(\omega) \leq \frac{K_r}{q}\left[U(\omega)\right]^2 + \frac{1}{q}U(\omega) \tag{2.129}$$

for $V = 0$.

The expression (2.128) has a positive value for:

$$\left(U + \frac{1}{2K_r}\right)^2 + V^2 \geq \left(\frac{1}{2K_r}\right)^2 \tag{2.130}$$

and the extreme at:

$$V_P'(\omega) = 2\left(U + \frac{1}{2K_r}\right) = 0, \text{ or } U_e = -\frac{1}{2K_r}$$

The value of the extreme is:

$$V_P(\omega)_{\min} = -\frac{K_r}{q}\cdot\frac{1}{4K_r^2} = -\frac{1}{4qK_r} \text{ (for } V = 0)$$

Similarly, for $V_p = 0$, the intersection of the parabola and the real axis is obtained:

$$U + \frac{1}{2K_r} = \pm\frac{1}{2K_r}$$

and further:

$$U_1 = -\frac{1}{K_r} \textit{ and } U_2 = 0$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

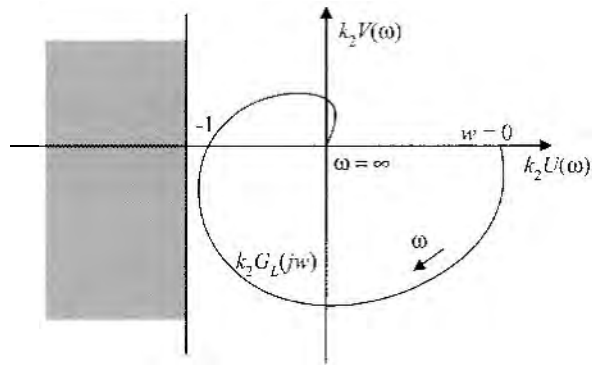


Figure 2.33: Example of absolute stable time-varying nonlinear system.

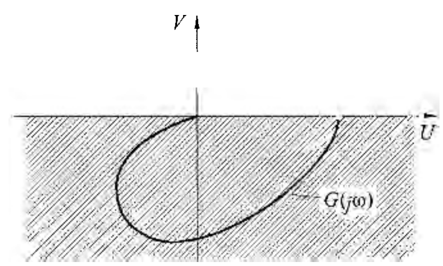


Figure 2.34: Frequency characteristic of stable linear part $G_L(j\omega) = G(j\omega)$.

From inequality (2.126) it follows that the Popov line with the slope q^{-1} may now pass through the origin, while inequality (2.129) means that the Popov plot may not enter inside the region which is bordered by the parabola with the peak in the point S with coordinates $[-1/2K_r, -1/(4qK_r)]$—the parabola intersects the real axis at the points $b = -1/K_r$ and $c = 0$ (Fig. 2.33). When $K_r = 0$, the conditions (2.129) and (2.126) are the same.

For the cases when the nonlinear characteristic is of the class $[0, \infty]$ or $[K_r, \infty]$, inequality (2.130) requires that the frequency characteristic may not enter inside the region which is limited by the circle with the center on the negative real axis at the point marked with a: $[-1/2K_r, 0]$ and the radius $1/2K_r$.

By analogy, for the systems with single-valued time-invariant relay character-

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

istics $y_N = F(x)$ with the properties:

$$\begin{aligned} F(0) &= 0 \\ F(x)x &> 0;\ x \neq 0 \\ 0 < \frac{F(x)}{x} &< \infty;\ x \neq 0 \end{aligned} \tag{2.131}$$

the sufficient conditions for absolute stability emerge from the criterion (2.93):

$$\begin{aligned} Re\{j\omega G_L(j\omega)\} &> 0;\ q > 0,\ \omega \geq 0 \\ Im\{G_L(j\omega)\} &< 0;\ \omega \geq 0 \end{aligned} \tag{2.132}$$

From (2.132) it follows that the frequency characteristic $G_L(j\omega)$ of the stable, linear part of the system must be situated in the lower half of the plane (Fig. 2.34) in order that the nonlinear system be absolutely stable.

EXAMPLE 2.22

It is necessary to determine the conditions for the stability of a relay system in the block diagram, Fig. 2.35a. The nonlinear element is a three-position relay with hysteresis and its static characteristic is given in Fig. 2.35b.

Remark. *As can be seen, the linear part of the system has one pole at the origin and doesn't satisfy the condition (2.95).*

Solution. *The multi-valued nonlinear characteristic (Fig. 2.35b) can be represented by an equivalent structure which contains a single-valued nonlinear element — three-position relay without hysteresis, $y_N = F(x_n)$ with a positive feedback H as is given in Fig. 2.36.*

By replacing $F(x,\dot{x})$ with the equivalent nonlinear characteristic $y_N = F(x_n)$, the system from Fig. 2.35 can be replaced by an equivalent structure as in Fig. 2.37 (Netushil, 1983).

In the system from Fig. 2.37 nonlinear characteristic $y_N = F(x_n)$ is single-valued of the class $[0, 1/x_a]$, while the equivalent linear part is determined by the transfer function:

$$G_{Ln}(s) = \frac{K_L}{sT(1+sT)} - H \tag{2.133}$$

The transfer function of the equivalent linear part of the system (2.133) doesn't satisfy the necessary conditions for the closed-loop system to be absolutely stable for two reasons:

1. *$G_{Ln}(s)$ has one pole at zero, i.e. $G_{Ln}(s)$ is neutrally stable and doesn't satisfy (2.95),*
2. *$G_{Ln}(s)$ doesn't satisfy Theorem 2.8, i.e. $\lim_{\omega \to \infty} G_{Ln}(j\omega) \neq 0$.*

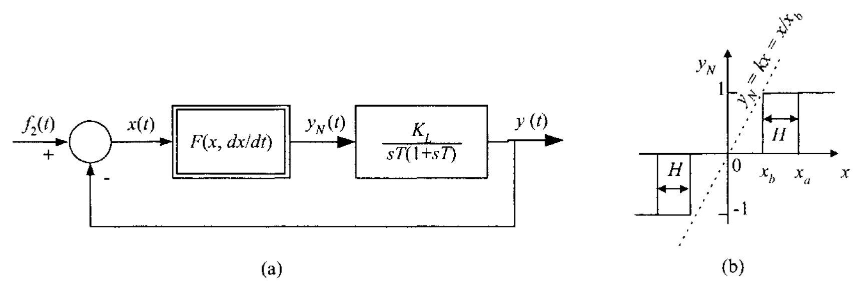


Figure 2.35: Block diagram of nonlinear system (a) and static characteristic of nonlinear element (b).

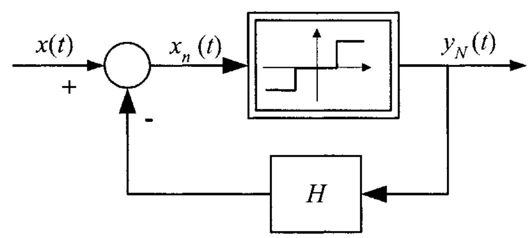


Figure 2.36: Equivalent structure for multi-valued nonlinear characteristic of three-position relay with hysteresis (Fig. 2.35b).

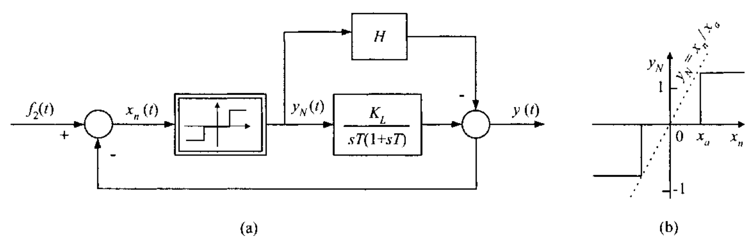
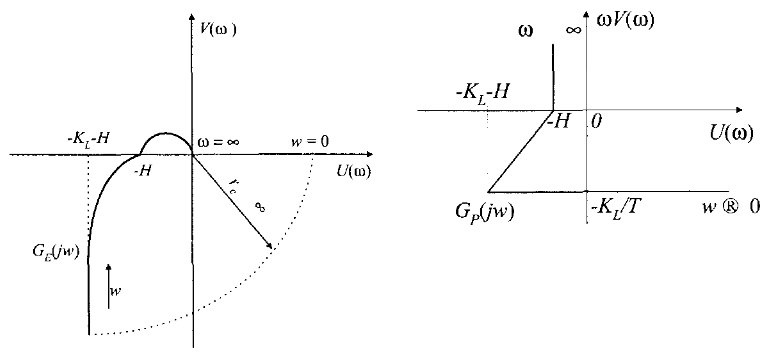


Figure 2.37: Equivalent structure (a) and equivalent nonlinear characteristic (b) for Example 2.22.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016



(a) (b)

Figure 2.38: Magnified frequency characteristic (a) equivalent $G_E(j\omega)$ and (b) Popov $G_P(j\omega)$.

In order to apply the criterion of absolute stability, it is necessary to determine the equivalent transfer function $G_E(s)$ *which will satisfy the condition of stability of the linear part as well as Theorem 2.8 or* $\lim_{\omega \to \infty} G_E(j\omega) = 0$. *For such a purpose the following equivalent transfer function can be used:*

$$G_E(s) = \frac{G_L(s)}{(1+\alpha\, s)+\beta\; G_L(s)} \tag{2.134}$$

where α *and* β *are small parameters* $\alpha << T$, $\beta << 1$.

From (2.134) and (2.133) results:

$$\begin{aligned} G_E(s) &= \frac{K_L - HsT(1+sT)}{sT(1+\alpha s)(1+sT)+\beta\;\;[K_L - HsT(1+sT)]} \\ &\approx \frac{K_L - HsT(1+sT)}{\alpha T^2 s^3 + T^2 s^2 + sT + \beta\; K_L} \end{aligned} \tag{2.135}$$

As can be seen, the equivalent transfer function (2.135) complies now with the required conditions if $T > 0$ *and* $K_L > 0$.

The plot $G_E(j\omega)$ *in Fig. 2.38a consists of three parts:*

1. *The range of low frequencies* ($\omega <<$), *where the higher-order terms of* ω

are neglected and $G_E(j\omega)$ can be approximately expressed by:

$$G_E(j\omega) \approx \frac{K_L}{\beta K_L + j\omega T}$$

In low-frequency range the plot $G_E(j\omega)$ is a circle with radius $r_c = 1/\beta$, which goes to infinity when $\beta \to 0$.

2. *The range of medium frequencies, where the terms with small parameters α and β can be neglected:*

$$G_E(j\omega) \approx \frac{K_L}{j\omega T(1+j\omega T)} - H$$

and the plot $G_E(j\omega)$ behaves as an inertial integration element shifted for H (Fig. 2.38a).

3. *The range of high frequencies where low-power terms of ω can be neglected. For the HF range it is true that:*

$$G_E(j\omega) \approx -\frac{H}{1+j\omega\alpha}$$

In the HF range the plot $G_E(j\omega)$ is a half-circle with radius $0.5H$ in the second quadrant (Fig. 2.38a).

From the $G_E(j\omega)$ plot, the Popov plot $G_P(j\omega)$ can be constructed in such a way that the ordinates $G_E(j\omega)$ are multiplied by the frequency ω (Fig. 2.38b). In the low-frequency range the arc of the circle is of infinite radius[28]*, i.e. it becomes a straight line,*[29] *starting at the point $(-K_L - H, -jK_L/T)$, and it runs parallel to the real axis and ends at $+\infty$ as $\omega \to 0$. In the medium-frequency range $G_P(j\omega)$ becomes a straight-line segment from the point $(-K_L - H, -jK_L/T)$ to the point $(-H, j0)$. In the HF range the circle with the radius $0.5H$ (for $G_E(j\omega)$) becomes in the Popov plot $G_P(j\omega)$ a straight line parallel to the imaginary axis from the point $(-H, j0)$ as is shown in Fig. 2.38b. The Popov line can be drawn on a segment of the real axis between $-\infty$ and $(-K_L - H)$, i.e. to the left of the point $(-K - H, j0)$. The condition for absolute stability is:*

$$\frac{1}{x_a} < \frac{1}{K_L + H} \tag{2.136}$$

or:

$$K_L < x_a - H = x_b \tag{2.137}$$

[28] For $G_E(j\omega)$.
[29] For $G_E(j\omega)$.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

or:

$$k_{2max} = \frac{1}{K_L} < \frac{1}{x_b} \tag{2.138}$$

The condition for absolute stability for a given system can be interpreted in such a way that $y_N(x)$ must be inside the sector $[0, k]$, defined by the angle $\beta = y_N/x \leq k = 1/x_b < 1/K_L$ (see the dashed line in Fig. 2.35b). ■

2.7 Absolute Stability of Forced Nonlinear Systems

The Popov frequency method gives sufficient conditions for the absolute stability of the equilibrium state of an unforced system. The method is not appropriate to assure the stability of a nonlinear system which has forced disturbances (reference quantities, external disturbances, measurement noise, etc.). The procedure which gives sufficient conditions for absolute stability for such cases has been proposed by Naumov and Tsypkin (1964). This will be presented next.

The dynamics of the nonlinear system of the structure as in Fig. 2.39, forced by an external quantity $f_1(t)$, is described by a nonlinear integral equation:

$$x(t) = f_1(t) - \int_0^t g(t-\tau)F[x(\tau)]\,d\tau \tag{2.139}$$

where:

$G(p)$ – transfer function of the stable linear part of the system, expressed by the derivative operator $p = d/dt$,

$F(x)$ – single-valued nonlinear characteristic of the nonlinear part of the system,

$f_1(t)$ – external bounded action which doesn't vanish, while satisfying condition (2.79).

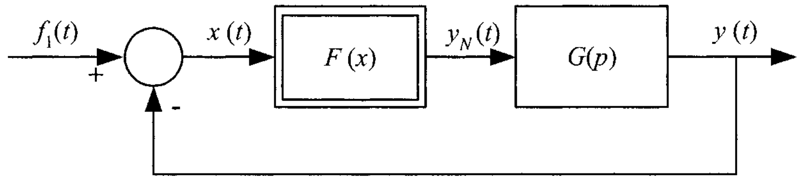


Figure 2.39: Block diagram of forced nonlinear system.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

If a bounded external action $f_1(t)$ acts in the time interval $0 < t < \infty$, the restored state[30] of the forced state or zero-state response is described by the nonlinear integral equation:

$$x_{ZS}(t) = f_1(t) - \int_0^t g(t-\tau)F\left[x_{ZS}(\tau)\right]d\tau \tag{2.140}$$

The sufficient condition for the restored state (zero-state response) of the forced system (Fig. 2.39) to be absolutely stable is satisfaction of the following theorem:

THEOREM 2.11
(ABSOLUTE STABILITY OF FORCED NONLINEAR SYSTEM, NAUMOV AND TSYPKIN)
The restored state of the system given by the structure in Fig. 2.39 will be absolutely stable if the following conditions are met:

1. *The external action to the system is bounded, see (2.79),*
2. *The equilibrium state of the linear part of the system is stable,*
3. *The nonlinear characteristic $F(x)$ is of the class $[0,k_2]$, and its derivative satisfies:*
$$0 \leq \frac{dF(x)}{dx} \leq k_2 \; ; \; 0 < k_2 < \infty \tag{2.141}$$
4. *The following inequality is true:*
$$Re\{G_L(j\omega)\} + \frac{1}{k_2} \geq 0 \; ; \; \omega \geq 0 \tag{2.142}$$

Graphical Interpretation of Theorem 2.11

The expression (2.142) is obtained from Popov condition (2.93) when $q = 0$. It can be graphically interpreted so that for the frequency characteristic of the linear part of the system $G_L(j\omega)$ the forbidden region is to the left of the straight line which passes through the point on the real axis $(-k_2^{-1}, j0\,)$. This line[31] is parallel to the imaginary axis ($q = 0$). Besides, the slope of the nonlinear characteristic must be greater than or equal to zero, and smaller than or equal to k_2. An example of the permissible nonlinear characteristic is given in Fig. 2.40a, while the example for a nonlinear characteristic that is not permissible is shown in Fig. 2.40b. In the

[30] The notion 'state' is used here for the signal at the input to the nonlinear part of the system, $x(t)$. Some authors use the notion 'process' instead.

[31] See the straight line parallel with the imaginary axis in Fig. 2.28b.

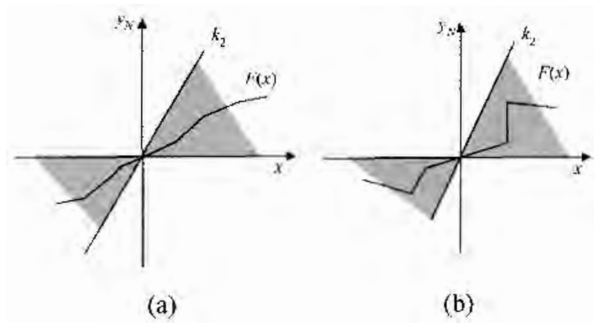


Figure 2.40: Example of allowed (a) and prohibited (b) nonlinear characteristics.

plane $[k_2U(\omega), k_2V(\omega)]$ from (2.142) can be derived:

$$Re\{k_2 \cdot G_L(j\omega)\} + 1 \geq 0 \tag{2.143}$$

and it follows that:

$$Re\{k_2U(\omega) + jk_2V(\omega)\} + 1 \geq 0 \tag{2.144}$$

Geometrically condition (2.144) means that the plot $k_2G_L(j\omega)$ for all $\omega \geq 0$—which is obtained by the substitution of the nonlinear element $F(x)$ with the linear operator of the gain k_2—is situated to the right of the line $k_2U(\omega) = -1$ (Fig. 2.41).

2.7.1 Absolute Stability of Forced Nonlinear Systems with an Unstable Linear Part

If the linear part is unstable, it is necessary to stabilize it by a negative feedback K_r and at the same time, an equivalent transfer function of the linear part of the system must be taken into account. Then:

$$G_E(s) = \frac{G_L(s)}{1 + K_rG_L(s)}; \lim_{\omega \to \infty} G_L(j\omega) = 0 \tag{2.145}$$

$$F_E(x) = F(x) - K_rx \tag{2.146}$$

$$K_r < \frac{F(x)}{x} < k_2 ; xF(x) > 0, x \neq 0 \tag{2.147}$$

The criterion of absolute stability of the equilibrium states of a forced nonlinear system with an unstable linear part, under the excitation of the bounded non-vanishing external quantities, has been formulated by Tsypkin (1977), Naumov (1972), Yakubovich (1965), Voronov (1979); Netushil (1983); and Naumov and Tsypkin (1964).

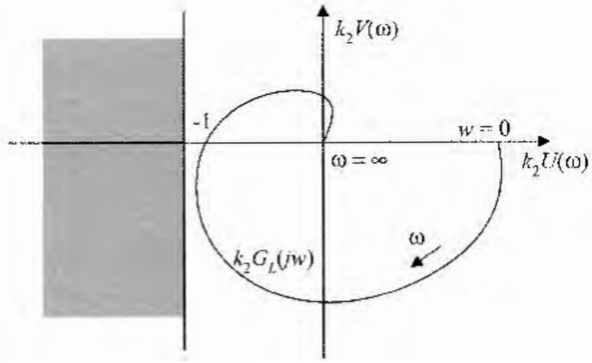


Figure 2.41: Prohibited region for frequency characteristic of linear stable part $k_2G_L(j\omega)$.

THEOREM 2.12
(ABSOLUTE STABILITY OF FORCED NONLINEAR SYSTEM WITH AN UNSTABLE LINEAR PART, NAUMOV AND TSYPKIN)
A sufficient condition for the absolute stability of the equilibrium states of a forced nonlinear system with an unstable linear part is to satisfy the frequency criterion:

$$Re\left\{\frac{G_L(j\omega)}{1+K_rG_L(j\omega)}\right\}+\frac{1}{k_2-K_r}\geq 0\;;\;\omega\geq 0 \tag{2.148}$$

and the condition:

$$K_r+\gamma=\frac{dF(x)}{dx}\leq k_2+\gamma \tag{2.149}$$

where γ is an arbitrary small positive quantity (see Fig. 2.42).

Condition (2.149) is valid for the nonlinearities which pass through the first and third quadrants[32], while condition (2.148) results from the Popov frequency criterion (2.105) when $q=0$. Comparing both criteria (2.105) and (2.148), it is obvious that the conditions for absolute stability of the equilibrium states of a forced nonlinear system are much more strict than those for an unforced nonlinear system, with additional constraints (2.149) which the derivative of the nonlinear function must satisfy.

In situations when $K_r=0$ in (2.148), i.e. when the linear part of the system $G_L(j\omega)$ is stable, instead of (2.148) the expression for the frequency criterion of the absolute stability (2.142) is obtained.

[32]See the properties (2.91)

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

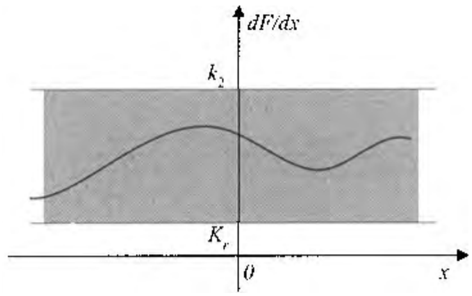


Figure 2.42: The gradient of nonlinear characteristic must be limited.

Graphical Interpretation of Theorem 2.12

Frequency criterion (2.148) can possess different geometrical interpretations, depending on the choice of the complex plane. With the use of a polar coordinate system it is convenient to make use of the constant amplitude and phase contours in the complex plane (Kuljača and Vukić, 1985; Netushil, 1983). Substitution of $K_r > 0, A = k_2/K_r > 1$ in (2.148) yields:

$$Re\left\{\frac{k_2G_L(j\omega)}{A+k_2G_L(j\omega)}\right\}+\frac{1}{A-1}\geq 0 \tag{2.150}$$

By inserting $G_L(j\omega) = U(\omega) + jV(\omega)$ into expression (2.150) and rearranging, equations for circles in the plane $[k_2U(\omega), k_2V(\omega)]$ are obtained. These circles correspond to different sector dimensions inside which the nonlinear characteristic can reside—sector dimensions are expressed through the ratio $A = k_2/K_r =$ const.:

$$\left[k_2U(\omega)+\frac{1}{2}(A+1)\right]^2+k_2V(\omega)=\frac{1}{4}(A-1)^2 \tag{2.151}$$

The family of circles in the plane $[k_2U(\omega), k_2V(\omega)]$ given by (2.151) for various $A = k_2/K_r$ are presented in Fig. 2.43. All of them pass through the point $(-1, j0)$, have the radius $R = (A-1)/2$ and the center on the negative real axis which is situated left of the straight line $k_2U(\omega) = -1$. Based on the equation (2.150) the *circle criteria of absolute stability* can be formulated (Tsypkin, 1977; Nelepin, 1971; Netushil, 1983; Mohler, 1991; Cook, 1986).

Criterion (2.150) will be satisfied if the plot $k_2G_L(j\omega)$ for $K_r \neq 0$ is outside the corresponding A circle. From Fig. 2.43 it is obvious that to every circle there corresponds a value $A = k_2/K_r$ which can change from $A = 1$ to $A = \infty$. For $A \rightarrow \infty$ ($k_2 \rightarrow \infty$ or $K_r = 0$), the circle deforms to a vertical line which passes through

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

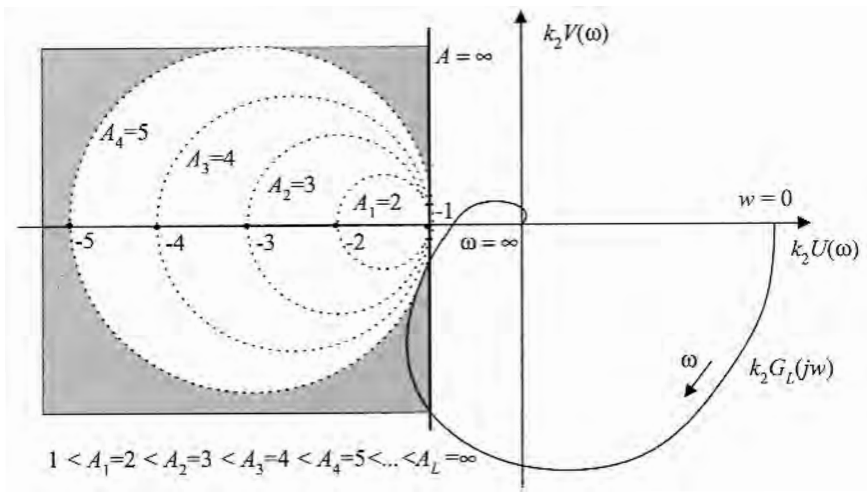


Figure 2.43: Circle criterion of absolute stability in the plane $[k_2U(\omega), k_2V(\omega)]$.

the point $(-1, j0)$. For $A = 1$[33] the circle transforms into a point $(-1, j0)$—the sufficient conditions for the absolute stability of the nonlinear system become equal to the necessary and sufficient conditions of the linear system's stability for $k_2 = K_r$.

To determine the absolute stability of the forced system whose nonlinear characteristic is known, we may test the position of the plot $k_2G_L(j\omega)$ in relation to the circle A which is determined by the boundaries of the sector where the nonlinear characteristic resides. With the proper choice of the parameters of the linear part for which the plot $k_2G_L(j\omega)$ is to the right of the circle A, it is possible to design an absolutely stable system.

When $G_L(j\omega)$ is defined, it is possible to determine the dependence $K_r = f(k_2)$ by setting different values of $k_2 =$ const. and obtaining $k_2G_L(j\omega)$ which touches the circle. When k_2 and A are known, the corresponding value of K_r can be found, by which the absolute stability is obtained.

2.8 Conclusion

Stability is the property of an equilibrium state. We say that a system is stable if all its equilibrium states are stable. For proper operation of any technical[34]

[33] The nonlinear part of the system becomes a linear one.

[34] Or any other system like economic, social, biological, etc.

system, stability is the main requirement. When stability is ensured, then other requirements need to be obtained such as dynamic performance indices, accuracy, sensitivity, robustness, etc. The topic of stability of nonlinear systems given here has shown that in general nonlinear systems are richer than linear systems in varieties of possible behavior. Linear systems have one equilibrium state, contrary to nonlinear systems. Because of that we can speak in linear system theory about stability of a system. Stability here does not depend on initial conditions or type of the excitation signal. However, nonlinear systems can have one equilibrium state which is locally stable, another which is locally unstable and so on. Stability must be analyzed in respect of a small (local) or large (global) change of state around each equilibrium state. Stability can depend not only on initial conditions, but also on the type of excitation signal, as well as on characteristics of a system at and around a particular state[35].

This chapter has explained some very important achievements of the stability theory. Lyapunov's work revolutionized our thinking and perspective about stability. New insights gave researchers new tools for stability analysis, enabling design of control systems with ensured stability throughout the system's state space. Absolute stability by V.M. Popov brought us the method of stability analysis for a specific class of unforced nonlinear systems. Due to its frequency interpretation it became popular among control engineers. Later, generalization of this method to include forced systems was also well accepted in the control community. Stability is a very important property and because of that this chapter is so extensive.

In what follows we will talk about linearization of nonlinear systems and in the rest of the book the harmonic linearization will be covered in more detail. The reader should bear in mind that any linearization is valid only for a specific local area of the state space close to the working point of a system. Due to that some stability issues given in this chapter will be lost, because only one equilibrium point will be of our prime concern.

[35]Equilibrium state.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Chapter 3

Linearization Methods

Dealing with nonlinear systems is difficult because known mathematical methods do not give us powerful enough means to analytically attack many problems which we encounter here. In order to simplify things and make them more manageable, linearization is quite often used. By linearizing nonlinear system about a single equilibrium state, the linear systems theory achievements can be explored. Some of our introductory comments from the first chapter should again be seen in view of this new perspective. Conventional linearization methods which are often used in engineering practice will be presented. Some new techniques, such as feedback linearization method[1], input-output pseudolinearization[2], and linearization about a trajectory[3], are left out because we believe that they are more appropriate to be covered in a text dealing with the design of nonlinear control systems.

Often it is most convenient and simple to analyze the dynamics of a nonlinear system by using a linearized mathematical model of the real system. With such analysis a nonlinear mathematical model is replaced by a linear one. Nevertheless caution must be taken, since such substitution is not always possible. Namely such substitution yields a valid result only if the linear effects are dominant. On the other hand if nonlinear effects are dominant, the linearization can only aggravate the description of the system, which leads to wrong conclusions.

As already mentioned, a large proportion of real technical systems can be represented by a structure consisting of a linear and a nonlinear part, the latter being described by static characteristics. Four linearization procedures are generally applied: conventional linearization, harmonic linearization, statistical linearization and a combined one (harmonic plus statistical) (Csáki, 1972; Popov and Pal'tov,

[1]See Jacubczyk and Respondek (1980); Isidori (1995); Krstić et al. (1995).

[2]This is a special case of feedback linearization; see Jacubczyk and Respondek (1980).

[3]Linearization along a trajectory differs from the linearization about an equilibrium because this linearization does not restrict the nonlinear system to stay close to a single equilibrium point, as is the case with the linearization about equilibrium.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

1960; Gelb and Vander Velde, 1968).

Conventional linearization is applied when the static characteristic of a system (or its elements) is a smooth function which can be presented graphically and analytically. The procedures of harmonic and statistical linearization are used in the case when the static characteristics of nonlinear elements are not smooth functions.

The references Jacubczyk and Respondek (1980); Lawrence and Rugh (1994) and Isidori (1995) give an extensive overview of this important topic for nonlinear systems.

3.1 Graphical Linearization Methods

Graphical procedures are used in the case when the static characteristics of a nonlinearity are defined graphically and have the property of a smooth function. Linearization is achieved either by tangent or by averaging (secant) methods. Such a linearization is acceptable only if in the vicinity of the operating point, the nonlinear characteristics have only small deviations, i.e. when an approximating line remains close to the nonlinear characteristic. The greater the region where the coincidence of the straight line and nonlinear static characteristic is acceptable, the better is linearization, and the linearized model can be used for more operating points. The basic demand for a linearized model, i.e. to be acceptable for the greatest possible number of operating points is thus fulfilled.

Tangent Method

The method is illustrated in Fig. 3.1. Fig. 3.1a illustrates the linearization method of the static characteristic of a nonlinearity at nominal operating point $0_1\,(x^0, y_N^0)$. With the assumption that oscillations in the system are within prescribed limits, the static characteristic can be approximated by a tangent with variables Δx and Δy_N:

$$\Delta y_N = K \cdot \Delta x \tag{3.1}$$

With tracking systems, oscillations in the nominal operating mode are not allowed, i.e. the static characteristic is approximated by the tangent equation:

$$y_N = Kx \tag{3.2}$$

Secant Method (Averaging Method)

The method is shown in Fig. 3.2. It is applied in engineering practice when the graphical representation of the nonlinear static characteristic in the range of input

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

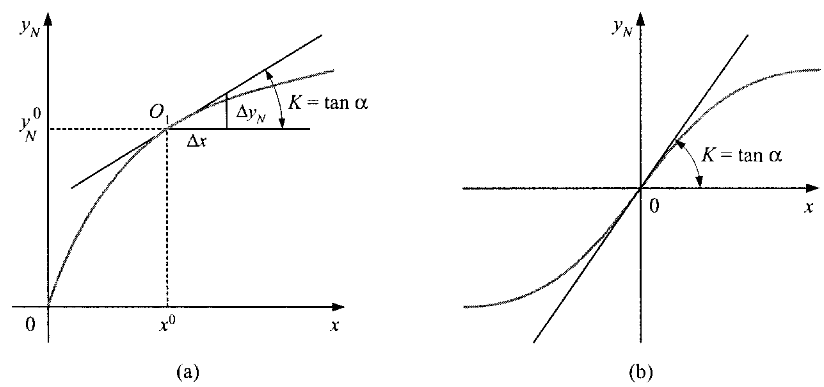


Figure 3.1: Linearization of smooth nonlinear static characteristics by the tangent line.

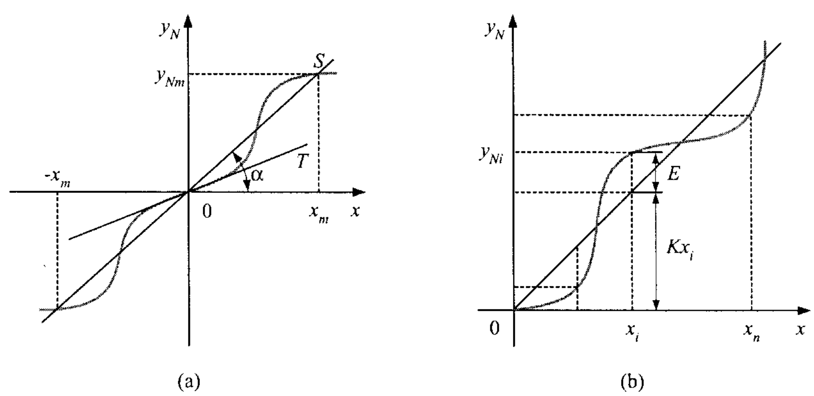


Figure 3.2: Linearization of smooth nonlinear static characteristic by the secant line.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

variations[4] $-x_m < x < x_m$ can be approximated by a line:

$$y_N = K_S x \tag{3.3}$$

where $K_s = \tan\alpha$. Secant approximation lines can be more accurate than the tangent method (line T), which is shown in Fig. 3.2a where the line S is more accurate than the line T, except in the vicinity of the point $x = y_N = 0$. Coefficient K_S can be still better defined by the least-squares method (Fig. 3.2b). Here, the curved static characteristic $y_N = F(x)$ is approximated by the line $y_N = Kx$, whereby K is calculated from the equation:

$$\min\left[E^2\right] = \min\left[\sum_{i=1}^{n}(Kx_i - y_{Ni})^2\right] \tag{3.4}$$

i.e. the sum of squares of the difference of Kx_i (secant points) and $y_{Ni} = F(x)$ (function points) must be a minimum. To obtain the extreme value, the equation has to be differentiated and the result equated with zero:

$$\frac{dE^2}{dK} = \sum_{i=1}^{n} 2(Kx_i - y_{Ni})x_i = 0$$

respectively:

$$K\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_{Ni}x_i \tag{3.5}$$

From (3.5) follows:

$$K = \frac{\sum\limits_{i=1}^{n} y_{Ni}x_i}{\sum\limits_{i=1}^{n} x_i^2} \tag{3.6}$$

By approximating a static characteristic by either of two above-mentioned methods, it is possible, for many nonlinear systems, first to carry out linear analysis of the dynamic behavior, and thereafter by using nonlinear theory to obtain more precise results.

3.2 Algebraic Linearization

For a large number of nonlinear functions of one variable which appear frequently in mathematical calculations, for small values of the argument Δx it is appropriate to apply expressions obtained by algebraic approximation. The most frequent are the functions presented in Table 3.1.

[4]It must be emphasized that in this chapter all input signals to a nonlinear element are denoted with x, which is otherwise reserved for state variables.

Table 3.1: Algebraic approximations of the most frequent nonlinear functions.

Nonlinear function	Linear approximation	First neglected term
$\dfrac{1}{1+\Delta x}$	$1-\Delta x$	$(\Delta x)^2$
$(1+\Delta x)^n$	$1+n\Delta x$	$\dfrac{n(n-1)}{2}(\Delta x)^2$
$\sqrt{1+\Delta x}$	$1+\dfrac{1}{2}\Delta x$	$-\dfrac{1}{8}(\Delta x)^2$
$\dfrac{1}{\sqrt{1+\Delta x}}$	$1-\dfrac{1}{2}\Delta x$	$\dfrac{3}{8}(\Delta x)^2$
$\sqrt[3]{1+\Delta x}$	$1+\dfrac{1}{3}\Delta x$	$-\dfrac{1}{9}(\Delta x)^2$
$\dfrac{1}{\sqrt[3]{1+\Delta x}}$	$1-\dfrac{1}{3}\Delta x$	$\dfrac{2}{9}(\Delta x)^2$
$e^{\Delta x}$	$1+\Delta x$	$\dfrac{1}{2}(\Delta x)^2$
$a^{\Delta x}$	$1+(\ln a)\Delta x$	$\dfrac{(\ln a)^2}{2}(\Delta x)^2$
$\ln(1+\Delta x)$	Δx	$-\dfrac{1}{2}(\Delta x)^2$
$\sin\Delta x$, $\sinh\Delta x$, $\mathrm{arcsinh}\Delta x$, $\arcsin\Delta x$	Δx	$-\dfrac{1}{6}(\Delta x)^3$, $\frac{1}{6}(\Delta x)^3$
$\cos\Delta x$, $\cosh\Delta x$, $\mathrm{arccosh}\Delta x$, $\arccos\Delta x$	1	$\dfrac{1}{2}(\Delta x)^2$, $-\frac{1}{2}(\Delta x)^2$
$\tan\Delta x$, $\tanh\Delta x$, $\mathrm{arctanh}\Delta x$, $\arctan\Delta x$	Δx	$\dfrac{1}{3}(\Delta x)^3$, $-\frac{1}{3}(\Delta x)^3$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

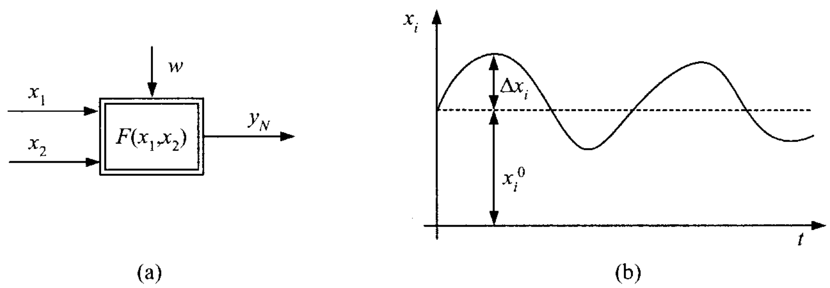


Figure 3.3: Nonlinear system with three inputs (excitations) and one output (response).

3.3 Analytical Linearization Method (Linearization in the Vicinity of the Operating Point)

The analytical linearization method in the vicinity of the operating point is based on the substitution of the nonlinear system by a linearized one. Linearization is carried out by expanding the nonlinear function into a Taylor series at the nominal operating point, i.e. in the static operation mode. The method is applied to smooth nonlinearities with small deflections from the operating point. This condition is fulfilled in stabilizing closed-loop control systems or control systems which solve the regulator problem. Fig. 3.3 presents a nonlinear system with two inputs (x_1,x_2), one output (y_N) and one disturbance (w).

The dynamic behavior of the system in Fig. 3.3 is described by the differential equation:

$$F(x_1,x_2,\dot{x}_2,y_N,\dot{y}_N,\ddot{y}_N,\dddot{y}_N) = \varphi(w,\dot{w}) \tag{3.7}$$

where F and φ are nonlinear functions.

With static operating mode at the operating point $x_1 = x_1^0$, $x_2^0 = x_2^0$, $y_N = y_N^0$ and disturbance $w = w^0$, we obtain:

$$F(x_1^0,x_2^0,0,y_N^0,0,0,0) = \varphi(w^0,0) \tag{3.8}$$

Assuming that the excitation variables x_1, x_2 and response y_N can be expanded in Taylor series, and the oscillations Δx_1, Δx_2 and Δy_N are sufficiently small (Fig. 3.3b), the system dynamics can be reasonably well described by the linear approximation:

$$\begin{aligned} &x_1(t) = x_1^0 + \Delta x_1(t),\ x_2(t) = x_2^0 + \Delta x_2(t),\ \dot{x}_2 = \Delta\dot{x}_2 \\ &y_N(t) = y_N^0 + \Delta y_N(t),\ \dot{y}_N = \Delta\dot{y}_N,\ \ddot{y}_N = \Delta\ddot{y}_N,\ \dddot{y}_N = \Delta\dddot{y}_N \end{aligned} \tag{3.9}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The disturbance variable $w(t)$ is independent of the system itself. It can be arbitrary and may not be linearized.

By expanding the function (3.7) in a Taylor series at the operating point (3.8) the dynamic equation of the system obtains the following form:

$$F(x_1^0, x_2^0, 0, y_N^0, 0, 0, 0) + \left(\frac{\partial F}{\partial x_1}\right)^0 \Delta x_1 + \left(\frac{\partial F}{\partial x_2}\right)^0 \Delta x_2 + \left(\frac{\partial F}{\partial \dot{x}_2}\right)^0 \Delta \dot{x}_2$$
$$+ \left(\frac{\partial F}{\partial y_N}\right)^0 \Delta y_N + \left(\frac{\partial F}{\partial \dot{y}_N}\right)^0 \Delta \dot{y}_N + \left(\frac{\partial F}{\partial \ddot{y}_N}\right)^0 \Delta \ddot{y}_N + \left(\frac{\partial F}{\partial \dddot{y}_N}\right)^0 \Delta \dddot{y}_N$$
$$+ \ldots + \text{higher order terms} = \varphi(w, \dot{w}) \tag{3.10}$$

By neglecting higher order terms in (3.10), the linearized equation of the system is written:

$$\left(\frac{\partial F}{\partial x_1}\right)^0 \Delta x_1 + \left(\frac{\partial F}{\partial x_2}\right)^0 \Delta x_2 + \left(\frac{\partial F}{\partial \dot{x}_2}\right)^0 \Delta \dot{x}_2 + \left(\frac{\partial F}{\partial y_N}\right)^0 \Delta y_N$$
$$+ \left(\frac{\partial F}{\partial \dot{y}_N}\right)^0 \Delta \dot{y}_N + \left(\frac{\partial F}{\partial \ddot{y}_N}\right)^0 \Delta \ddot{y}_N + \left(\frac{\partial F}{\partial \dddot{y}_N}\right)^0 \Delta \dddot{y}_N = \varphi(w, \dot{w}) - \varphi(w^0, 0) \tag{3.11}$$

Linearized equations (3.11) and nonlinear equations (3.7) describe the dynamics of the system in Fig. 3.3. The differences between the equations (3.11) and (3.7) are following:

- Equation (3.11) is an approximation, since higher-order terms are neglected.
- Variables in (3.11) are excitation deflections Δx_1, Δx_2 and response deflection Δy_N from their respective operating points x_1^0, x_2^0, y_N^0.
- Equation (3.11) is a linear differential equation in relation to deflections $\Delta x_1, \Delta x_2, \Delta y_N, \Delta \dot{x}_2, \ldots, \Delta \dddot{y}_N$ with constant coefficients $\left(\frac{\partial F}{\partial x_1}\right)^0, \left(\frac{\partial F}{\partial x_2}\right)^0, \ldots$ These coefficients are variable coefficients when the nonlinear function F is also a function of the argument t, or when the operating point is defined by variable quantities $x_1^0(t)$, $x_2^0(t)$, $y_N^0(t)$, as is the case in programmable control.

Analytical linearization method by expansion in Taylor series is equivalent to the graphical linearization. Figure 3.4 presents a graphical dependence of $F(x_1)$ with fixed values of remaining variables $x_2 = x_2^0$, $\dot{x}_2 = 0$, $y_N = y_N^0$, $\dot{y}_N = \ddot{y}_N = \dddot{y}_N = 0$.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

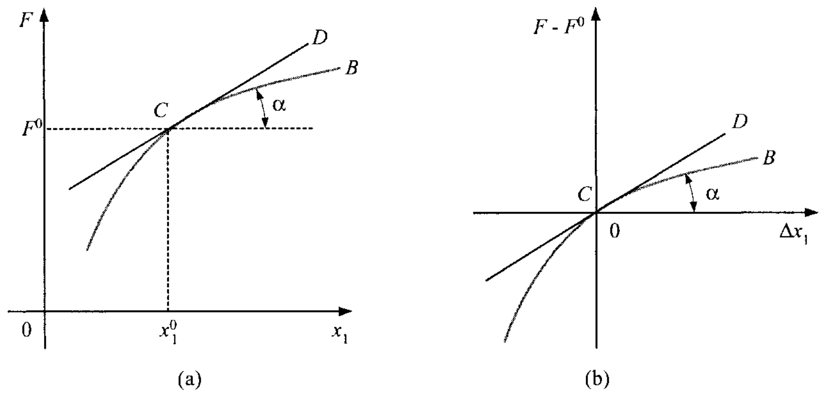


Figure 3.4: Graphical dependence $F(x_1)$ with other variables constant.

Nonlinear function $F - F^0 = F(x_1^0, x_2^0, 0, y_N^0, 0, 0, 0)$ at the operating point $C(x_1^0, F^0)$ is approximated by the tangent CD with the slope:

$$K_1 = \frac{\partial F}{\partial x_1} = \tan\alpha$$

which is equal to the first term of equation (3.11). In Fig. 3.4b the linearization is shown by substituting $x_1 = x_1^0 + \Delta x_1$ and by choosing the point C as the origin of coordinates. From the above discussion follows the conclusion that by analytic linearization method with Taylor-series expansion the nonlinear function $y_N = F(x_1, ..., x_n)$ described by a plane in n-dimensional space[5] is replaced by a tangential plane at the operating point, i.e. a linear relationship of small deflections Δy_N and Δx_i with the coefficients of proportionality (linearization) which are equal to partial derivatives of the nonlinear function at the operating point (Fig. 3.5b).

The equation for the system in Fig. 3.5a can be written as a scalar equation with vector argument:

$$y_N = F(\mathbf{x}) \tag{3.12}$$

and the linearized equation for the system in Fig. 3.5b in the form of a scalar product:

$$\Delta y_N \approx \mathbf{grad}\, F|_C \Delta\mathbf{x} = \nabla F|_C \Delta\mathbf{x} = \left.\frac{dF}{d\mathbf{x}}\right|_C \Delta\mathbf{x} \tag{3.13}$$

[5]Index n represents the number of inputs to the nonlinear element and may not be mixed up with the dimension of the state vector.

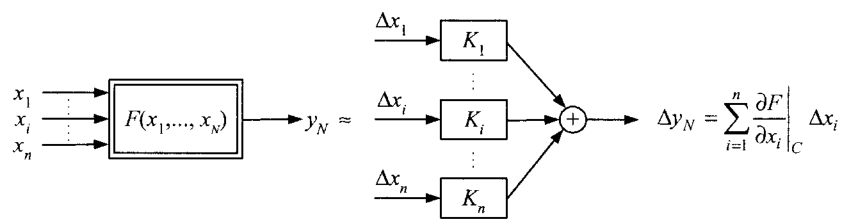


Figure 3.5: Substitution of nonlinear function by a linear one, where deflections Δy_N and Δx_i are correlated with linearization coefficients K_i.

or in the matrix form:

$$\Delta y_N \approx [\mathbf{grad}\, F|_C]^T \Delta\mathbf{x} = [\nabla F|_C]^T \Delta\mathbf{x} = \left[\frac{dF}{d\mathbf{x}}\bigg|_C \right]^T \Delta\mathbf{x} = \frac{dF}{d\mathbf{x}^T}\bigg|_C \Delta\mathbf{x} \tag{3.14}$$

where **grad** F is the gradient (column vector) of the scalar function F, ∇ is column vector of operator nabla ($\nabla = [\partial/\partial x_1, ..., \partial/\partial x_n]^T$) and ^{T} means transposition (from a column to a row vector and vice versa).

The next example will illustrate this:

$$\mathbf{grad}\, F = \nabla F = \frac{dF}{d\mathbf{x}} = \begin{bmatrix} \frac{\partial F}{\partial x_1} \\ \vdots \\ \frac{\partial F}{\partial x_n} \end{bmatrix} = \left[\frac{\partial F}{\partial x_1} \cdots \frac{\partial F}{\partial x_n} \right]^T$$

$$\mathbf{grad}^T\, F = [\nabla F]^T = \left[\frac{dF}{d\mathbf{x}} \right]^T = \frac{dF}{d\mathbf{x}^T} = \left[\frac{\partial F}{\partial x_1} \cdots \frac{\partial F}{\partial x_n} \right]$$

Whether a column or a row vector will be used must be known from the begining, since the dimensions must be proper. The choice itself doesn't change the result, but it is important that the choice must not change the dimension of the output. In this particular case the output is a scalar quantity.

Equation (3.14) can therefore be written in the form:

$$\Delta y_N \approx [\Delta\mathbf{x}]^T\, \mathbf{grad}\, F|_C = [\Delta\mathbf{x}]^T\, \nabla F|_C = [\Delta\mathbf{x}]^T \frac{dF}{d\mathbf{x}}\bigg|_C \tag{3.15}$$

The equations (3.14) and (3.15) give the same result (scalar Δy) if the right choice of the vector form (column or row vector) is taken.

For multivariable nonlinear systems the mathematical model is:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}[t, \mathbf{x}(t), \mathbf{u}(t)] \\ \mathbf{y}(t) &= \mathbf{h}[t, \mathbf{x}(t), \mathbf{u}(t)] \end{aligned} \tag{3.16}$$

where
$\mathbf{u}(t) = [u_1\ u_2\ \ldots\ u_m]^T$ is a vector of m control signals to the nonlinear system,
$\mathbf{y}(t) = [y_1\ y_2\ \ldots y_r]^T$ is a vector of r output signals from the nonlinear system,
$\mathbf{f}(t,\mathbf{x},\mathbf{u}) = \begin{bmatrix} f_1(t,x_1,x_2,\ldots,x_n,u_1,\ldots,u_m) \\ f_2(t,x_1,x_2,\ldots,x_n,u_1,\ldots,u_m) \\ \vdots \\ f_n(t,x_1,x_2,\ldots,x_n,u_1,\ldots,u_m) \end{bmatrix}$ is a vector of nonlinear functions of every state variable component and
$\mathbf{h}(t,\mathbf{x},\mathbf{u}) = \begin{bmatrix} h_1(t,x_1,x_2,\ldots,x_n,u_1,\ldots,u_m) \\ h_2(t,x_1,x_2,\ldots,x_n,u_1,\ldots,u_m) \\ \vdots \\ h_r(t,x_1,x_2,\ldots,x_n,u_1,\ldots,u_m) \end{bmatrix}$ is a vector of nonlinear functions of every output signal component.

In case of a time-invariant nonlinear system where $\mathbf{f}(\mathbf{x},\mathbf{u})$ is continuously differentiable at the equilibrium[6] $(\mathbf{x}_e,\mathbf{u}_e)$, then the Jacobian matrices $\mathbf{J}_x$ and $\mathbf{J}_u$ can be used to approximate $\mathbf{f}(\mathbf{x},\mathbf{u})$. A multivariable Taylor series expansion has the form:

$$\begin{aligned}\frac{d}{dt}\mathbf{x}(t) = \mathbf{f}(\mathbf{x},\mathbf{u}) &= \mathbf{f}(\mathbf{x}_e,\mathbf{u}_e) + \mathbf{D}_1\mathbf{f}(\mathbf{x}_e,\mathbf{u}_e)\,[\mathbf{x}(t)-\mathbf{x}_e] \\ &\quad + \mathbf{D}_2\mathbf{f}(\mathbf{x}_e,\mathbf{u}_e)\,[\mathbf{u}(t)-\mathbf{u}_e] + \mathbf{r}\,[\mathbf{x}(t),\mathbf{u}(t)]\end{aligned} \tag{3.17}$$

where:
$\mathbf{J}_x = \mathbf{D}_1\mathbf{f}(\mathbf{x}_e,\mathbf{u}_e) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$ is a Jacobian $n \times n$ matrix,
$\mathbf{J}_u = \mathbf{D}_2\mathbf{f}(\mathbf{x}_e,\mathbf{u}_e) = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial u_1} & \cdots & \frac{\partial f_n}{\partial u_m} \end{bmatrix}$ is a Jacobian $n \times m$ matrix,
$\mathbf{r}\,[\mathbf{x}(t),\mathbf{u}(t)]$ is the remainder (higher order terms) of the Taylor series expansion.

If the remainder $\mathbf{r}(\mathbf{x},\mathbf{u})$ satisfies:

$$\lim_{(\mathbf{x},\mathbf{u})\to(\mathbf{x}_e,\mathbf{u}_e)} = \frac{\mathbf{r}(\mathbf{x},\mathbf{u})}{\sqrt{|\mathbf{x}-\mathbf{x}_e|^2 + |\mathbf{u}-\mathbf{u}_e|^2}} = \mathbf{0}$$

the approximation will be accurate up to first order.

Defining the deviation from the equilibrium as $\Delta\mathbf{x}(t) = \mathbf{x}(t) - \mathbf{x}_e$ and $\Delta\mathbf{u}(t) = \mathbf{u}(t) - \mathbf{u}_e$ and assuming fixed equilibriums, i.e. $\frac{d}{dt}\mathbf{x}_e = \mathbf{0}$ with $\mathbf{f}(\mathbf{x}_e,\mathbf{u}_e) = \mathbf{0}$, we

[6]The pair $(\mathbf{x}_e,\mathbf{u}_e)$ is called equilibrium if $\mathbf{f}(\mathbf{x}_e,\mathbf{u}_e) = \mathbf{0}$.

will have:

$$\Delta\dot{\mathbf{x}}(t) \approx \mathbf{J}_x\Delta\mathbf{x}(t) + \mathbf{J}_u\Delta\mathbf{u}(t) \tag{3.18}$$

Equation 3.18 represents the linearized dynamics of the nonlinear system about the equilibrium point $(\mathbf{x}_e, \mathbf{u}_e)$.

Linearization of the algebraic output equation can be done in a similar way, i.e. the following Taylor series expansion can be obtained for the time-invariant nonlinear system:

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x},\mathbf{u}) = \mathbf{h}(\mathbf{x}_e,\mathbf{u}_e) + \mathbf{K}[\mathbf{x}(t) - \mathbf{x}_e] + \mathbf{r}[\mathbf{x}(t),\mathbf{u}(t)]$$

where $\mathbf{K} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_r}{\partial x_1} & \cdots & \frac{\partial h_r}{\partial x_n} \end{bmatrix}$ is a Jacobian $r \times n$ matrix,

$\mathbf{r}[\mathbf{x}(t),\mathbf{u}(t)]$ is the remainder (higher order terms) of the Taylor series expansion.

For the nonlinear unforced[7] system, if the function $\mathbf{h}(\mathbf{x})$ is continuously differentiable at $\mathbf{x}_e$, the Taylor series expansion takes the form:

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}) = \mathbf{h}(\mathbf{x}_e) + \mathbf{K}\Delta\mathbf{x}(t) + \mathbf{r}[\mathbf{x}(t)] \tag{3.19}$$

If $\mathbf{h}(\mathbf{x}_e) = \mathbf{0}$ and also if the remainder satisfies:

$$\lim_{\mathbf{x}\to\mathbf{x}_e} = \frac{\mathbf{r}(\mathbf{x})}{\sqrt{|\mathbf{x}-\mathbf{x}_e|^2}} = \lim_{\Delta\to\mathbf{0}} = \frac{|\mathbf{r}(\mathbf{x}_e+\Delta)|}{|\Delta|} = \mathbf{0}$$

the approximation will be accurate up to first order, and can be given by:

$$\Delta\mathbf{y}(t) \approx \mathbf{K}\Delta\mathbf{x}(t) \tag{3.20}$$

where $\mathbf{K}$ is a Jacobian $r \times n$ matrix defined as:

$$\mathbf{K} = \mathbf{h}(\mathbf{x})\nabla^T = \begin{bmatrix} h_1(x_1,x_2,\ldots,x_n) \\ h_2(x_1,x_2,\ldots,x_n) \\ \vdots \\ h_r(x_1,x_2,\ldots,x_n) \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \cdots & \frac{\partial}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_r}{\partial x_1} & \cdots & \frac{\partial h_r}{\partial x_n} \end{bmatrix}$$

For $r = n$ the Jacobian matrix $\mathbf{K}$ becomes a square $n \times n$ matrix.

[7]Vector of control signals is a null vector.

3.4 Evaluation of Linearization Coefficients by Least-Squares Method

The method is essentially analog to the graphical approximation of a nonlinear static characteristic of one variable $y_N = F(x)$ by a secant line (Fig. 3.2). For nonlinear functions with several variables, a phase plane $y_N = F(x_1, \ldots, x_n)$ is approximated by a plane which cuts the phase plane in a prescribed number of points in the vicinity of the operating point.

To evaluate the linearization coefficients, it is appropriate to apply the least-squares method. The nonlinear function is described by the equation:

$$y_N = F(x_1, x_2, \ldots, x_n) \tag{3.21}$$

and the operating point is:

$$y_N^0 = F^0(x_1^0, x_2^0, \ldots, x_n^0) \tag{3.22}$$

In the vicinity of the operating point (3.22), m points for each variable x_i are chosen so that the condition $m > n$ is fulfilled. The values of the variable y_N at these points are:

$$y_{Nk} = F(x_{1k}, x_{2k}, \ldots, x_{nk});\ k = 1, 2, \ldots, m \tag{3.23}$$

The distances between the chosen points and the operating point are given by the equations:

$$\left.\begin{array}{c} x_{1k} - x_1^0 = \Delta x_{1k} \\ \vdots \\ x_{nk} - x_n^0 = \Delta x_{nk} \end{array}\right\} \quad (k = 1, 2, \ldots, m) \tag{3.24}$$

and

$$y_{Nk} - y_N^0 = \Delta y_{Nk}\ ;\ k = 1, 2, \ldots, m \tag{3.25}$$

From (3.24) and (3.25) the linearized model follows:

$$\Delta y_N \approx \sum_{i=1}^{n} K_i \Delta x_i = K_1 \Delta x_1 + \ldots + K_n \Delta x_n \tag{3.26}$$

Combining equations (3.4) and (3.26) gives the sum of squares of deflections for the chosen points:

$$E = \sum_{k=1}^{m} (K_1 \Delta x_{1k} + K_2 \Delta x_{2k} + \ldots + K_n \Delta x_{nk} - \Delta y_{Nk})^2 \tag{3.27}$$

The minimum of the sum of squares for E is obtained when the first partial derivatives are equal to zero:

$$\frac{\partial E}{\partial K_1} = \frac{\partial E}{\partial K_2} = \ldots = \frac{\partial E}{\partial K_n} = 0 \tag{3.28}$$

$$\frac{\partial E}{\partial K_i} = \sum_{k=1}^{m} 2(K_1 \Delta x_{1k} + K_2 \Delta x_{2k} + \ldots + K_n \Delta x_{nk} - \Delta y_{Nk})\Delta x_{ik} = 0 \quad (i = 1, 2, \ldots, n) \tag{3.29}$$

and further:

$$K_1 \sum_{k=1}^{m} \Delta x_{1k} \Delta x_{ik} + K_2 \sum_{k=1}^{m} \Delta x_{2k} \Delta x_{ik} + \ldots + K_n \sum_{k=1}^{m} \Delta x_{nk} \Delta x_{ik} = \sum_{k=1}^{m} \Delta y_{Nk} \Delta x_{ik} \quad (i = 1, 2, \ldots, n) \tag{3.30}$$

Coefficients of linearization K_i are defined by the solution of the system of n equations (3.30). For a function of one variable $y_N = F(x)$, the equation of the linearized model is:

$$\Delta y_N \approx K \cdot \Delta x \tag{3.31}$$

where the coefficient of linearization or equivalent gain is:

$$K = \frac{\sum_{k=1}^{m} \Delta y_{Nk} \Delta x_k}{\sum_{k=1}^{m} (\Delta x_k)^2}, \; m > 1 \tag{3.32}$$

The least-squares method is applied in cases when approximations by simpler methods yield no satisfying results. In the case when the nonlinear function cannot be represented analytically, the sums in the expression (3.32) can be replaced by integrals. In such a way, for the function of one variable $y_N = F(x)$, the coefficient of linearization by the least-squares method is:

$$K = \frac{\int_{-\Delta x}^{\Delta x} \Delta F(\Delta x) \Delta x d(\Delta x)}{\int_{-\Delta x}^{\Delta x} (\Delta x)^2 d(\Delta x)} \tag{3.33}$$

Although the least-squares method is better than the simpler tangent or secant methods, it is seldom used in practice because its realization is difficult, especially in cases of more complex nonlinear functions.

EXAMPLE 3.1
(COMPARISON OF LEAST-SQUARES AND TANGENT METHODS OF LINEARIZATION)

The difference between these two methods is illustrated in a simple example of the nonlinear function $y_N = F(x) = x^n$ in the vicinity of the operating point[8] $x^0 = X \neq 0$. Approximation by the tangent yields (see equation (3.13)):

$$\frac{\partial F}{\partial x}\Big|_{x=x^0} = nx^{n-1}\Big|_{x=x^0} = nX^{n-1}$$

and the linearized model is:

$$\Delta y_N = nX^{n-1}\Delta x$$

According to equation (3.24), the least-squares method yields:

$$y_N^0 + \Delta y_N = (x^0 + \Delta x)^n$$
$$Y + \Delta y = (X + \Delta x)^n$$

respectively:

$$\Delta y_N = (X + \Delta x)^n - X^n \tag{3.34}$$

Using equation (3.33), the coefficient of linearization is:

$$K = \frac{\int\limits_{-\Delta x}^{\Delta x} [(X + \Delta x)^n - X^n]\,\Delta x d(\Delta x)}{\int\limits_{-\Delta x}^{\Delta x} (\Delta x)^2 d(\Delta x)}$$

By integrating:

$$K = \frac{\sum\limits_{h=1}^{[(n+1)/2]} \frac{2}{2h+1}\binom{n}{2h-1} X^{n-2h+1}(\Delta x)^{2h+1}}{\frac{2}{3}(\Delta x)^3}$$

where the brackets signify an integer. For $n = 1$, we will have $K = 1$, while for $n = 2$, K will take the value:

$$K = \frac{\frac{4}{3}X(\Delta x)^3}{\frac{2}{3}(\Delta x)^3} = 2X$$

For $n = 1$ and $n = 2$, the tangent method ($\Delta y_N = K\Delta x = nX^{n-1}\Delta x$) will give $K = 1$ and $K = 2$, respectively. As may be seen, both methods give the same result for

[8] Instead of the operating point notations x^0 and y_N^0 in this example X and Y will be used because of simpler notation.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

$n = 1$ and $n = 2$, whereas for $n > 2$ the coefficients of linearization differ. For example, for $n = 3$, the least-squares method yields:

$$K = \frac{2X^2(\Delta x)^3 + \frac{2}{5}(\Delta x)^5}{\frac{2}{3}(\Delta x)^3} = 3X^2 + \frac{3}{5}(\Delta x)^2$$

It is evident that the coefficient of linearization obtained by the least-squares method depends on X and Δx, and differs from the coefficient of linearization obtained by the tangent method: $K = 3X^2$. The same is true for $n = 4$. With the least-squares method:

$$K = \frac{\frac{8}{3}X^3(\Delta x)^3 + \frac{8}{5}X(\Delta x)^5}{\frac{2}{3}(\Delta x)^3} = 4X^3 + \frac{12}{5}X(\Delta x)^2$$

while the tangent method yields $K = 4X^3$. It can be shown that the coefficients of linearization will differ for all $n > 2$.

3.5 Harmonic Linearization

In technical calculations, because of the complexity of an exact description of all phenomena, the system is always replaced by an idealized mathematical model which may contain all the basic properties of the system and will enable as simple as possible methodology to be used for defining the dynamic behavior. Since the linear models are best researched, every real system ought to be replaced by an equivalent linear system. The analytical approximations by tangent or secant methods in the vicinity of the operating point cannot be used in the cases when the input (excitation) variable instantaneously changes its value from $x = 0$ to $x = X$, as shown in Fig. 3.6a,b.

In Fig. 3.6 straight lines are shown $y = Kx$, $y_N = K_{eq}x$ and $y_N = K_m x$ by which linearization of nonlinear static characteristic $y_N = F(x)$ is possible. Linarization by the first straight line $y_N = Kx$ is justified in all situations when the change of the input (excitation) variable x in the vicinity of the operating point (i.e. the point "0" at the coordinate origin) is small. If the change is rapid and large from $x = 0$ to $x = X$, the point C on the curve $y_N = F(x)$ (Fig. 3.6a) is defined by the equation of the straight line:

$$y_N = K_m x;\ K_m = \frac{F(x)}{X} \tag{3.35}$$

For an arbitrary change of the excitation from $x(t) = 0$ to $x(t) = X$, the response $y_N = F(x)$ will be closer to the linear system with equivalent coefficient of proportionality:

$$K_m < K_{eq} < K \text{ or } K < K_{eq} < K_m \tag{3.36}$$

which depends on the form of the nonlinear characteristic, Fig. 3.6a and b.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

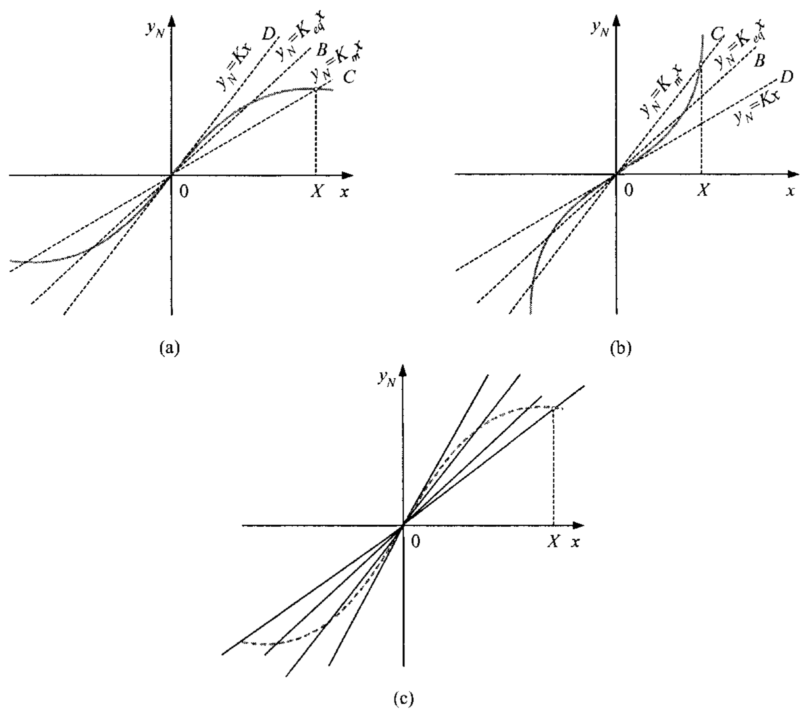


Figure 3.6: Linearization of static characteristics with straight lines.

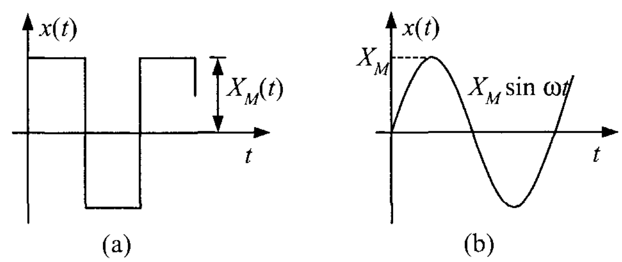


Figure 3.7: Types of input signals to a nonlinear element $y_N = ax + bx^3$.

The value of this coefficient for every point $C(x, y_N)$ of the characteristic $y_N = F(x)$ and for the known laws of change of $x(t)$ can be evaluated if the condition of equivalence is defined, such as equal areas under the curves $y_N = F(x)$ and $y_N = K_{eq}x$, or the energy equality, etc. For a nonlinear function, e.g.

$$y_N = ax + bx^3 \tag{3.37}$$

if the excitation to nonlinear element $x(t)$ has the form of the square wave periodic function (Fig. 3.7a), the response $y_N(t) = F(x)$ will have the form of square wave oscillations $F(X_m)$. The equivalent gain at the particular point $C(x, y_N)$ of the curve $y_N = F(x)$ results from the equation (3.35):

$$K_{eq} = \frac{F(X_m)}{X_m} = a + bX_m^2 \tag{3.38}$$

and the equivalent linearization (known as the rectangular linearization) of the nonlinearity (3.37) has the form:

$$y = K_{eq}x = (a + bX_m^2)x \tag{3.39}$$

From the expression (3.39) it is evident that the equivalent coefficient of linearization depends exclusively on the amplitude of the square wave periodic input function.

When the excitation $x(t)$ is a sinusoid (Fig. 3.7b):

$$x(t) = X_m \sin \omega t \tag{3.40}$$

the response $y_N = F(x) = ax + bx^3$ will take the form of a complex periodic function. In the case when $y = F(x)$ is a linear function, the response will be a sinusoidal one with the same frequency as the input signal:

$$y = X_l \sin \omega t \tag{3.41}$$

If for the condition of equivalence the equivalent gain is chosen:

$$K_{eq} = \frac{X_I}{X_m} \tag{3.42}$$

oscillations of the response y_N will be equal to the first harmonic of complex nonlinear oscillations of the variable $y_N = F(X_m \sin \omega t)$. The first harmonic of this function is defined by Fourier series expansion. For the odd functions in Fig. 3.6, only the sinusoidal terms exist, and the first harmonic is:

$$y_I = X_I \sin \omega t \tag{3.43}$$

where:

$$X_I = \frac{1}{\pi} \int_0^{2\pi} F(X_m \sin \varphi) \sin \varphi d\varphi \quad (\varphi = \omega t) \tag{3.44}$$

From the condition of equivalent linearization (3.42) and (3.44), the equivalent gain can be written down:

$$K_{eq} = \frac{X_I}{X_m} = \frac{1}{\pi X_m} \int_0^{2\pi} F(X_m \sin \varphi) \sin \varphi d\varphi \tag{3.45}$$

For the nonlinear characteristic given in Fig. 3.6b, described by equation $y_N(t) = ax + bx^3$, where $x(t) = X_m sin\omega t$, the equivalent gain is according to (3.45) given by:

$$K_{eq} = \frac{a}{\pi X_m} \int_0^{2\pi} X_m \sin^2 \varphi d\varphi + \frac{b}{\pi X_m} \int_0^{2\pi} X_m^3 \sin^4 \varphi d\varphi \tag{3.46}$$

or:

$$K_{eq} = a + \frac{3}{4} b X_m^2 = \frac{X_I}{X_m} \tag{3.47}$$

According to (3.43), the first harmonic of complex oscillations at the output of the nonlinear element $y_N(t) = F(x)$ for $x(t) = X_m sin\omega t$ is:

$$y_I = X_I \sin \omega t = K_{eq} X_m \sin \omega t = \left(a + \frac{3}{4} b X_m^2\right) X_m \sin \omega t \tag{3.48}$$

The form of linearization (3.45) by which for excitation $x(t) = X_m sin\omega t$ the equality of the first harmonic of nonlinear oscillations $y_N(t) = F(x)$ with the amplitude of oscillations of equivalent linear system $y_I(t) = X_I sin\omega t$ is achieved is called harmonic linearization. The equation of harmonic linearization (3.44) approximates the equation of nonlinearity $y_N(t) = F(x)$ with the actuating excitation $x(t)$.

As in the case of rectangular linearization (3.39), with harmonic linearization the equivalent gain depends upon the amplitude of oscillations of the excitation

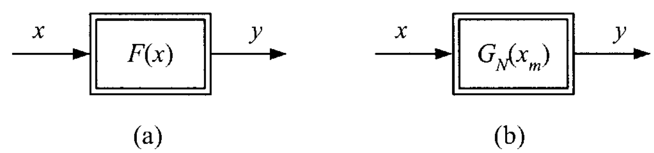


Figure 3.8: Nonlinear block (a) and block with describing function (b).

$x(t)$. For different values of the amplitude X_m the equivalent gain will have corresponding constant values.

In contrast to graphical and analytical methods of linearization of a nonlinear function in the vicinity of the operating point (linearization in time domain), harmonic linearization (linearization in frequency domain) allows a nonlinear characteristic $y_N(t) = F(x)$ to be substituted not with one line but with a "bundle of line", the slopes of which depend on the amplitude of oscillations of the variable $x(t)$, i.e. upon the "operating interval" of $y_N(t) = F(x)$ spread by oscillations $x(t) = X_m sin\omega t$ (Fig. 3.6c).

By analogy, depending on the changes of the excitation $x(t)$ and on the equivalence conditions, other forms of linearization can be applied which will result in corresponding equivalent gains of the particular nonlinearity $y_N = F(x)$.

3.6 Describing Function

The term describing function refers to the equivalent gain or low-frequency gain defined by the harmonic linearization method of a nonlinear static characteristic. Application of the describing function determines the periodic operating modes of the nonlinear system with $y_N = F(x)$ and $y_N = F(x, \dot{x})$ which most often appear in practice (Netushil, 1983; Goldfarb, 1965; Popov and Pal'tov, 1960; Gelb and Vander Velde, 1968; and others).

Generally if the input to a nonlinear element, Fig. 3.8, is a harmonic function

$$x(t) = X_m \sin \omega t \tag{3.49}$$

the periodic process at the output of a nonlinear element $y_N(t) = F[x(t)]$ is given by the Fourier series:

$$y_N(t) = Y_0 + \sum_{k=1}^{\infty} Y_{Pk} \sin k\omega t + \sum_{k=1}^{\infty} Y_{Qk} \cos k\omega t \tag{3.50}$$

where the coefficients Y_{Pk} and Y_{Qk} are:

$$Y_{Pk} = \frac{1}{\pi}\int_0^{2\pi} F(X_m \sin\omega t)\sin k\omega t \; d(\omega t) \tag{3.51}$$

$$Y_{Qk} = \frac{1}{\pi}\int_0^{2\pi} F(X_m \sin\omega t)\cos k\omega t \; d(\omega t) \tag{3.52}$$

Generally these coefficients depend on the amplitude of the input signal X_m as well as on its frequency ω, i.e. $Y_{Pk} = Y_{Pk}(X_m, j\omega)$, $Y_{Qk} = Y_{Qk}(X_m, j\omega)$. For unforced nonlinear elements with symmetrical nonlinear characteristics, periodic oscillations of $y_N(t)$ will be symmetrical in relation to the time axis, i.e. $Y_0 = 0$. By applying harmonic linearization, a periodic signal $y_N(t)$ is approximated by its first harmonic:

$$y(t) \approx Y_{P1}\sin\omega t + Y_{Q1}\cos\omega t \tag{3.53}$$

or equivalently:

$$y(t) \approx \operatorname{Im}\left\{(Y_{P1} + jY_{Q1})e^{j\omega t}\right\} \tag{3.54}$$

where:

$$Y_{P1} = \frac{1}{\pi}\int_0^{2\pi} F(X_m \sin\omega t)\sin\omega t \; d(\omega t) \tag{3.55}$$

$$Y_{Q1} = \frac{1}{\pi}\int_0^{2\pi} F(X_m \sin\omega t)\cos\omega t \; d(\omega t) \tag{3.56}$$

The describing function (complex gain) of a nonlinear element is defined as the ratio between the first harmonic of output and input signals expressed in complex form:

$$G_N(X_m) = P(X_m) + jQ(X_m) = \frac{Y_{P1}}{X_m} + j\frac{Y_{Q1}}{X_m} = |G_N(X_m)|\, e^{j\varphi_N} \tag{3.57}$$

where:

$$P(X_m) = \frac{Y_{P_1}}{X_m} = \frac{1}{\pi X_m}\int_0^{2\pi} F(X_m \sin\omega t)\sin\omega t \; d(\omega t) \tag{3.58}$$

$$Q(X_m) = \frac{Y_{Q_1}}{X_m} = \frac{1}{\pi X_m}\int_0^{2\pi} F(X_m \sin\omega t)\cos\omega t \; d(\omega t) \tag{3.59}$$

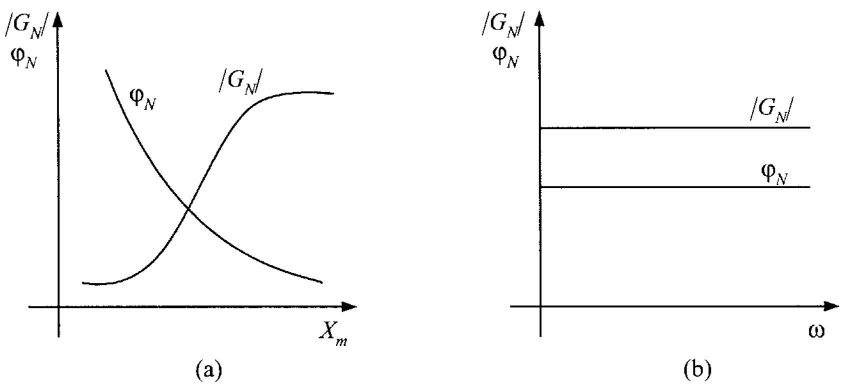


Figure 3.9: Dependence of amplitude and phase of describing function G_N of nonlinear element without inertia on input amplitude (a) and the input signal frequency (b).

$$|G_N(X_m)| = \frac{Y_{m_1}}{X_m} = \frac{\sqrt{Y_{P1}^2 + Y_{Q1}^2}}{X_m}; \ \varphi_N = \arctan\frac{Y_{Q1}}{Y_{P1}} \tag{3.60}$$

From (3.60) it follows that the describing function of an inertialess nonlinear element depends on the amplitude X_m of the input signal, and not on the input signal frequency, Fig. 3.9. The complex gain components of the describing function Y_{P1} and Y_{Q1} are called *harmonic linearization coefficients*.

The harmonic linearization coefficient $Q(X_m)$, i.e. the imaginary part of the describing function $G_N(X_m)$, exists only for multi-valued nonlinearities. Only in the case of multi-valued nonlinearity does a harmonic signal shift in phase while propagating through a nonlinear element. Equation (3.58) yields for multi-valued symmetrical nonlinearity:

$$\begin{aligned} Q(X_m) &= \frac{Y_{Q1}}{X_m} = \frac{1}{\pi X_m}\int_0^{2\pi} F(X_m \sin\omega t)\cos\omega t d(\omega t) \\ &= \frac{1}{\pi X_m^2}\oint F(X_m \sin\omega t) d(X_m \sin\omega t) = \frac{1}{\pi X_m^2}\oint F(x)dx = -\frac{1}{\pi X_m^2} S \end{aligned} \tag{3.61}$$

The contour integral in (3.61) is solved by integrating along the contour which the multi-valued nonlinear characteristic closes when x goes counterclockwise from $x = 0$ (at $t = 0$) to $x = X_m$ (at $t = \frac{\pi}{2\omega}$) and then from $x = X_m$ to $x = -X_m$ (at $t = \frac{3\pi}{2\omega}$) and finally from $x = -X_m$ to $x = 0$ (at $t = \frac{2\pi}{\omega}$)—see Fig. 3.10. The result of

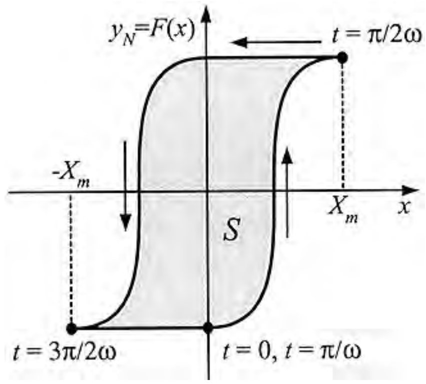


Figure 3.10: Multi-valued nonlinear static characteristic (hysteresis type).

integration is equal to the area S of the hysteresis loop with the negative sign[9].

$$S = \int_{0} F(x)dx + \int_{-X_m}^{X_m} F(x)dx + \int_{-X_m}^{0} F(x)dx \tag{3.62}$$

For single-valued symmetrical nonlinearities $S = 0$, i.e. there is no phase shift of a harmonic signal when passing through an inertialess single-valued nonlinear element.

Based on equation (3.60) and Fig. 3.8b, the nonlinearity $y_N = F(x)$ is replaced with the describing function $G_N(X_m)$:

$$y = G_N(X_m)x \tag{3.63}$$

or:

$$y(t) = Y_{P1}\sin\omega t + Y_{Q1}\cos\omega t = |G_N(X_m)|\sin(\omega t + \varphi_N) \tag{3.64}$$

Certain nonlinear elements can be replaced by a describing function which depends solely on the amplitude of the input signal. Nonlinearities where the output depends only on input amplitude are called *simple nonlinear elements* or *nonlinear elements of zero order*. Simple nonlinearities comprise typical (standard) nonlinear elements.

With complex nonlinearities, contrary to simple nonlinearities, the derivatives of the input and output coordinates are included, and the describing function is dependent not only on the input amplitude, but also on the input signal frequency, i.e. $G_N = G_N(X_m, \omega)$. Describing functions for most relay characteristics are

[9]The sign depends on the direction of scanning the multi-valued curve.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

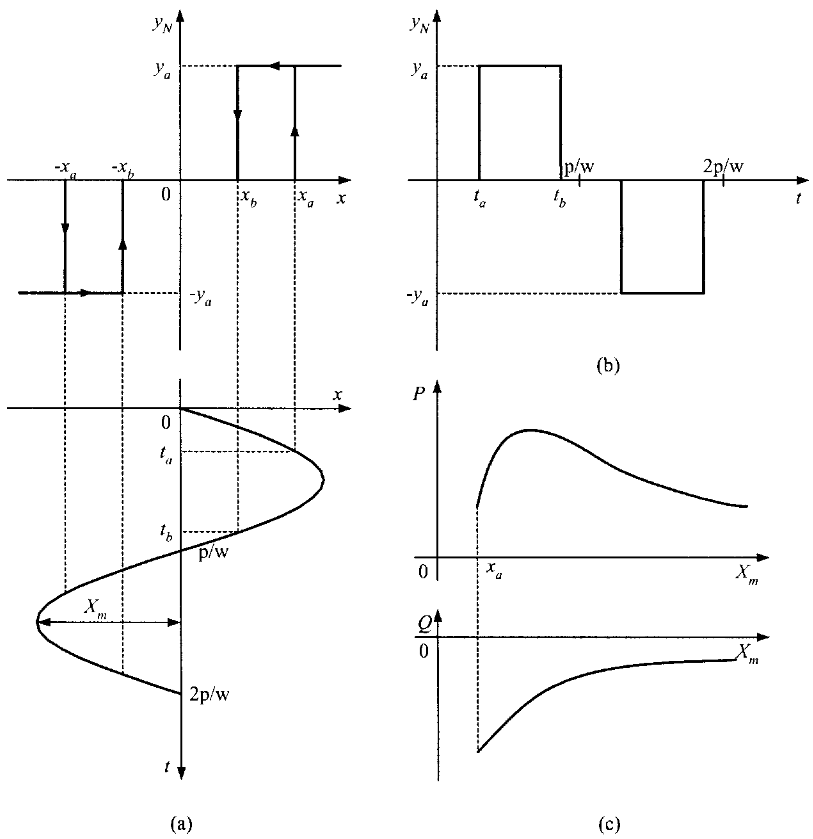


Figure 3.11: Static characteristics with input (a) and output (b), as well as with real and imaginary part (c) of the describing function of the three-position relay with hysteresis.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

derived from the describing function of a three-position relay with hysteresis, Fig. 3.11. For small input signals $X_m < x_a$, $y = 0$, and $G_N(X_m) = 0$. For $X_m \geq x_a$, the output signal has the form of rectangular pulses, Fig. 3.11b. The switching times t_a and t_b are given by:

$$\begin{aligned} X_m \sin \omega t_a &= x_a \\ \sin \omega t_a &= \frac{x_a}{X_m} \end{aligned} \tag{3.65}$$

and:

$$\begin{aligned} X_m \sin \omega \left(\frac{\pi}{\omega} - t_b \right) &= x_b \\ \sin \omega \left(\frac{\pi}{\omega} - t_b \right) &= \frac{x_b}{X_m} \end{aligned} \tag{3.66}$$

The real part of the describing function results from:

$$\begin{aligned} P(X_m) &= \frac{2}{\pi X_m} \int_{t_a}^{t_b} F(X_m \sin \omega t) \sin \omega t \; d(\omega t) = \frac{2}{\pi X_m} \int_{t_a}^{t_b} y_a \sin \omega t \; d(\omega t) \\ &= \frac{2 y_a}{\pi X_m} (\cos \omega t_a - \cos \omega t_b) \end{aligned} \tag{3.67}$$

From (3.65) and (3.66) follows:

$$\begin{aligned} \cos \omega t_a &= \sqrt{1 - \left(\frac{x_a}{X_m} \right)^2} \\ \cos \omega t_b &= \sqrt{1 - \left(\frac{x_b}{X_m} \right)^2} \end{aligned} \tag{3.68}$$

and:

$$P(X_m) = \frac{2 y_a}{\pi X_m} \left[\sqrt{1 - \left(\frac{x_a}{X_m} \right)^2} + \sqrt{1 - \left(\frac{x_b}{X_m} \right)^2} \right] \tag{3.69}$$

The imaginary part of the describing function is defined from the equation (3.61) and Fig. 3.11:

$$Q(X_m) = -\frac{2 y_a (x_a - x_b)}{\pi X_m^2} \tag{3.70}$$

so that:

$$\begin{aligned} G_N(X_m) &= P(X_m) + jQ(X_m) \\ &= \frac{2 y_a}{\pi X_m} \left[\sqrt{1 - \left(\frac{x_a}{X_m} \right)^2} + \sqrt{1 - \left(\frac{x_b}{X_m} \right)^2} \right] - j \frac{2 y_a (x_a - x_b)}{\pi X_m^2} \end{aligned} \tag{3.71}$$

The graphical display in Fig. 3.11c shows that the coefficients of harmonic linearization $P(X_m)$ and $Q(X_m)$ are discontinuous at the point $X_m = x_a$, i.e. at this point a pulse of finite duration appears.

For a single-valued relay characteristic, equation (3.71) becomes $x_a = x_b$, $Q(X_m) = 0$, and the describing function has only the real part:

$$G_N(X_m) = \frac{4y_a}{\pi X_m}\sqrt{1-\left(\frac{x_a}{X_m}\right)^2} \tag{3.72}$$

The describing function of a two-position relay follows from equation (3.72) for $x_a = 0$:

$$G_N(X_m) = \frac{4y_a}{\pi X_m} \tag{3.73}$$

In practice, normalized forms of describing function are used. For standard nonlinearities approximated by straight lines, the normalization of harmonic linearization occurs by substitution:

$$N = \frac{y_a}{x_a};\ A = \frac{X_m}{x_a};\ \lambda = \frac{x_b}{x_a} \tag{3.74}$$

where N is the normalizing factor, A is a dimensionless amplitude and λ is the reset coefficient. The normalized describing function of a three-position relay with hysteresis is:

$$\begin{aligned} G_N(A) &= NG_{N_0}(A) \\ G_{N_0}(A) &= \frac{2}{\pi A^2}\left[\sqrt{A^2-1}+\sqrt{A^2-\lambda^2}-j(1-\lambda)\right] \end{aligned} \tag{3.75}$$

where $G_{N_0}(A)$ is the normalized gain.

A graphical display of the function $G_N(X_m)$ in the complex plane is called the *amplitude characteristic of a nonlinear element*. Amplitude characteristic and describing function of basic standard nonlinear elements are shown in in Table A.1 in Appendix A.

As already mentioned, with the harmonic linearization method (linearization in frequency domain), nonlinear characteristic $y_N(t) = F(x)$ has instead of one line a "bundle of lines" with slopes that depend upon the amplitude of oscillations of variable $x(t)$, i.e. upon the "operating interval" of $y_N(t) = F(x)$ spread by the oscillations $x(t) = X_m sin\omega t$.

EXAMPLE 3.2

(DESCRIBING FUNCTION OF A TWO-POSITION RELAY WITHOUT HYSTERESIS)

The following example which shows clearly the significance of the describing function is the two-position relay without hysteresis, with harmonic input of various amplitudes.

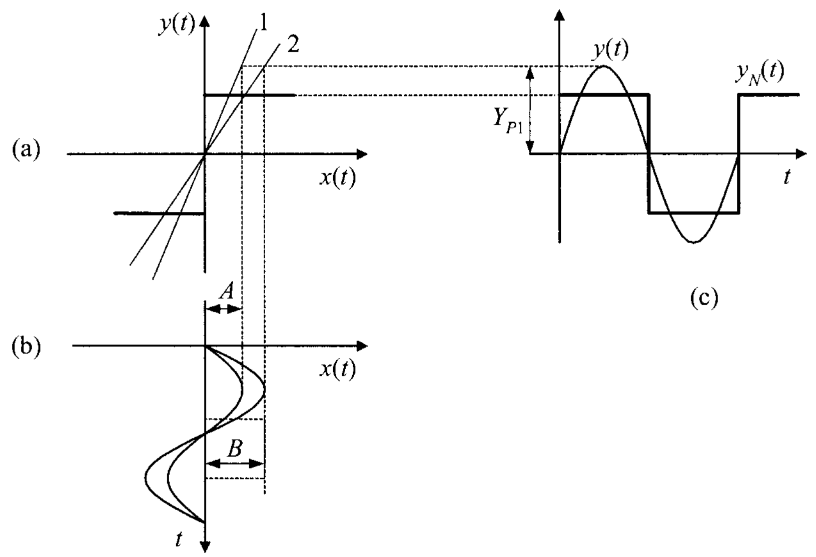


Figure 3.12: (a) Static characteristic of a two-position relay without hysteresis obtained by harmonic linearization in case of sinusoidal input of amplitude A (line 1) and amplitude B (line 2), (b) harmonic inputs, (c) output of the nonlinear element y_N with its first harmonic approximation $y(t)$.

If two harmonic inputs $x(t) = X_m \sin\omega t$ are independently applied at the input of a nonlinear element, with $X_m = A$ and $X_m = B$, respectively, the output will be in both cases the same rectangular signal with the same first harmonic (Fig. 3.12c), $y(t) = Y_{P1} \sin\omega t$ where Y_{P1} is the amplitude of the first harmonic. The coefficient of harmonic linearization in the case of the input signal having amplitude $X_m = A$ is equal to the slope of line 1 (Fig. 3.12a) and is given by the expression (see (3.57)):

$$G_N(X_m) = G_N(A) = P(A) = \frac{Y_{P_1}}{A} \tag{3.76}$$

If the amplitude changes to $X_m = B$, the coefficient of harmonic linearization is now equal to the slope of line 2 (Fig. 3.12a), which must be smaller since $B > A$, while Y_{P1} is in both cases the same. It follows:

$$G_N(B) = P(B) = \frac{Y_{P_1}}{B} < P(A)$$

The conclusion is that harmonic linearization illustrates the fact that a nonlinear element is equivalent to a linear element whose gain (coefficient of harmonic

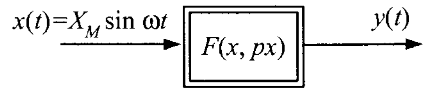


Figure 3.13: Nonlinear element with inertia subjected to harmonic input.

linearization) changes depending on the input signal amplitude.

$$G_N(A) = P(A) = \frac{Y_{P1}\sin\omega t}{A\sin\omega t} = \frac{Y_{P_1}}{A}$$
$$y = F(x) \approx G_N(A)x = \frac{Y_{P1}}{A}x$$

With an amplitude change in the input to the nonlinear element, the coefficient of harmonic linearization and the approximation of $y_N = F(x)$ *change, too.*

Describing Function of Nonlinearity $y_N(t) = F(x,\dot{x})$

If the signal $x(t) = X_m sin\omega t$ acts as the input to a nonlinear element with inertia, Fig. 3.13, the response at the output will be:

$$y_N(t) = F(X_m\sin\omega t, X_m\omega\cos\omega t) \tag{3.77}$$

By expansion of equation (3.77) into its Fourier series, asssuming that the nonlinearity with inertia is symmetrical, the output is:

$$y_N = P(X_m,\omega)x + \frac{Q(X_m,\omega)}{\omega}px + \text{higher order terms} \tag{3.78}$$

where $p = \frac{d}{dt}$ is a derivative operator and:

$$P(X_m,\omega) = \frac{1}{\pi X_m}\int_0^{2\pi} F(X_m\sin\varphi, X_m\omega\cos\varphi)\sin\varphi\; d\varphi$$
$$Q(X_m,\omega) = \frac{1}{\pi X_m}\int_0^{2\pi} F(X_m\sin\varphi, X_m\omega\cos\varphi)\cos\varphi\; d\varphi$$

The first harmonic of the output signal is:

$$y = P(X_m,\omega)x + \frac{Q(X_m,\omega)}{\omega}px \tag{3.79}$$

By substituting into equation (3.79) the derivative operator p with complex variable[10] $s = j\omega$, the first harmonic of the output can be described through describing

[10] Substitution is allowed since the harmonic excitation acts at the input of the nonlinear element.

function as:

$$y \approx G_N(X_m, \omega)x \tag{3.80}$$

where the describing function is:

$$G_N(X_m, \omega) = P(X_m, \omega) + jQ(X_m, \omega) \tag{3.81}$$

3.7 Statistical Linearization

When a regular signal $s(t)$ and a disturbance $n(t)$ act simultaneously on a control system, the nonlinearities can have a critical influence on the dynamic behavior of the system. As an example, consider the saturation nonlinearity shown in Fig. 3.14. The input signal is (Petrov et al., 1967; Netushil, 1983):

$$\begin{aligned} x(t) &= s(t) + n(t) \\ &= m_x(t) + v_x(t) \end{aligned} \tag{3.82}$$

where:

$s(t)$ - the useful (information-bearing) signal, with expected mean value $m_s(t)$,

$n(t)$ - the disturbance signal, with expected mean value $m_n(t)$,

$m_x(t)$ - the expected value of the total input signal $x(t)$, $(m_x(t) = m_s(t) + m_n(t))$,

$v_x(t)$ - the total variation of the input signal away from its expected value—dispersion.

Figure 3.14 illustrates the propagation of a noise-corrupted input signal through the saturation nonlinearity. Assuming that the input signal to the nonlinear element contains relatively little variation about its expected value, $|v_x(t)|_{max} << |m_x(t)|$ so that $x(t) \approx m_x(t)$, and assuming that $m_x(t) < x_a$, then the useful output signal of this particular nonlinear element will be proportional to the input signal:

$$m_y(t) \approx \alpha \cdot m_x(t) \tag{3.83}$$

At a low noise level in the input signal ($|n(t)| << |s(t)|$), the mathematical expectation of the complete input signal is essentially equal to the mathematical expectation of the useful signal ($m_x(t) \approx m_s(t)$). At a high noise level in the input signal ($|n(t)| >> |s(t)|$), the expected value of the input signal is small relative to the variation in the signal $|m_x| << |v_x(t)|$. In this case, the useful signal is damped to the extent that its effect essentially does not appear at the output of the nonlinear element—the large amplitude variations in the input cause the nonlinear element to be operating in the saturation region most of the time. Therefore, large fluctuations of the input signal will cause irregular oscillations of the output

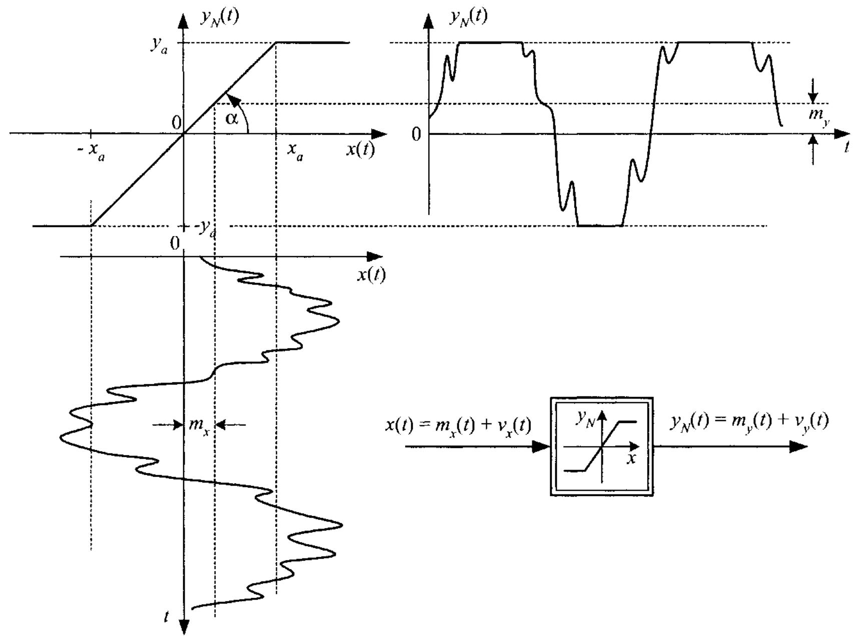


Figure 3.14: Noise corrupted excitation and response of a saturation nonlinearity.

signal, within the saturation limits. The appearance of irregular oscillations of the output signal causes changes in the spectral content of the output signal, i.e. it will be "rich" with both high and low harmonics. This can have a negative effect in feedback systems and cause undesired performance of the system.

Exact methods of determining the dynamic behavior of a nonlinear system under random input signals are not yet available. Analysis of the passing of useful signals and noise through a nonlinear element asks for approximate methods, of which the method of statistical linearization is most appropriate.

Essential to the procedure of statistical linearization is the replacement of a nonlinear element by—in a statistical sense—a linear one, which preserves some probabilistic characteristics of the output signal. These are commonly the mean value and dispersion. Two criteria are helpful:

1. Minimum of average least-squares difference of nonlinear and linearized system,

2. Equivalence of nonlinear and linearized system in mathematical expectation and dispersion.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

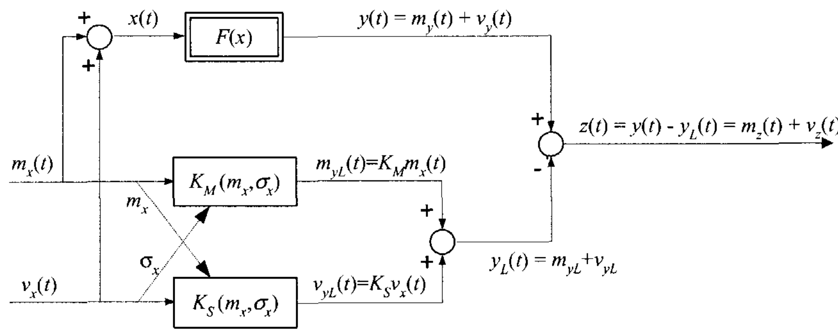


Figure 3.15: The substitution of a nonlinear element by an equivalent in a statistical sense linear element.

The substitution of a nonlinear element by an equivalent one, in a statistical sense linear element, is shown in Fig. 3.15. According to Fig. 3.15, the nonlinear dependency of the output and input signals of a nonlinear element without inertia is:

$$y_N(t) = F[x(t)] \tag{3.84}$$

The stationary stochastic process at the input of the nonlinear element is from equation (3.82):

$$x(t) = m_x(t) + v_x(t) \tag{3.85}$$

and the output signal $y_N = F(m_x + v_x)$ is also a stochastic process described by:

$$y_N(t) = m_y(t) + v_y(t) \tag{3.86}$$

where $m_y(t)$ and $v_y(t)$ are mathematical expectation (mean value) and variation (dispersion) [11] of the output signal around its expected value, respectively.

Nonlinear element $y_N = F(x)$ is replaced by an equivalent linear element having its gain dependent on the mathematical expectation $m_x(t)$ and dispersion of the process $v_x(t)$, i.e. the components of the stationary process of the input signals m_x and $v_x(t)$ propagate through two different channels with the gains K_M and K_S.

At the output of the ideal linear element, Fig. 3.15, we have:

$$y_L(t) = K_M m_x(t) + K_S v_x(t) = m_{y_L}(t) + v_{y_L}(t) \tag{3.87}$$

where K_M is the statistical gain coefficient of the regular component m_x of the input signal $x(t)$, and K_S is the statistical gain coefficient of irregular component $v_x(t)$ of the input signal $x(t)$.

[11] Centered component for which the expectation is zero.

According to the equation (3.87), coefficients of statistical linearization K_M and K_S are determined from the condition of "preserving" mathematical expectation and standard deviation of the nonlinear function, when it is replaced by a linear model. From equation (3.87) the relation for K_m can be written:

$$K_M m_x(t) = m_{y_L}(t) = m_y(t) \tag{3.88}$$

or:

$$K_M = \frac{m_y(t)}{m_x(t)} \tag{3.89}$$

and for K_S:

$$D_y = K_S^2 [v_x(t)]^2 = K_S^2 D_x \tag{3.90}$$

and further:

$$K_S = \sqrt{\frac{D_y}{D_x}} = \pm \frac{\sigma_y(t)}{\sigma_x(t)} \tag{3.91}$$

where D_y, D_x are variances and σ_x, σ_y are standard deviations (dispersions) of stochastic components of the output and input signals of the nonlinear element, respectively.

Expressions (3.89) and (3.90) for K_M and K_S are valid only for central symmetrical characteristics. The sign of K_S is given by the properties of the nonlinear function $y_N = F(x)$:

$$\operatorname{sign} K_S = \operatorname{sign} \left.\frac{dF(x)}{dx}\right|_{x=m_x} \tag{3.92}$$

Coefficients K_M and K_S can be evaluated from the minimum mean-square error of the difference $z(t)$ (Fig. 3.15):

$$E\left[z^2(t)\right] = E\left\{[y(t) - y_L(t)]^2\right\} = E\left\{[m_y(t) + v_y(t) - K_M m_x(t) - K_S v_x(t)]^2\right\} \tag{3.93}$$

This expands to:

$$\begin{aligned} E\left[z^2(t)\right] = E\Big\{&[m_y(t) + v_y(t)]^2 - 2[m_y(t) + v_y(t)][K_M m_x(t) + K_S v_x(t)] \\ &+ [K_M m_x(t) + K_S v_x(t)]^2\Big\} \end{aligned}$$

and further:

$$\begin{aligned} E\left[z^2(t)\right] &= E\left\{m_y^2(t) + 2m_y(t)v_y(t) + v_y^2(t)\right. \\ &\quad - 2m_y(t)K_M m_x(t) - 2m_y(t)K_S v_x(t) - 2v_y(t)K_M m_x(t) - 2v_y(t)K_S v_x(t) \\ &\quad \left. + K_M^2 m_x^2(t) + 2K_M m_x(t) K_S v_x(t) + K_S^2 v_x^2(t)\right\} \\ &= E\left\{m_y^2(t) + v_y^2(t) + K_M^2 m_x^2(t) + K_S^2 v_x^2(t)\right. \\ &\quad - 2m_y(t)K_M m_x(t) - 2v_y(t)K_S v_x(t) \\ &\quad \left. + [2m_y(t)v_y(t) + 2K_M m_x(t)K_S v_x(t) - 2m_y(t)K_S v_x(t) - 2v_y(t)K_M m_x(t)]\right\} \end{aligned}$$

With respect to (3.88) where $K_M m_x(t) = m_y(t)$ the term in brackets is equal to zero and the following relation is derived:

$$E\left[z^2(t)\right] = E\left\{m_y^2(t) + v_y^2(t) + K_M^2 m_x^2(t) + K_S^2 v_x^2(t) - 2m_y(t)K_M m_x(t) - 2v_y(t)K_S v_x(t)\right\} \tag{3.94}$$

Applying the substitutions:

$$E\left\{\left[v_x^2(t)\right]\right\} = \sigma_x^2(t)\,; \quad E\left\{\left[v_y^2(t)\right]\right\} = \sigma_y^2(t) \quad \text{and}$$
$$E\left\{2v_y(t)K_S v_x(t)\right\} = 2K_S E\left\{v_y(t)v_x(t)\right\} = 2K_S R_{xy}(t)$$

where R_{xy} is the cross-correlation of v_x and v_y. Finally from (3.94) follows the mean-square deviation of $z^2(t)$ as:

$$E\left\{z^2(t)\right\} = \bar{z}^2(t) = m_y^2(t) + \sigma_y^2(t) + K_M^2 m_x^2(t) + K_S^2\sigma_x^2(t) - 2m_y(t)K_M m_x(t) - 2K_S R_{xy}(t)$$

Coefficients K_M and K_S are determined from the condition $\bar{z}^2(t) = \min$:

$$\frac{\partial \bar{z}^2}{\partial K_M} = 2K_M m_x^2(t) - 2m_x(t)m_y(t) = 0 \tag{3.95}$$

from which:

$$K_M = \frac{m_y}{m_x} \tag{3.96}$$

and from:

$$\frac{\partial \bar{z}^2}{\partial K_S} = 2\cdot K_S\sigma_x^2 - 2\cdot R_{xy}(t) = 0$$

it follows that:

$$K_S = \frac{R_{xy}(t)}{\sigma_x^2} \tag{3.97}$$

In order to define the coefficients K_M and K_S it is necessary to provide the one-dimensional probability density $p(x)$ of the signal $x(t)$ at the input of the nonlinear element. The expressions (3.89), (3.91) and (3.97) are defined by equations:

$$K_M m_x = m_y = \int_{-\infty}^{\infty} F(x)p(x)dx \tag{3.98}$$

$$K_{S_1} = \pm\left\{\frac{1}{D_x}\int_{-\infty}^{\infty} F^2(x)p(x)dx - m_y^2\right\}^{\frac{1}{2}} \tag{3.99}$$

$$K_{S_2} = \frac{1}{D_x}\int_{-\infty}^{\infty} (x - m_x)F(x)p(x)dx \tag{3.100}$$

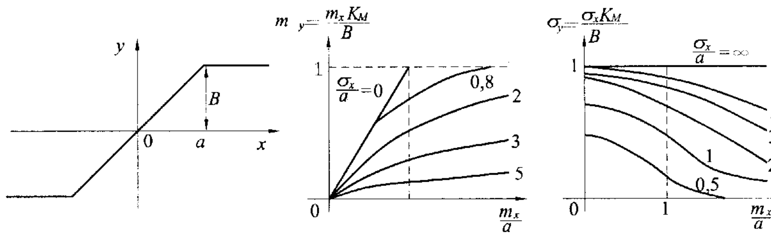


Figure 3.16: Normalized K_M and K_S for the nonlinearity of the type saturation.

When the coefficients of linearization K_S are determined with the expressions (3.99) and (3.100), the results are near to exact results. Still better values of K_S are achieved by taking the mean arithmetic value:

$$K_S = \frac{K_{S_1} + K_{S_2}}{2} \tag{3.101}$$

From the expressions for K_M and K_S it follows that the coefficients of linearization depend not only on the form of nonlinearity $y_N = F(x)$, but also on the probability density of the random quantity $p(x)$ as well. In order to determine the coefficients K_M and K_S, the random process which acts upon the input to the nonlinearity, must be subdued to the normal distribution law. As most real automatic control systems are described by a combination of a basic nonlinearity plus a linear part that behaves as a low-pass filter, the change in the form of the distribution law in broad limits is not essential for the accuracy of evaluating the coefficients K_M and K_S. With normal distribution law, the probability density is uniquely determined by the mathematical expectation m_x and the dispersion σ_x of the random process $x(t)$. The coefficients K_M and K_S with known m_x and σ_x, i.e. $K_M = K_M(m_x, \sigma_x)$ and $K_S = K_S(m_x, \sigma_x)$ are also unique. Namely, by applying linearization (Fig. 3.15), the coefficients K_M and K_S are functions of the useful input signal and of the added noise. For basic nonlinearities, coefficients of linearization are evaluated both analytically and graphically. Figs. 3.16 and 3.17 give normalized graphical relations of K_M and K_S for typical nonlinerities, such as saturation and three-position relay without hysteresis.

According to this graphical display, obviously the presence of a random component of the input signal $x(t)$ "smooths" the nonlinear function $y_N = F(x)$ in terms of the mean value of the signal. For $\sigma_x = 0$, i.e. for $v_x(t) = 0$, the function $m_y = f(m_x)$ is equal to the static characteristic of the nonlinear element $y_N = F(x)$. By increasing σ_x, the slope of the nonlinear static characteristic $m_y = f(m_x)$ is diminished. The random component of the input signal "linearizes" the characteristic of the nonlinear element for a deterministic component (m_x, mean value) of the input signal. For instance, a relay element behaves as an element with a conti-

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

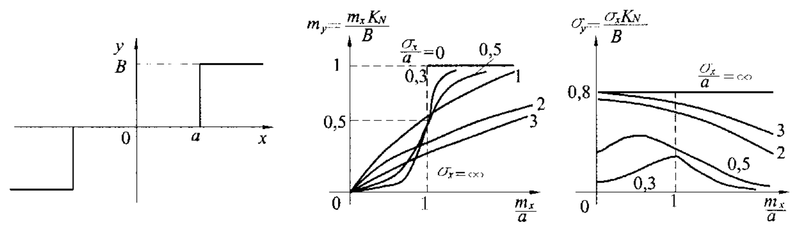


Figure 3.17: Normalized K_M and K_S for the nonlinearity of the type three-position relay without hysteresis.

nuous static characteristic when a random input component $v_x(t)$ acts instead of a deterministic component $m_x(t)$. By increasing m_x for these nonlinearities, the gain of the random component K_S decreases, i.e. the propagation of a random component is diminished since the element is saturated with the deterministic component of the signal.

The application of the statistical linearization method is limited by the demands that the input signal is subject to a normal distribution law of the probability density, and that at the output a linear inertial element is attached with low-pass filter behavior. With these conditions, the statistical gain K_S can be defined by:

$$K_S = \frac{dm_y}{dm_x} \tag{3.102}$$

instead with the equation (3.100).

3.8 Combined (Dual-Input) Describing Functions

As the name implies dual-input (or combined)[12] describing functions have more than one signal at the input to the nonlinear element. They are useful in analyzing nonlinear systems when subharmonic oscillations occur, i.e. when harmonic inputs of a given frequency ω cause an output with subharmonics.

Combined describing functions are applied to study the behavior of time-invariant nonlinear elements $y_N = F(x)$ or $y_N = F(x,\dot{x})$ in situations when an input to the nonlinear element consists of two signals. In such situations, three possible cases appear most often (Csáki, 1972):

1. Input signal is formed by two sinusoidal signals,

[12] Sometimes in the literature the term dual-input describing function is used to describe nonlinear elements excited by the input signal $x(t) = x^0 + X_\omega \sin \omega t$, where the amplitude x^0 is relatively small compared to X_ω. However, two-input describing function is sometimes used if the input to the nonlinear element is as (3.103).

2. Input signal is formed by one sinusoidal and one random signal,

3. Input signal is formed by two random signals.

When at the input of the nonlinear element two random signals with normal distribution are applied, the behavior of the nonlinear element can be analyzed by the use of statistical linearization with one random input. Namely, the sum of two signals with normal distribution has also the normal distribution, and the total dispersion is the sum of two dispersions.

In order to find out the properties of nonlinear systems, the dual-input describing function is an appropriate choice. When the input signal consists of two harmonic components with proportional frequencies, the input will be:

$$x(t) = X_\omega \sin(\omega t + \varphi) + X_\Omega \sin \Omega t \tag{3.103}$$

where $\Omega = m\omega$ and $m = b$ or $m = 1/b$, b is an integer. The output signal of the nonlinear element is:

$$y_N(t) = P_1 \sin(\omega t + \varphi) + Q_1 \cos(\omega t + \varphi) + P_m \sin m\omega t + Q_m \cos m\omega t + \ldots \tag{3.104}$$

From this equation, the describing functions for the fundamental component ω and harmonic component $m\omega$ are written down:

$$G_{N_1}(X_\omega, X_\Omega, \varphi, m) = \frac{1}{X_\omega}\left[P_1(X_\omega, X_\Omega, \varphi, m) + jQ_1(X_\omega, X_\Omega, \varphi, m)\right] \tag{3.105}$$

$$G_{N_m}(X_\omega, X_\Omega, \varphi, m) = \frac{1}{X_\Omega}\left[P_m(X_\omega, X_\Omega, \varphi, m) + jQ_m(X_\omega, X_\Omega, \varphi, m)\right] \tag{3.106}$$

It must be noted that the above describing functions depend on X_ω, X_Ω, ω, Ω, φ, and possibly (not given in 3.103) upon the phase shift of the second harmonic component. Otherwise, the existence of two describing functions complicates the examination of the system behavior, as the evaluation of the describing function is a complex and laborious process since six variables are in play. The method of dual-input describing function is appropriate when solving problems linked to the dynamic behavior of nonlinear systems, the stability of forced oscillations, resonance jump, analysis of subharmonic oscillations, the study of dynamics for a system with two nonlinear elements, and finally for the application of dither signals in nonlinear systems.

EXAMPLE 3.3
(DUAL-INPUT DESCRIBING FUNCTION OF A CUBIC NONLINEARITY)
The build-up of subharmonic oscillations can be shown for the case of a cubic nonlinearity:

$$y_N(t) = F(x) = ax^3(t) \tag{3.107}$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

If the input is a monoharmonic signal $x(t) = X_\omega sin\omega t$, the output will be[13]*:*

$$y_N(t) = F(x) = a(X_\omega \sin\omega t)^3 = \frac{3}{4}aX_\omega^3 \sin\omega t - \frac{1}{4}aX_\omega^3 \sin 3\omega t \tag{3.108}$$

and the describing function is:

$$G_N(X_\omega) = P(X_\omega) = \frac{3}{4}aX_\omega^2 \tag{3.109}$$

In the case of a polyharmonic input $x(t) = X_\omega sin(\omega t + \varphi) + X_\Omega sin\Omega t$, where $\Omega = m\omega$, the respective output will be:

$$y_N(t) = F(x) = ax^3(t) = a\left[X_\omega \sin(\omega t + \varphi) + X_\Omega \sin m\omega t\right]^3 \tag{3.110}$$

After raising to a third power and trigonometric rearranging there follows:

$$\begin{aligned} y_N(t) = {} & \frac{3}{4}a(X_\omega^2 + 2X_\Omega^2)X_\omega \sin(\omega t + \varphi) + \frac{3}{4}a(2X_\omega^2 + X_\Omega^2)X_\Omega \sin m\omega t \\ & - \frac{1}{4}aX_\omega^3 \sin 3(\omega t + \varphi) - \frac{1}{4}aX_\Omega^3 \sin 3m\omega t \\ & - \frac{3}{4}aX_\omega^2 X_\Omega \{\sin[(m-2)\omega t - 2\varphi] + \sin[(m+2)\omega t + 2\varphi]\} \\ & + \frac{3}{4}aX_\omega X_\Omega^2 \{\sin[(2m-1)\omega t - \varphi] - \sin[(2m+1)\omega t + \varphi]\} \end{aligned} \tag{3.111}$$

From (3.111) follows that output signal $y_N(t)$ contains components with frequencies ω and $m\omega$ as well as higher harmonic components. The output containing only components with frequencies ω and $m\omega$ is:

$$y(t) = y_1(t) + y_m(t) \tag{3.112}$$

respectively:

$$y_1(t) = G_{N1}(X_\omega, X_\Omega, \varphi, m)X_\omega \sin(\omega t + \varphi) \tag{3.113}$$
$$y_m(t) = G_{Nm}(X_\omega, X_\Omega, \varphi, m)X_\Omega \sin m\omega t \tag{3.114}$$

where G_{N1} and G_{Nm} are dual-input describing functions of the input signal components.

According to (3.53), (3.54) and (3.57), equations (3.113) and (3.114) can be rewritten as:

$$y_1(t) = Y_{P1} \sin(\omega t + \varphi) + Y_{Q1} \cos(\omega t + \varphi) \tag{3.115}$$
$$y_m(t) = Y_{Pm} \sin m\omega t + Y_{Qm} \cos m\omega t \tag{3.116}$$

[13] After applying trigonometric relations and rearranging.

respectively:

$$y_1(t) = \mathrm{Im}\left\{ \left(\frac{Y_{P1}}{X_\omega} + j\frac{Y_{Q1}}{X_\omega} \right) X_\omega e^{j(\omega t+\varphi)} \right\} \tag{3.117}$$

$$y_m(t) = \mathrm{Im}\left\{ \left(\frac{Y_{Pm}}{X_\Omega} + j\frac{Y_{Qm}}{X_\Omega} \right) X_\Omega e^{jm\omega t} \right\} \tag{3.118}$$

For $m = 1/3$, from (3.111) and relations for $y_1(t)$ and $y_m(t)$, follows that the dual-input describing function for cubic nonlinearity and for the input signal of frequency ω is given by:

$$G_{N1}(X_\omega, X_\Omega, \varphi, m) = \frac{3}{4}a\left(X_\omega^2 + 2X_\Omega^2\right) \tag{3.119}$$

while for the input harmonic component of frequency $\Omega = m\omega$, the dual-input describing function is:

$$G_{Nm}(X_\omega, X_\Omega, \varphi, m) = \frac{3}{4}a\left(2X_\omega^2 + X_\Omega^2\right) \tag{3.120}$$

Note that neither dual-input describing function depends on the phase shift and therefore are real.

For $m = 1/3$ and from (3.111) and (3.117):

$$\begin{aligned} &\mathrm{Im}\left\{ G_{N1}(X_\omega, X_\Omega, \varphi, \frac{1}{3}) \right\} \\ &= \frac{3}{4}a\left(X_\omega^2 + 2X_\Omega^2\right) X_\omega \sin(\omega t + \varphi) - \frac{1}{4}aX_\Omega^3 \sin\omega t \end{aligned} \tag{3.121}$$

The dual-input describing function for subharmonic ($m = 1/3$) input component is real, and given by:

$$G_{N_{1/3}}(X_\omega, X_\Omega, \varphi, \frac{1}{3}) = \frac{3}{4}a\left(2X_\omega^2 + X_\Omega^2\right) \tag{3.122}$$

while the dual-input describing function for the first harmonic is complex:

$$\begin{aligned} G_{N_1}(X_\omega, X_\Omega, \varphi, \frac{1}{3}) = &\frac{3}{4}a\left(X_\omega^2 + 2X_\Omega^2\right) - \frac{1}{4}a\frac{X_\Omega^3}{X_\omega}\cos\varphi \\ &+ j\frac{1}{4}a\frac{X_\Omega^3}{X_\omega}\sin\varphi \end{aligned} \tag{3.123}$$

For $m = 3$ from (3.111) and relations for $y_1(t)$ (3.117) and $y_m(t)$ (3.118), follows:

$$G_{N1}(X_\omega, X_\Omega, \varphi, 3) = \frac{3}{4}a(X_\omega^2 + 2X_\Omega^2 - X_\omega X_\Omega \cos 3\varphi) + j\frac{3}{4}aX_\omega X_\Omega \sin 3\varphi \tag{3.124}$$

$$G_{N3}(X_\omega, X_\Omega, \varphi, 3) = \frac{1}{4}a\left[3(2X_\omega^2 + X_\Omega^2) - \frac{X_\omega^3}{X_\Omega}\cos 3\varphi\right] - j\frac{1}{4}a\frac{X_\omega^3}{X_\Omega}\sin 3\varphi \tag{3.125}$$

If at the input of cubic nonlinear element the signal:

$$x(t) = X_\omega \cos \omega t + X_\Omega \cos\left(\frac{\omega t}{3} + \varphi\right)$$

is applied[14]*, the output will be:*

$$\begin{aligned} y_N(t) =& a\left(\frac{3X_\omega^3}{4} + \frac{3X_\omega X_\Omega^2}{2}\right)\cos \omega t + a\left(\frac{3X_\omega^2 X_\Omega}{2} + \frac{3X_\Omega^3}{4}\right)\cos\left(\frac{\omega t}{3} + \varphi\right) \\ &+ a\frac{3X_\omega X_\Omega^2}{4}\cos\left(\frac{-\omega t}{3} + 2\varphi\right) + a\frac{3X_\Omega^3}{4}\cos(\omega t + 3\varphi) + \text{higher order terms} \end{aligned} \tag{3.126}$$

The dual-input describing functions are now:

$$G_{N_1}(X_\omega, X_\Omega, \varphi, \frac{1}{3}) = a\frac{3}{4}X_\omega^2 + a\frac{3}{2}X_\Omega^2 + a\frac{1}{4}\left(\frac{X_\Omega^3}{X_\omega}\right)e^{j3\varphi} \tag{3.127}$$

$$G_{N_{1/3}}(X_\omega, X_\Omega, \varphi, \frac{1}{3}) = a\frac{3}{2}X_\omega^2 + a\frac{3}{4}X_\Omega^2 + a\frac{3}{4}X_\omega X_\Omega e^{-j3\varphi} \tag{3.128}$$

$G_{N_{1/3}}$ *is the describing function of* $m = 1/3$ *subharmonic and has the phase shift:*

$$\tan\beta = \frac{-b\sin 3\varphi}{2 + b^2 + b\cos 3\varphi} \tag{3.129}$$

where $b = X_\Omega / X_\omega$.

Maximal and minimal phase shift is approximately ± 21 deg. *We can conclude from this that if the control system has the cubic nonlinear element and the linear part which gives the phase shift not bigger than* 159 deg, *the* 1/3 *subharmonics will not be present in the system excited by the input signal of the form* $x(t) = X_\omega \cos \omega t + X_\Omega \cos(\frac{\omega t}{3} + \varphi)$.

This example shows the difficulties which are encountered with dual-input describing functions. If the nonlinear element is more complex, the difficulties are even more aggravating. Common dual-input describing functions are given in Gelb and Vander Velde (1968).

Analogous to the procedure of obtaining the dual-input describing function, a procedure is worked out for combined harmonic and statistical linearization[15]. The method is helpful when the input signal has harmonic and random components:

$$x(t) = X_\omega \sin \omega t + m_w(t) + \mathring{w}(t) \tag{3.130}$$

where $m_w(t)$ is the mathematical expectation of the random signal $w(t)$, and $\mathring{w}(t)$ is the centered random component of the random signal $w(t)$.

[14] Note the difference from before when the input was $x(t) = X_\omega \sin(\omega t + \varphi) + X_\Omega \sin \Omega t$.

[15] In following considerations the method of combined and statistical linearization will not be treated, so further comments will be omitted.

3.9 Conclusion

The classical dilemma for control engineers is always "to linearize or not to linearize the mathematical model." For physicists this dilemma does not exist, because they need the most accurate models. However, control engineers are aware of the fact that sophisticated mathematical models provide accurate descriptions of the system behavior, that the linearized model gives only a local description of nonlinear system behavior, and that some of the intricacies of the nonlinear system behavior may be lost through linearization process. Control engineers must always weight the accuracy of the nonlinear model against the simplicity of the linear model, which is more amenable for analysis and design of control systems. If linearization is acceptable, meaning that the consequences of linearization are tolerable, then it should be recommended. However, we should be aware of consequences which linearization can make, and this chapter hopefully can help in that. Several linearization methods about equilibrium state are covered in this chapter. They are very often used in control applications. Harmonic linearization has great appeal in the control community because engineers are well prepared to analyze systems in the frequency domain. Due to the fact that nonlinear systems often operate in the oscillatory way, this method of linearization lends itself naturally in this case, and this is the main reason why in the following chapters this kind of linearization will be used.

The main limitations in the utility of linearizations are (a) control design based on a linearized mathematical model of the plant will behave poorly, especially if the plant operating point moves away from the equilibrium point about which the linearization is made, (b) the system does not possess filtering capability needed for applying the harmonic linearization.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Chapter 4

Operating Modes and Dynamic Analysis Methods

This chapter serves as a bridge between the first and the second part of the book. While in the first part we intended to be more general and talk about nonlinear systems in general, in the second part of the book nonlinear control systems are our prime topic. So, the operating modes in which a nonlinear control system usually operate are very briefly presented in this chapter. Dynamic analysis methods are also briefly mentioned. Some of them are very common in engineering practice, and they are elaborated in subsequent chapters.

4.1 Operating Modes of Nonlinear Control Systems

An important property of nonlinear automatic control systems is the possibility to restore oscillations as one of possible dynamic states. Periodic oscillations can emerge as a consequence of either input periodic excitation (forced oscillations) or due to influence of the system, when oscillations are established without external excitation (self-oscillations), i.e. when the system operates in the stabilizing mode (Rosenwasser, 1969; Popov, 1960; Naumov, 1972; Netushil, 1983). Periodic oscillations which occur in nonlinear control systems are very often a normal operating mode. The periodic operating mode is normal in guidance of moving objects such as airplanes, rockets, ships, etc. Although the oscillations are unwanted especially for tracking control systems, in many cases they cannot be avoided. The control engineer should know appropriate methods which can reduce the oscillations to a tolerable level or even completely suppress them. Often in order to achieve this, the designer will use harmonic excitation signals with the frequency below the bandpass of the linear part of the system. The effect of

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

this will be that the system is forced to oscillate with a desired applied frequency and amplitude. This technique is called forced oscillations. Another way to attack this problem is by use of dither signals, whose frequency must be higher than the bandpass frequencies of the linear part of the system. These signals can change the dynamic properties of the nonlinear part of the system so that the desired effects of elimination or reduction of self-oscillations are achieved.

4.1.1 Self-Oscillations

Self-oscillations are inherent only in nonlinear systems, and the term itself occurs only in the theory of nonlinear oscillations. Linear control theory analyzes linear oscillations by examining the second-order linear process in an oscillatory dynamic mode:

$$m\ddot{x}+h\dot{x}+kx=f(t) \tag{4.1}$$

Similarly, in the theory of nonlinear systems, the integrals of the nonlinear dynamic component of the second order are treated:

$$m\ddot{x}+\varphi(\dot{x})+\psi(x)=f(t) \tag{4.2}$$

or in the general form:

$$m\ddot{x}+\psi(x,\dot{x},t)=0 \tag{4.3}$$

In equation (4.2) the damping force $\varphi(x)$ is a nonlinear function of the velocity. The existence of such force is indicated by the argument x in the function ψ in expression (4.3). Elastic force $\psi(x)$ is a nonlinear function of deflection of the shift x, respectively. The equations (4.2) and (4.3) describe oscillatory processes in various physical systems.

In the case when an external signal $f(t)$ doesn't act upon the system (4.2), or when the system (4.3) is time-invariant, the equation of the system assumes the form:

$$m\ddot{x}+\varphi(\dot{x})+\psi(x)=0 \tag{4.4}$$

respectively:

$$m\ddot{x}+\psi(x,\dot{x},0)=0 \tag{4.5}$$

In this case the system is time-invariant and unforced.

If the nonlinear system is not excited, the oscillations which are established as a consequence of initial conditions which differ from zero are called self-oscillations.

The integral term $k\frac{x^2}{2}$ in the linear system (4.1) and the term $\int_0^x \psi(x)dx$ of the nonlinear system (4.2) represent momentary potential energy level, while the term $m\frac{\dot{x}^2}{2}$ represents the momentary value of kinetic energy.

The terms $h\dot{x}$ and $\varphi(\dot{x})$ in equations (4.1) and (4.2), respectively, characterize the energy dependence between the system and the surroundings — for $h\dot{x}>0$ or

$\varphi(\dot{x}) > 0$ the system is dissipating the initial energy outwards, while for $h\dot{x} < 0$ or $\varphi(\dot{x}) < 0$ the system is receiving external energy. Generally, the systems with $h\dot{x} \neq 0$ or $\varphi(\dot{x}) \neq 0$ are called *dissipative systems*. Systems with $h\dot{x} = 0$ or $\varphi(\dot{x}) = 0$ have all the energy from the initial conditions, so the energy remains constant. Such systems with no exchange of energy with the surroundings are called *conservative systems*.

If one of the solutions of the equations (4.1), (4.2) and (4.3) is periodic function:

$$x(t) = x(t + kT),\ k = 1,2,3,\dots \tag{4.6}$$

where T is the oscillation period, the system being analyzed is in the periodic operating mode. Periodic oscillations (4.6) can be forced oscillations or self-oscillations. The latter are inherent to unforced systems. The forced oscillations occur when the control system is subject to a periodic input variable $f(t)$.

In linear systems oscillations are possible only in conservative systems. However, those oscillations are not self-oscillations because they can occur in unforced systems. In unforced dissipative linear second-order systems possible oscillations (damped or undamped) are described by:

$$x(t) = C_1 e^{\pm \frac{h}{2m} t} \sin(\omega_1 t + \varphi_1) \tag{4.7}$$

Note that oscillations described by (4.7) are not periodic in the sense defined by (4.6). Self-oscillations are not possible in linear systems. Namely, self-oscillations can occur in unforced nonlinear systems and an unforced linear system does not possess energy which will maintain oscillations. Contrary to that, unforced nonlinear dissipative systems can maintain oscillations due to energy resources which are charged from internal energy sources (such as power supply of an electronic amplifier). Amplitude, frequency and slope of self-oscillations depend on characteristics of nonlinear and linear parts of a system. Their occurrence will also depend on initial conditions. However, initial conditions do not have any influence on the amplitude, frequency or slope of self-oscillations.

The exact mathematical treatment of periodic operating modes of nonlinear systems is very complex, since it represents one of the possible solutions of nonlinear differential equations.

The problem of self-oscillations in automatic control systems has a great practical importance. The self-oscillations are not allowed in "normal" operating modes, but sometimes they are intentionally introduced in order to improve the dynamic properties of the system. For all these reasons, it is indispensable to know what causes establishing of self-oscillations, their parameters, stability and ways of elimination.

The exposition which follows will encompass the approximate engineering methods for the dynamic analysis of nonlinear control systems, given by the block diagram in Fig. 1.2.

4.1.2 Forced Oscillations

Contrary to self-oscillations, forced oscillations are not maintained by the system alone, but arise from an external periodic signal with the purpose that the system repeats these external oscillations with the same frequency.

The net effect is that the system is forced to oscillate with a frequency which is more acceptable. If the unwanted oscillations cannot be avoided, then such a treatment can bring the system to a more favorable operating mode. More about forced oscillations will be said in Section 6.2

4.1.3 Effects of High-Frequency Signal—Dither

It is well known that by injecting a high-frequency signal to the input of a nonlinear element, the dynamic behavior of the system can improve. The *dither signal* frequency normally lies above the bandpass frequencies of the system, with the purpose that the high-frequency signal cannot be detected at the output of the system. When this dither signal is used to eliminate or to damp oscillations, then the notion *signal stabilization* is used. Moreover, it is customary with this method to reduce the effects of static friction in mechanical systems, with the appropriate notion *dynamic lubrication*. Dither signals were for the first time applied in electromechanical servomechanisms with two-position controllers.

EXAMPLE 4.1
(INFLUENCE OF DITHER SIGNALS TO DEAD ZONE NONLINEARITY)
A servo system with dead zone, otherwise unresponsive to small amplitude signals, can become responsive with the help of a corresponding dither signal. To explain this phenomenon, the nonlinear element of the dead zone type is taken (Fig. 4.1a) to which a harmonic information bearing sygnal smaller than the dead zone is applied (Fig. 4.1b). Obviously no output signal will appear. However, if a high-frequency dither is superimposed on the harmonic input signal so that the combined amplitude is greater than the dead zone, Fig. 4.1c, the output signal will differ from zero and will contain a component of the same frequency and phase as the low-frequency signal (the carrier of information at the input). If we concentrate on this information-bearing signal, and neglect the frequency component of the dither signal, $\omega_0 >> \omega$*, which will be filtered out by the low-pass linear part of the system (positioned after the nonlinear element), then the analysis can be carried out by a dual-input describing function.*

EXAMPLE 4.2 (EFFECT OF THE DITHER SIGNAL ON THE SYSTEM'S GAIN)
The dither signal has a significant effect on the system's gain. This is illustrated on a nonlinear element with cubic characteristic (see Example 3.3, with $a = 1$*).*

With the input $x(t) = X_\omega \cos \omega t$*, the simple describing function (gain at low frequencies) will be:*

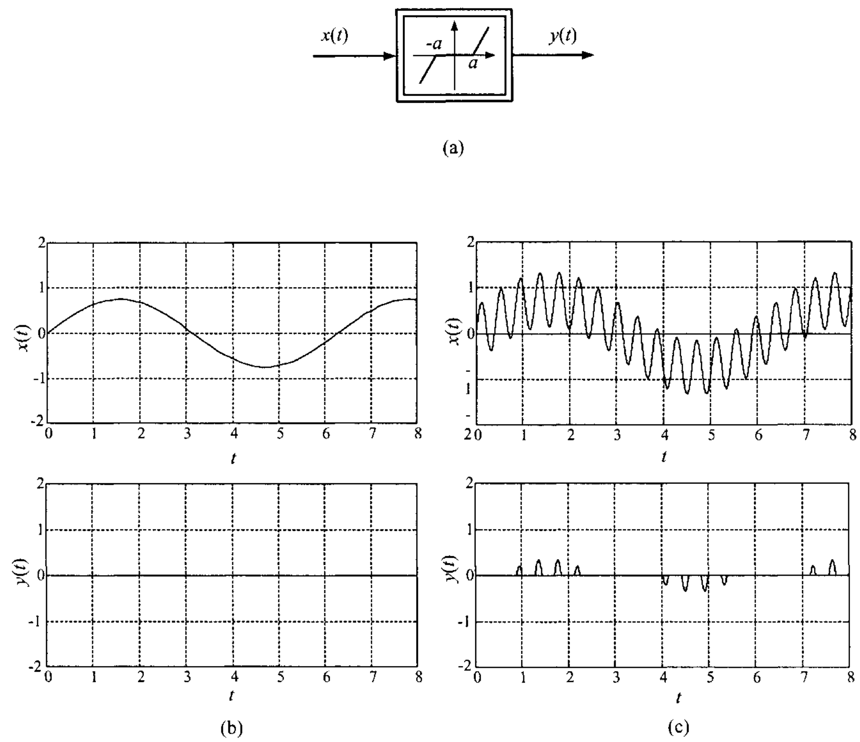


Figure 4.1: (a) Static characteristic of nonlinear element dead zone, (b) harmonic signal (information carrier), and (c) superimposed dither at the input and the signal of the output of the dead zone element.

$$G_N(X_\omega) = \frac{3}{4}X_\omega^2 \tag{4.8}$$

When at the input is added a high-frequency dither signal, the excitation will become:

$$x(t) = X_\omega \cos \omega t + X_\Omega \cos \Omega t \tag{4.9}$$

If we concentrate on the low-frequency component with frequency ω, the dual-input describing function will be:

$$G_{N_1}(X_\omega, X_\Omega) = \frac{3}{4}(X_\omega^2 + 2X_\Omega^2) \tag{4.10}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The amplitude of the high-frequency dither signal has a significant effect on the low-frequency gain. For instance, if $X_\omega = X_\Omega$, the gain will increase 200%:

$$G_N(X_\omega, X_\Omega) = 3G_N(X_\omega) \tag{4.11}$$

By analyzing both examples, the conclusion is that the effect of a dither signal applied at the input of the system is the following:

- It creates a large spectrum of new frequency components at the output of the nonlinear element.
- The amplitude of the dither signal has a large effect on the low-frequency gain (dual-input describing function).
- The dither signal frequency Ω has a negligible effect on gain (dual-input describing function).

Concept of Equivalent Nonlinearity

This concept plays a significant role in nonlinear systems, as it enables the change of dynamic properties of a nonlinear system by simple superposition of the dither signal with the input. Namely, the frequency and the form of the dither signal has the same effect as a change of the static characteristic of the nonlinear element. It is known that the dual-input describing function depends on the amplitude, as well as on the form, of the dither signal. This fact means that by a simple change of the form and amplitude of the dither signal we can obtain corresponding change in the nonlinear element, manifested by a change in the static characteristic. This concept can be best explained in an example.

EXAMPLE 4.3
(EQUIVALENT NONLINEARITY OF TWO-POSITION RELAY WITH RECTANGULAR DITHER)
In Fig. 4.2a the block diagram of a system with a two-position controller is given. If the input in Fig. 4.2b (information carrier) is combined with a high-frequency dither signal (Fig. 4.2c), it can be shown that at the output of the two-position controller a signal is formed as in Fig. 4.2d. This signal is the input to a low-frequency filter (process), and it will "feel" the signal as in Fig. 4.2e. Such a signal could be obtained if the input without dither would pass the nonlinearity of type three-position relay with dead zone, whose static characteristic is given in Fig. 4.3. This nonlinearity is called equivalent nonlinearity.

This simple example explains the basic idea of equivalent nonlinearity, which is obtained by a proper choice of the form, frequency and amplitude of the superimposed dither signal. What kind of equivalent nonlinearity is obtained depends nevertheless on the form of input signal. Thus, it is possible by a simple change of

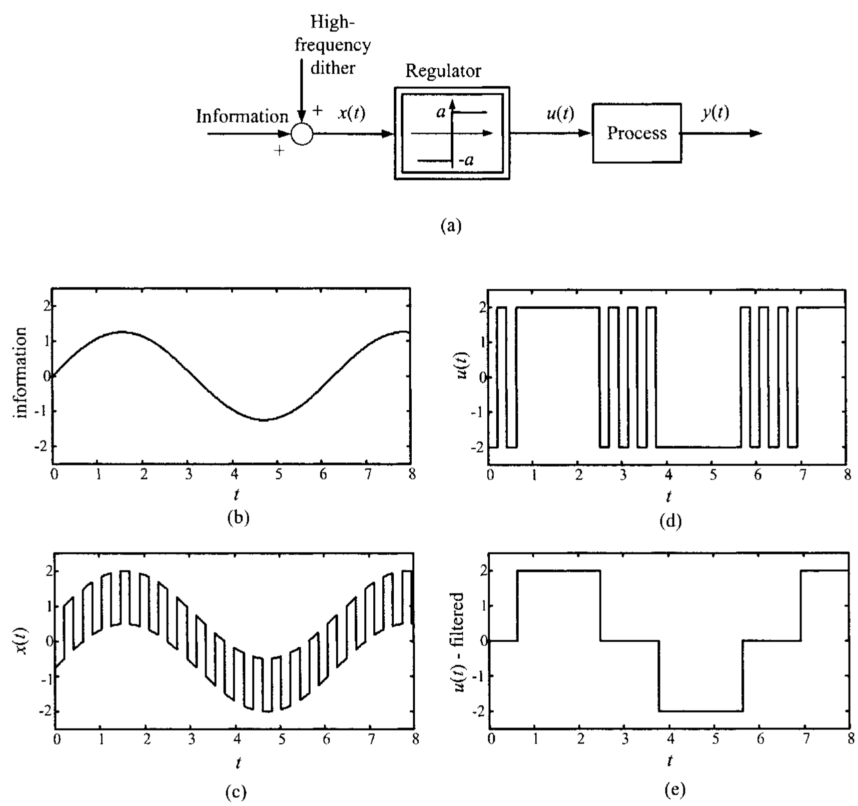


Figure 4.2: (a) Block diagram of two-position control of process, (b) excitation at the input of two-position control without dither, (c) excitation with superimposed rectangular dither signal, (d) output of the two-position regulator, (e) filtered signal as "seen" by the process.

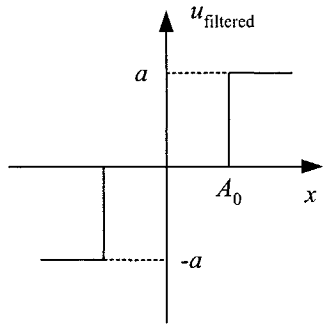


Figure 4.3: Equivalent nonlinearity for two-position controller with superimposed rectangular dither signal.

the form of dither signal to realize the desired dynamics of the nonlinear element. At the same time the closed-loop control dynamics depend upon the operating condition of the system. Thus, an adaptive control system can be realized. Such a technique belongs to *signal adaptation*, which is often useful in adaptive control systems.

4.2 Methods of Dynamic Analysis of Nonlinear Systems

The theory of nonlinear control systems, based upon the theory of nonlinear differential equations, is wider in scope than the theory of linear control systems. Due to the fact that in general[1] it is not possible to find the exact solution of the nonlinear differential equation, a variety of approximating solution methods was developed over time. These methods give us the possibility to approximately resolve the dynamics of a nonlinear system given by the block diagram in Fig. 1.2. A number of approximative methods are used in engineering practice, because a general method which can be applicable for all systems does not exist. All approximate methods can be classified[2] to:

- Classical (conventional) methods and
- Unconventional methods.

[1] Except in some specific cases.
[2] This is only one of many possible classifications.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Conventional methods are based upon qualitative methods of solving the nonlinear differential equations as well as on linearization methods (Chapter 3). Qualitative methods of solving nonlinear differential equations are based on the work of:

- Connection of solutions (Proell, 1884; Lecornu, 1898). For nonlinearities which are linear by segments (relay without hysteresis, saturation, dead zone and similar) dynamics can be analyzed by treating the system as linear in a particular segment and finding the complete solution by connecting solutions from each segment where the solution is passing through. The final point of a solution at the border of a segment presents the initial condition for the solution in the following segment.

- Phase trajectories (Léauté, 1885). Phase trajectories display dynamic behavior of the system and are convenient for systems up to third order. More about them will be said in Chapter 5.

- Mapping of points from the phase trajectory (Andronov and Bautin, 1944).

- Graphoanalytical method (Popov, 1960).

- Numerical and simulation methods. Simulation methods are lately very popular due to the fact that the theory of nonlinear systems is far from being so complete and general as is the case with the theory of linear systems. Advances in numerical analysis and technology are causing the use of numerical and/or simulation packages such as Matlab/Simulink[3], Matrix_X[4], Simnon[5], Easy5[6] and many others in education and research. For any control engineer use of simulation packages is indispensable and it should be stressed that almost every analysis and design attempt for nonlinear control systems must in one or the other stage of analysis or design use simulation.

All these methods are mathematically strict methods. By their application it is possible to obtain exact solution for many nonlinear mathematical models. Their main drawback is that they are applicable only to relatively simple nonlinear systems. Application to higher order nonlinear systems is unsuitable because complex and exhaustive mathematical operations are required. Linearization methods are based on:

- The theory of small parameter (disturbance) (Lyapunov, 1892; Poincaré, 1928),

- The describing function (see Chapter 5), and

[3]Matlab and Simulink are the trademarks of the MathWorks Inc. (USA).
[4]Matrix_X is the trademark of Integrated Systems Inc. (USA).
[5]Simnon is the trademark of the SSPA (Sweden).
[6]Easy5 is the trademark of Boeing Corp. (USA).

- Statistical linearization method (see Chapter 3).

By applying these methods, nonlinear mathematical models are replaced by linear or quasilinear mathematical models. Dynamic analysis is then possible by use of linear system theory.

Unconventional methods are methods which use fuzzy logic and/or methods which use neural networks. These methods are characterized by the fact that they do not need the mathematical model of the system, but instead are more heuristically oriented. Based on very successful solution of some quite demanding control problems it is certain that in the future unconventional methods will play a very important[7] role in control of nonlinear systems.

The use of approximate methods in dynamic analysis of nonlinear systems has some advantages as well as disadvantages. The universality and relative simplicity in application certainly belong to advantages of approximating methods. However, inability to estimate correctness of the obtained result and their mathematical difficulty when applied to higher order nonlinear systems are their main disadvantages. Despite aforementioned disadvantages, approximating methods are still the basic methods used by engineers for analysis and synthesis of nonlinear control systems. In the following chapters methods based on phase trajectories and the describing function applied for the nonlinear control systems with the structure given in Fig. 1.2 will be presented.

[7] If not a major role.

Chapter 5

Phase Trajectories in Dynamic Analysis of Nonlinear Systems

Of all methods used for nonlinear control system analysis, the method of phase trajectories is the most exact one, because the nonlinear differential equation has to be solved in one way or another. Before, when computers were not so common, this method was often used. Today, with powerful computers and software packages, the method mainly serves for educational purposes. Because of that it was decided that this topic should be included in our text, due to its capability to give better insight in the physical nature of nonlinear control systems behavior.

The state of the unforced linear time-invariant (LTI) system can be described by the homogenous differential equation of n-th order (external description):

$$\sum_{i=0}^{n} a_i \frac{d^i x(t)}{dt^i} = a_n \frac{d^n x(t)}{dt^n} + a_{n-1} \frac{d^{n-1} x(t)}{dt^{n-1}} + \ldots + a_2 \ddot{x} + a_1 \dot{x} + a_0 x = 0 \qquad (5.1)$$

or with n first-order differential equations written in matrix form (internal or state variable description):

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \qquad (5.2)$$

where:

a_i – coefficients of differential equation (5.1)

$\mathbf{A}$ – $n \times n$ system matrix

$\mathbf{x}$ – $1 \times n$ state vector

The state is at any moment t determined by knowledge of the coordinate x and its derivatives (from (5.1) or from the solution of (5.2)). Geometrical interpretation of the general solution of equation (5.1) will be a family of curves in n-dimensional space given by the coordinates $x,\dot{x},\ldots,x^{(n-1)}$. Such a space is called *phase space* or *state space*, and the curves which are solutions of equation (5.1) are called *phase* or *state trajectories*[1]. The group of all phase trajectories determined by possible initial conditions has the name *phase* or *state* portrait (Krasovskii and Pospelov, 1962; Besekerskii and Popov, 1975; Nelepin, 1971; and Léauté, 1885).

At any moment of time the state of the system (5.1) is determined either by a point in the phase space or by a point on the phase trajectory which corresponds to the particular moment t.

The coefficients $a_0,\cdots,a_n$ of the unforced system can be constants for LTI systems (linear equation), time-variant for linear time-variant systems (linear equation with variable coefficients), or functions of the variable x and its derivatives $x^{(i)}$ for nonlinear systems (nonlinear equation).

In phase space, the region of the coordinate origin $\frac{dx^i}{dt^i}=0,\ i=0,1,\cdots,n-1$, i.e. the region of equilibrium state, is of special interest. For linear systems, phase trajectories have the same form in the whole phase space. In other words, knowledge of the phase portrait (phase trajectories) in the region of small displacements provides knowledge of the state of the linear system anywhere in phase space. We can say that for linear systems local and global properties are equal.

With nonlinear systems the form of the phase trajectories can be different in various subspaces of the state space. Except for trajectories which are typical for linear systems, the phase portraits of nonlinear systems have specific phase trajectories, which are seen only in nonlinear systems, for instance a separatrix or a limit cycle. The dynamic behavior of a nonlinear system is in general not the same in various subspaces of phase space. However, it can be shown that any nonlinear system defined by an analytic function may be represented in a neighborhood of an equilibrium point by a linear system. If this is the case, then the local behavior of such a system can be characterized entirely in terms of linear theory. So, the local qualitative behavior of (5.2) is determined by the eigenvalues of $\mathbf{A}$, which we denote by $\lambda_i(i=1,2\ldots,n)$. For the second-order systems the nature of the eigenvalues of the matrix $\mathbf{A}$ may be visualized more clearly if we recall that the characteristic polynomial of $\mathbf{A}$ is:

$$\lambda^2 - tr(\mathbf{A})\lambda + \det(\mathbf{A})$$

where $tr(A)$ is the trace of the matrix[2], while $\det(\mathbf{A})$ is determinant of the matrix $\mathbf{A}$. The discriminant Δ is defined by:

$$\Delta = [tr(\mathbf{A})]^2 - 4\det(\mathbf{A}) \tag{5.3}$$

and the eigenvalues of $\mathbf{A}$ are:

$$\frac{1}{2}\left[tr(\mathbf{A}) \pm \sqrt{\Delta}\right] \tag{5.4}$$

[1]The notion of phase trajectories is used almost exclusively with grapho-analytical methods.
[2]The sum of the diagonal elements of $\mathbf{A}$.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

By analyzing the phase plane of the second-order linear system we can conclude about the patterns of phase trajectories of the nonlinear system with local deviations from the equilibrium point under consideration. While for linear systems these patterns are global and are valid in the whole phase plane, for nonlinear systems they will be valid only "in the small".

The method of phase trajectories is very convenient for the analysis of unforced nonlinear systems (with initial conditions differing from zero), when differential equations are of first and second order. When the nonlinear systems are of higher order, the method of phase trajectories is not suitable because of the complexity of the procedure and of the impossibility of a visual display of the phase trajectory or phase portrait. For some classes of nonlinear systems of higher order bifurcation theory and other numerical methods can be applied.

5.1 Phase Plane

5.1.1 Phase Trajectories of Linear Systems

The dynamic behavior of a second-order linear system can be described by differential equation (5.1) for $n = 2$ as:

$$\frac{d^2x(t)}{dt^2} + b_1\frac{dx(t)}{dt} + b_0x(t) = 0 \tag{5.5}$$

where $b_1 = \frac{a_1}{a_2}$ and $b_0 = \frac{a_0}{a_2}$.

Coefficients b_1 and b_0 are usually given by the *damping ratio*[3] (ζ) and *undamped natural frequency*[4] (ω_n) as $b_1 = 2\zeta\omega_n$ and $b_0 = \omega_n^2$. The geometric presentation of the general solution to differential equation (5.1) for $n = 2$ is a family of phase trajectories in the phase plane with coordinates x and $\dot{x}$. From (5.5) for $y = \dot{x} = \frac{dx(t)}{dt}$ two differential equations of first order emerge:

$$\begin{aligned}\frac{dy(t)}{dt} &= -b_1y(t) - b_0x(t)\\ \frac{dx(t)}{dt} &= y(t)\end{aligned} \tag{5.6}$$

By dividing the first of these two equations by the second (with $x \neq 0$ and $y \neq 0$), a differential equation of the phase trajectory is obtained:

$$\frac{dy(t)}{dx(t)} = -b_1 - b_0\frac{x(t)}{y(t)} \tag{5.7}$$

[3]The damping ratio is defined as the ratio of the actual damping constant to the critical value of the damping constant, $\zeta = \frac{a_1}{2\sqrt{a_0a_2}}$.

[4]The undamped natural frequency is defined as the frequency of the sustained oscillation of the transient if the damping is zero, $\omega_n = \sqrt{\frac{a_0}{a_2}}$.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The general solution of equation (5.7) is represented by a family of phase trajectories $\dot{x}_i = f(x_i)$ and with initial conditions $\dot{x}_{0i}$ and x_{0i}. The particular phase trajectory defines the transient process $x_i(t)$ in time domain for a particular initial condition. At the origin of the coordinate system $x = y = 0$, the slope of the tangent to the phase trajectory $\frac{dy}{dx} = -b_1 - b_0 \frac{0}{0}$ is undefined, so depending on the roots of the characteristic equation of (5.5), given by:

$$\lambda_{1,2} = -\frac{b_1}{2} \pm \sqrt{\frac{b_1^2}{4} - b_0} = -\zeta\omega_n \pm j\omega_n\sqrt{1-\zeta^2} = -\sigma \pm j\omega_d \tag{5.8}$$

six types of singular points[5] can be distinguished: center, stable focus, unstable focus, stable node, unstable node, and saddle (see Fig. 5.1). This figure classifies local structures, describing the patterns of phase trajectories "in the small" at small deviations from the singular point of the nonlinear system too. A singular point of the type center (Fig. 5.2) results from the general solution of equation (5.5) for the case when the roots of the characteristic equation are pure imaginary ($\Delta < 0$ and $tr(A) = 0$), i.e $\lambda_{1,2} = \pm j\sqrt{b_0} = \pm j\omega_n$, $b_1 = 0$ ($\zeta = 0$), $b_0 > 0$. The state of the linear system which originates from any initial condition will behave like an undamped oscillatory process, i.e. the linear system is on the edge of instability. Phase coordinates of the system are given by expressions:

$$x(t) = A\sin(\omega_n t + \varphi)\,;\; y(t) = \frac{dx}{dt} = \omega_n A\cos(\omega_n t + \varphi)\,;\; \omega_n = \sqrt{b_0} \tag{5.9}$$

with constants A and φ which are dependent on the initial conditions x_0 and y_0. The equation of phase trajectory (5.7) for the system (5.5) is a parametric equation of an ellipse with half-axes A and $\omega_n A$; for $\omega_n = 1[s^{-1}]$, the phase trajectory becomes a circle:

$$\frac{x^2}{A^2} + \frac{y^2}{(\omega_n A)^2} = 1 \tag{5.10}$$

From (5.10) it follows that undamped periodic oscillations of a linear system with constant amplitude and frequency are represented in the phase plane by a closed trajectory of a regular geometrical pattern—ellipse or circle, Fig. 5.2.

A singular point of the type stable focus follows from the general solution of equation (5.5) for the case when roots (5.8) are conjugate complex with negative real parts: $b_1^2 < 4b_0$, $b_1 > 0$, $b_0 > 0$ (see region 2 in Fig. 5.1). The general solution and state of the system tends to an equilibrium state by oscillating with an exponentially damped amplitude. Phase characteristics of the system are given by expressions:

$$x(t) = Ae^{-\sigma t}\sin(\omega_d t + \varphi),\; y(t) = \frac{dx}{dt} = \gamma A e^{-\sigma t}\cos(\omega_d t + \varphi + \Theta) \tag{5.11}$$

[5]Singular points are equilibrium points where the slope of the tangent to the trajectory is undefined.

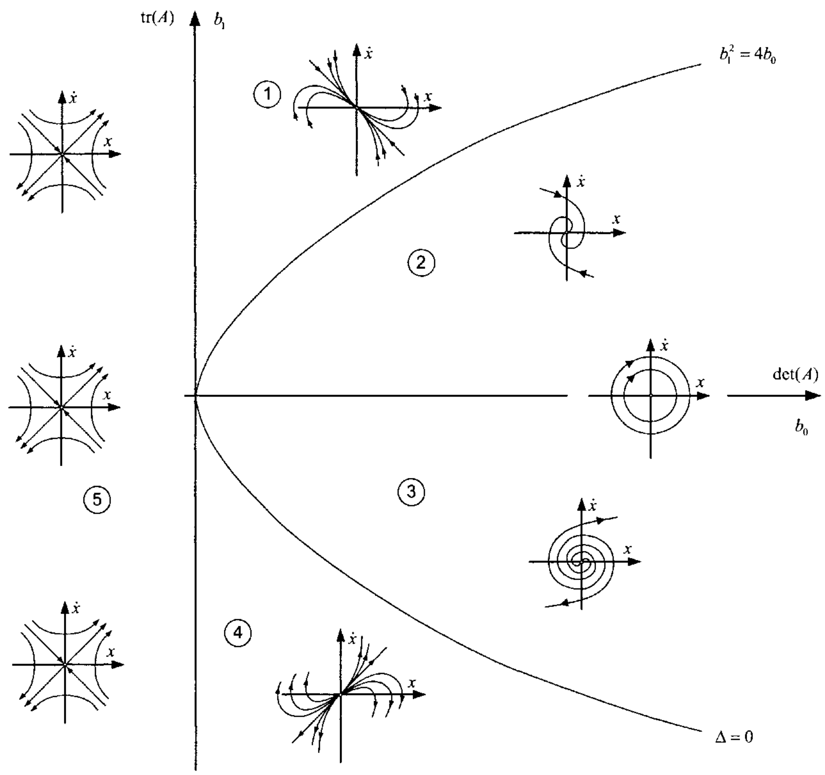
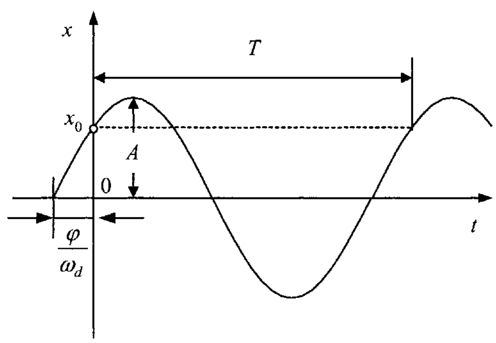
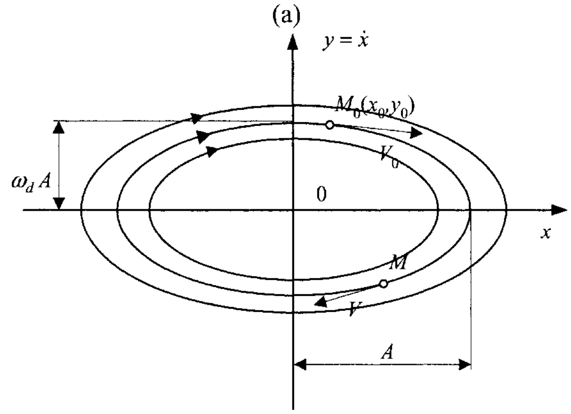


Figure 5.1: Classifying local structures.



(a)



(b)

Figure 5.2: State behavior of linear system (a) and phase portrait for case of singular point of the type center (b).

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

where:

$$\begin{aligned}
\sigma &= \zeta\omega_n = \frac{b_1}{2};\ \omega_d = \sqrt{b_0 - \left(\frac{b_1}{2}\right)^2} = \omega_n\sqrt{1-\zeta^2}; \\
\gamma &= \sqrt{b_0} = \omega_n;\ \Theta = \tan^{-1}\left(\frac{\sigma}{\omega_d}\right)
\end{aligned}$$

The constants A and φ are determined from initial conditions x_0 and y_0. The transient process and phase portrait of the system (5.11) are shown in Fig. 5.3.

From Fig. 5.3 it follows that damped periodic oscillations of a linear system are represented in the phase plane by a spiral trajectory, where the state moves from initial point $M_0(x_0,y_0)$ towards the origin $x = 0$, $\dot{x} = 0$.

A singular point of the type unstable focus results from a general solution of the equation (5.5) for the case when the roots (5.8) are conjugate complex with positive real parts $b_1^2 < 4b_0$, $b_1 < 0$, $b_0 > 0$ (see region 3 in Fig. 5.1). The general solution is a periodic oscillation with exponentially increasing amplitude, i.e. the linear system is unstable. Phase characteristics of the system are given by expressions (5.11) for $\sigma = b_1/2$. The transient process and phase portrait are shown graphically in Fig. 5.4.

The result is that periodic oscillations of the linear system with increasing amplitude take in the phase plane the form of a spiral trajectory, where the state moves with time from the initial point $M_0(x_0,y_0)$ further from the origin $x = 0$, $\dot{x} = 0$.

A singular point of the type stable node results from equation (5.5) for the case when roots (5.8) are real and negative $b_1^2 > 4b_0$, $b_1 > 0$, $b_0 > 0$ (see region 1 in Fig. 5.1). The general solution is a non-oscillatory transient process, i.e. the linear system is stable. Phase characteristics correspond to the expressions:

$$\begin{aligned}
x(t) &= C_1 e^{-\sigma_1 t} + C_2 e^{-\sigma_2 t} \\
y(t) = \frac{dx(t)}{dt} &= -\sigma_1 C_1 e^{-\sigma_1 t} - \sigma_2 C_2 e^{-\sigma_2 t}
\end{aligned} \tag{5.12}$$

where $\sigma_{1,2} = \lambda_{1,2} > 0$.

Possible forms of state trajectories of the linear system and its phase portraits (5.12) are shown in Fig. 5.5.

In Fig. 5.5a are shown two possible behaviors of the state in time (curves 1 and 2) and corresponding phase trajectories 1 and 2 in Fig. 5.5b. For responses such as those given by curve 1, it can be stated that for all time the sign of $x(t)$ and $y(t)$ does not change, i.e. $x(t) > 0$ and $y(t) = \dot{x}(t) < 0$; $\forall t$. However, for responses represented by curve 2, the sign of $x(t)$ and $y(t)$ will change only once. The boundary lines for trajectories 1 and 2 are defined by the lines in phase plane $y = -\sigma_1 x$ and $y = -\sigma_2 x$, that follows from (5.12) for $\sigma_1 = 0$ and $\sigma_2 = 0$, respectively.

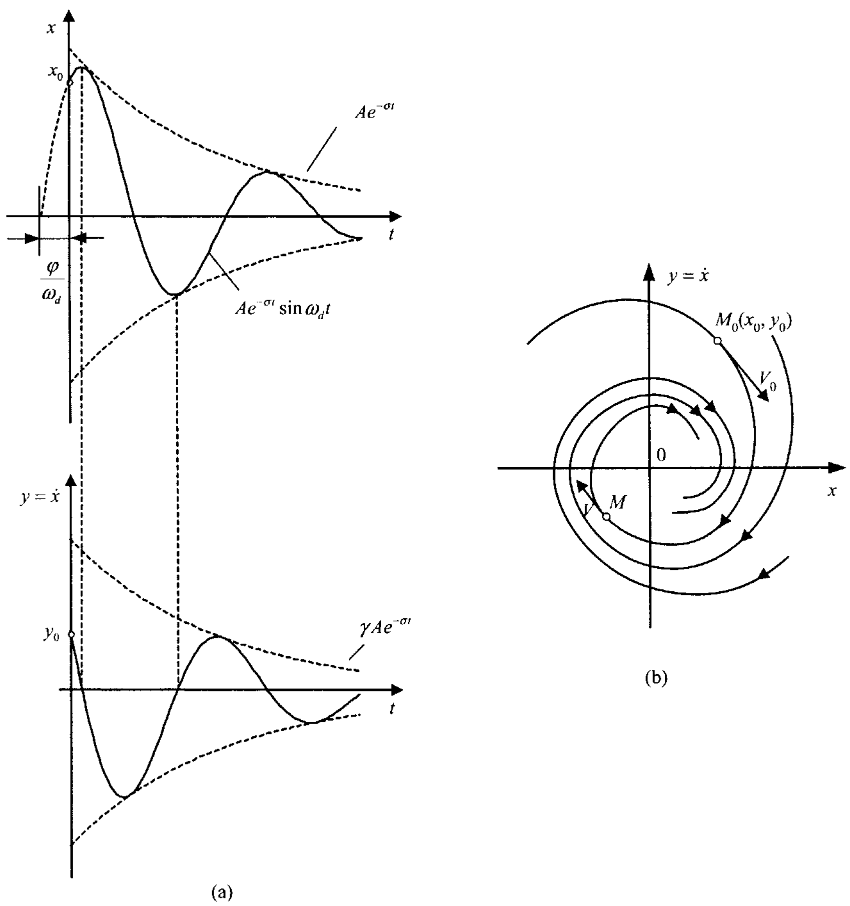


Figure 5.3: The state of a stable oscillatory process (a) and phase portrait for singular point of the type stable focus (b).

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

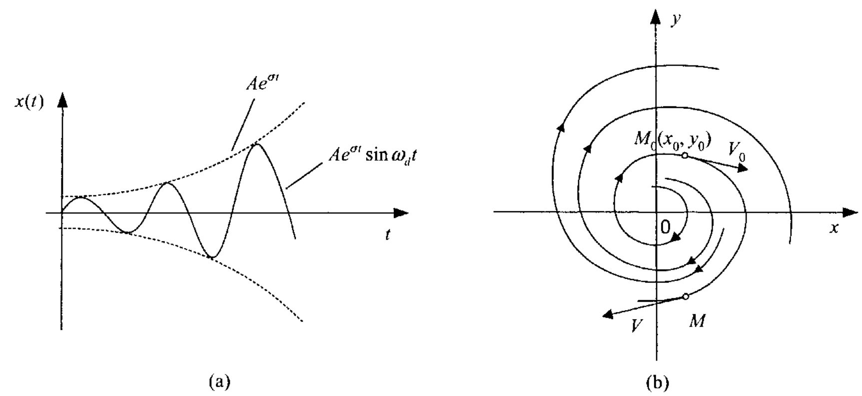


Figure 5.4: Phase trajectory (a) and phase portrait (b) for the singular point of type unstable focus.

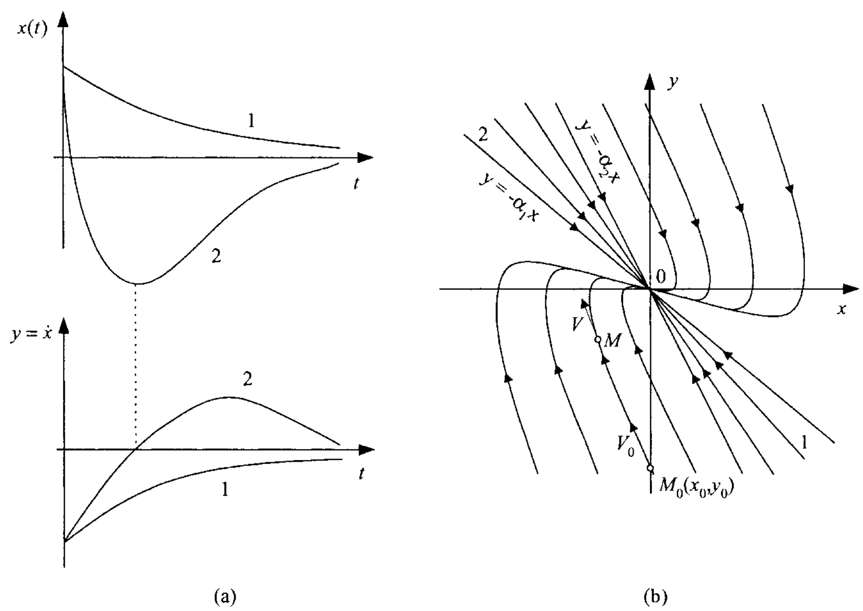


Figure 5.5: Possible trajectories (a) and phase portrait (b) of the linear system with singular point of type stable node.

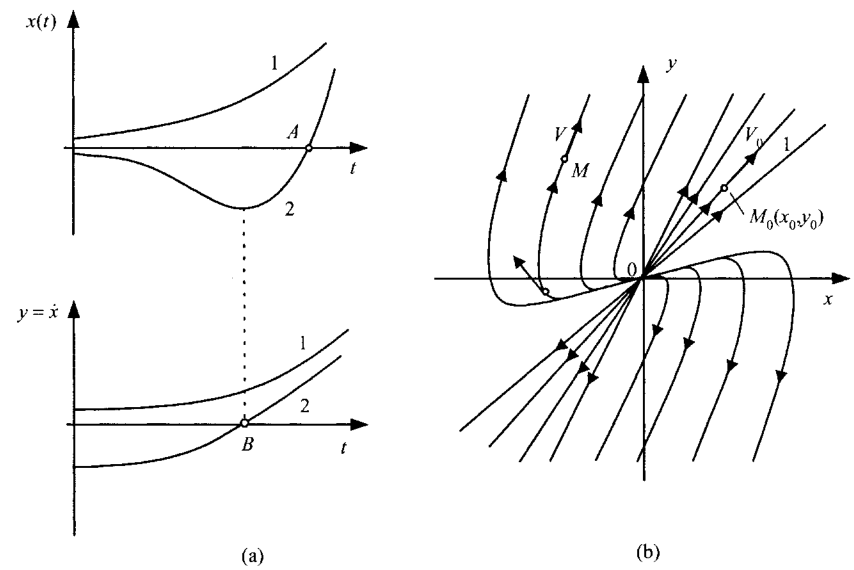


Figure 5.6: Possible phase trajectories (a) and phase portrait (b) of linear system with singular point of type unstable node.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

From Fig. 5.5 follows that the stable non-oscillatory process of a linear system has phase trajectories where the state approaches the coordinate origin from the initial point asymptotically with time t.

A singular point of the type unstable node follows from the general solution of equation (5.5) for the case when the roots (5.8) are real and positive: $b_1^2 > 4b_0$, $b_1 < 0$, $b_0 > 0$ (see region 4 in Fig. 5.1). The general solution is non-oscillatory and increase with time. The response of the linear system is unstable. Phase trajectories of the system are defined by equations (5.12) for $\sigma_1 < 0$ and $\sigma_2 < 0$. The possible forms of the phase trajectories and the phase portrait for the linear system with an unstable node are presented in Fig. 5.6.

The result from Fig. 5.6 is that expanding aperiodic processes of the linear system have in the phase plane phase trajectories where the state starts at the initial point $M_0(x_0,y_0)$ and moves away from the coordinate origin.

A singular point of the type saddle is derived again from the general solution of equation (5.5) for the case that the roots (5.8) are real and of opposite sign. The general solution is a non-oscillatory and increasing process, i.e. the linear system is unstable. Phase trajectories of the system (see region 5 in Fig. 5.1) are given by equations (5.12) for various signs of σ_1 and σ_2. For the case $b_0 < 0$, $b_1 = 0$, $\sigma_1 = \sigma_2 = \sigma = \sqrt{-b_0}$, the phase trajectories are given by:

$$\frac{dy(t)}{dx(t)} = \sigma^2 \frac{x}{y} \tag{5.13}$$

and:

$$\frac{x^2}{c^2} - \frac{y^2}{(\sigma c)^2} = 1 \tag{5.14}$$

Possible forms of the phase trajectories and the phase portrait for the linear system with the singular point of the type saddle are given in Fig. 5.7.

From Figs. 5.6 and 5.7 it follows that non-oscillatory and diverging processes of a linear system are represented in the phase plane by phase trajectories with singular points of the type unstable node or saddle.

5.1.2 Phase Trajectories of Nonlinear Systems

The dynamic behavior of an unforced nonlinear system with the structure as in Fig. 1.2 is described by nonlinear differential equation (1.5) with $f(t) = 0$:

$$A(p)x(t) + B(p)F(x, px) = 0 \tag{5.15}$$

Figure 5.8 shows the essential properties of an unforced nonlinear system in Fig. 1.2, which results from the possible forms of the solution $x(t)$ given in Fig. 5.8. Figure 5.9 shows various phase portraits of nonlinear systems.

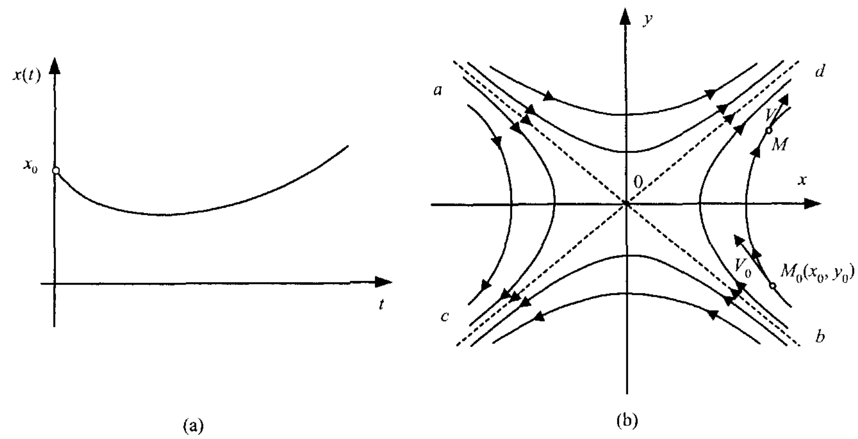


Figure 5.7: Possible form of phase trajectory (a) and phase portrait (b) for linear system with singular point of type saddle.

Contrary to linear systems, where only two possible regions in the parameter plane exist — stability and instability regions—the nonlinear systems can have the following regions in the parameter plane:

1. Region of stability of equilibrium state with constant value of controlled variable,
2. Region of stable self-oscillations,
3. Region of instability of the equilibrium states, and
4. Regions in which more complex dynamic behaviors are exercised.

If the solution of the system has the form as in Fig. 5.8a or the phase portrait as in Fig. 5.9a, the equilibrium state of the system ($x_e = 0$) is unstable. Namely, as the trajectories 1 and 2, independent of the initial conditions, finish in oscillations of constant frequency and amplitude and do not approach the equilibrium state ($x_e = 0$), the equilibrium state cannot be regarded as asymptoticaly stable.

Figure 5.8b shows two cases. In the first case (curve 1), equilibrium state of the system $x_e = 0$ is *locally stable*, as a trajectory which starts from an initial condition $x(0) < a$ in the vicinity of the equilibrium state will finish at the equilibrium state $x(\infty) = x_e = 0$.

In the second case (curve 2), the equilibrium state is *globally unstable* since for all initial conditions which are not in the vicinity of equilibrium state $x_e > a$, the trajectory always increases in amplitude. This is the case of an unstable

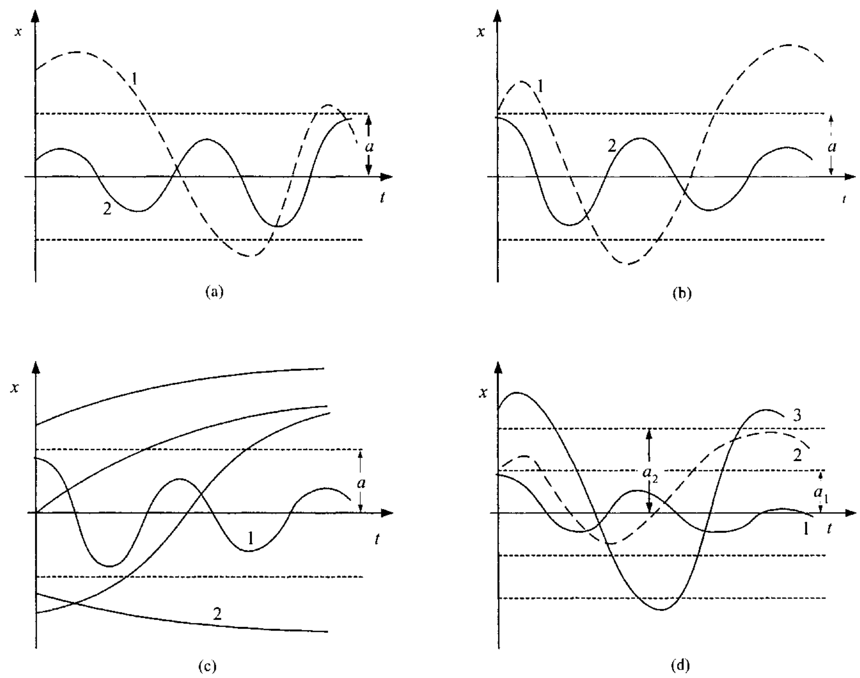


Figure 5.8: Possible forms of solution of an unforced nonlinear system.

periodic process with unstable self-oscillations. The trajectory $x(t)$ departs from the initial periodic process, with either an oscillatory or a non-oscillatory response as a consequence. The trajectory cannot be held at the frequency and the amplitude of initial oscillations (Fig. 5.9b).

In Figs. 5.8d and 5.9c are presented examples of various equilibrium states for a nonlinear system (5.15)—curve 1 shows the behavior of the system with locally stable equilibrium state at the origin $x_e = 0$. Curve 2 is the trajectory for the case of unstable self-oscillations, while curve 3 shows the case of stable self-oscillations. Figure 5.9c shows the situation in phase plane—the equilibrium state of the system $x_e = 0$ will be *locally stable*, but for all initial conditions that are not in the vicinity of the equilibrium state, stable self-oscillations of constant amplitude will be established. This means a *globally unstable* equilibrium state.

It is obvious that besides the mentioned singular points, in nonlinear systems are possible specific phase trajectories, Fig. 5.9.

By comparing phase portraits in Figs.5.9a and 5.4, it is obvious that near the coordinate origin the phase trajectories behave as singular points of the type unstable focus, implying an unstable equilibrium state. Contrary to a linear system, phase trajectories of the nonlinear system can move with time from the coordinate origin to some closed limit trajectory. Other trajectories outside the limit region are asymptotically approaching the limit, too. Phase trajectory in Fig. 5.9a is called *stable limit cycle*, and the solution $x(t)$ is a periodic oscillatory process of constant amplitude and frequency, Fig. 5.8a. Self-oscillations are established in the system with the amplitude $x = A$ and $y = \dot{x} = B$.

The case of local stability and global instability of an equilibrium state, Fig. 5.8b, corresponds to the phase portrait in Fig. 5.9b. Limiting values of initial conditions for which the system is stable correspond in the phase plane to the parameters of an unstable limit cycle. From here the spiral trajectories approach the coordinate origin, the so-called *local stable equilibrium state*, and other trajectories move towards infinity, the so-called *global unstable equilibrium state*.

The phase portrait in Fig. 5.9c corresponds to a stable limit cycle which is far away from the coordinate origin—this produces stable oscillations with large amplitudes (Fig. 5.8c). By comparing the portraits in Figs. 5.9b and 5.9c, it is obvious that with large initial conditions the form of phase trajectories is changing. It is, for example, possible that the non-oscillatory processes pass over to stable oscillations, and vice versa—with decreasing initial conditions a reverse process is possible, Figs. 5.8c and 5.9.

According to linear system theory, for a closed phase trajectory of a regular geometrical form (Fig. 5.2), the edge of the system stability is in the vicinity of the equilibrium state $x_e = 0$. With large initial conditions, i.e. outside the linear region of static characteristics, the form of phase portrait will generally change. One of the possible changes of phase trajectories is depicted in Fig. 5.9d. For relatively small initial conditions, the phase portrait is of the type center, while for

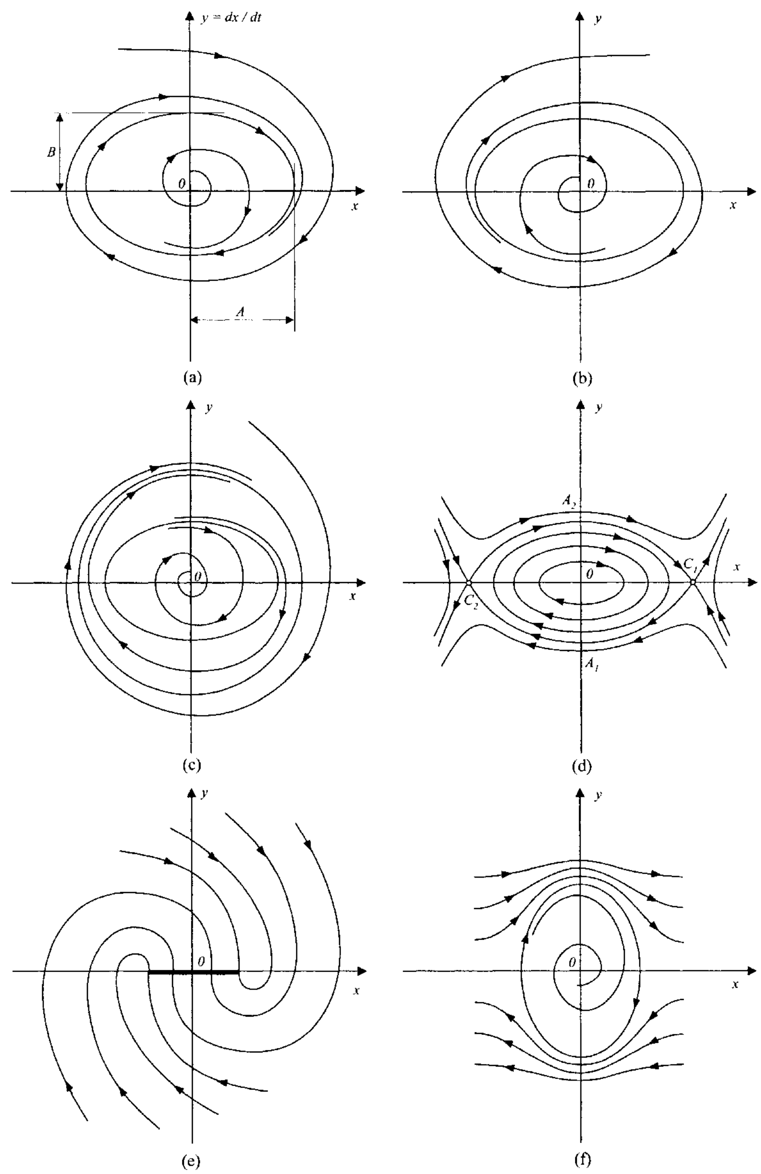


Figure 5.9: The types of phase portraits of nonlinear systems.

large initial conditions two saddles C_1 and C_2 appear, which means an unstable system. Each of the limiting phase trajectories which divide the phase plane into different phase portraits (phase trajectories pass through the saddles C_1 and C_2) is a specific one and is called a *separatrix*. Characteristically, the separatrices do not intersect the phase trajectories (because separatrices themselves are phase trajectories). That means the trajectory starts in the phase plane with certain properties and will not pass over to the other part of the phase plane with different properties.

In analogy with the above discussion, the phase trajectories of nonlinear systems, which for small initial conditions are impossible to analyze as linear structures, have qualitatively the same forms of phase portraits. Among such nonlinear systems are those with relay characteristics and with typical nonlinear elements. In many cases some of mentioned systems with large initial deflections can approach the behavior of corresponding linear systems. For instance, with nonlinear systems of the type dead zone or backlash the system can be treated as a linear one in cases when the input signal amplitude to the nonlinear part is larger than the width of above-mentioned nonlinearities.

In systems with dead zone and Coulomb friction there exist jam areas, i.e. where the equilibrium state $x = 0$ is determined by a series of points on the abscissa near the coordinate origin (Fig. 5.9e).

5.2 Methods of Defining Phase Trajectories

Methods of defining phase trajectories can be analytical, grapho-analytical and numerical.

Analytical Methods

The equation of phase trajectory $\dot{x} = f(x)$ can be determined in the cases when the differential equation has a solution, i.e. when $x(t)$ can be found analytically, and after that $\dot{x}(t)$. By eliminating the parameter t from the functions $x(t)$ and $\dot{x}(t)$ the function $\dot{x} = f(x)$ is obtained.

Let the differential equation be of the form:

$$\ddot{x} + \omega^2 x = 0 \tag{5.16}$$

The solution is:

$$x(t) = A \sin(\omega t + \varphi) \tag{5.17}$$

where A and φ are integration constants which depend upon initial conditions. The derivation of (5.17) gives:

$$\dot{x}(t) = A\omega \cos(\omega t + \varphi) \tag{5.18}$$

By combining (5.17) and (5.18), parameter t is eliminated, and phase trajectory $\dot{x} = f(x)$ is obtained:

$$\frac{x^2}{A^2} + \frac{\dot{x}^2}{A^2\omega^2} = 1 \tag{5.19}$$

In the majority of cases it is simpler to reduce the differential equation of second order to two equations of first order with separated variables $y = \dot{x}$ and x. When solving such equation of first order, the equation of phase trajectories follows. By substituting $y = \dot{x}$ in equation (5.16) we get:

$$\dot{y} + \omega^2 x = 0 \tag{5.20}$$

or:

$$\frac{\dot{y}}{\dot{x}} + \omega^2 \frac{x}{\dot{x}} = 0 \tag{5.21}$$

As $\frac{\dot{y}}{\dot{x}} = \frac{dy}{dx}$, from (5.21) follows:

$$\frac{dy}{dx} + \omega^2 \frac{x}{y} = 0 \tag{5.22}$$

Solution of this equation gives directly the phase trajectories:

$$\begin{gathered} \frac{y^2}{\omega^2} + x^2 = A^2 \text{ , or} \\ \frac{x^2}{A^2} + \frac{\dot{x}^2}{A^2\omega^2} = 1 \end{gathered} \tag{5.23}$$

Generally speaking, a stationary process is described by a differential equation with constant coefficients:

$$\ddot{x}(t) + a\dot{x}(t) + bx(t) = c \tag{5.24}$$

Substituting $y(t) = \dot{x}(t)$ and rearranging gives the equation of first order:

$$\frac{dy(t)}{dt} = c - bx(t) - ay(t) \tag{5.25}$$

To find the phase portrait $y = f(x)$, the variable t must be eliminated from (5.25). By dividing (5.25) with $y = \frac{dx}{dt}$, the equation of phase trajectories can be written:

$$\frac{dy}{dx} = \frac{c - bx - ay}{y} \tag{5.26}$$

or:

$$\begin{aligned} \frac{dy}{dt} &= Q(x,y) \\ \frac{dx}{dt} &= P(x,y) \end{aligned} \tag{5.27}$$

$$\frac{dy}{dx} = \frac{Q(x,y)}{P(x,y)} \tag{5.28}$$

The last equation is the differential equation of phase trajectories, and it determines the slope of the tangent at the point $M(x,y)$ on the phase trajectory.

An analytical solution of equation (5.27) is found only in cases when separation of variables is possible. In all other cases the solution of (5.27) is sought by graphical or numerical methods.

The equilibrium state of the system corresponds to points in the phase plane where the velocity and acceleration of coordinate x equal zero:

$$\begin{aligned} y = \frac{dx}{dt} &= 0 \\ \frac{dy}{dt} = \frac{d^2x}{dt^2} &= 0 \end{aligned} \tag{5.29}$$

or:

$$\begin{aligned} P(x,y) &= 0 \\ Q(x,y) &= 0 \\ \frac{dy}{dx} &= \frac{0}{0} \end{aligned} \tag{5.30}$$

Therefore, the equilibrium state is determined by equations (5.29) and (5.30). The points in the phase plane where the system is at the equilibrium state are called *singular points*. There are four types of them: *center, focus, node and saddle.*

Linear systems have only one equilibrium state. If the functions $P(x,y)$ and $Q(x,y)$ are linear, by satisfying condition (5.30) as the solution of the system, only one singular point N (Fig. 5.10a) is obtained. However, if the functions $P(x,y)$ and $Q(x,y)$ are nonlinear, by fulfilling the condition (5.30) several solutions are possible, i.e. there are several singular points (Fig. 5.10b).

Grapho-Analytical Methods

Among the most-used grapho-analytical[6] methods for constructing phase portraits are the *isocline method* and the *δ-method.* Isoclines are curves which pass through the points along the phase trajectories in which tangents on phase trajectories have the same slopes. They are found from the equation of phase trajectories (5.27). By putting into (5.27) $dy/dx = N = const.$, an equation with two unknowns x and y is obtained. For $N = N_1$ the equation of the isocline is:

$$\frac{dy}{dx} = N_1 = \frac{Q(x,y)}{P(x,y)} \tag{5.31}$$

[6]In situations when the mathematical model is not known, and only a graphical display of the phase trajectory is at our disposal (obtained, e.g., by experiment), the grapho-analytical methods must be applied.

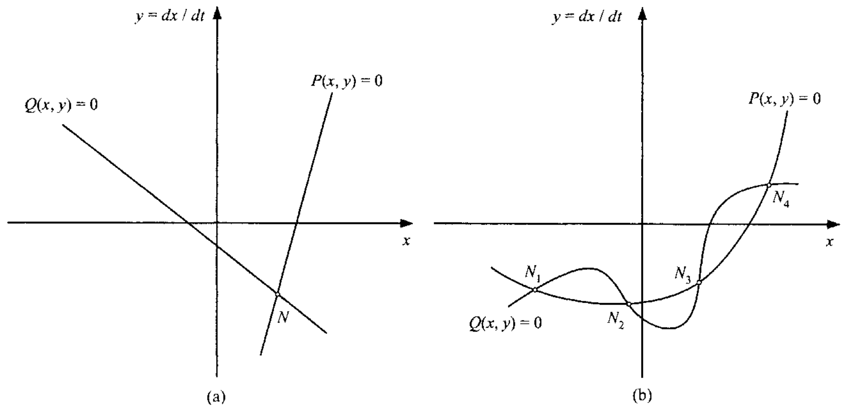


Figure 5.10: Singular point N of linear system (a), and singular points of nonlinear system (b).

with the corresponding curve in the phase plane. By setting various constant values $\frac{dy}{dx} = N_1$, $\frac{dy}{dx} = N_2, \ldots,$ $\frac{dy}{dx} = N_n$ a family of isoclines in the phase plane is found.

With the aid of the family of isoclines (Fig. 5.11), the phase trajectory with the initial point M_0 can be graphically constructed. Let the initial point be on the isocline N_1. The construction of a phase trajectory starts from an arbitrarily chosen initial point M_0. On each of the isoclines an arrow marks the slope of the phase trajectories at the intersection point with the isocline.

In order to construct the trajectory between two adjoining isoclines, from the point M_0 two straight lines with slopes N_1 and N_2 are drawn in the direction of the isocline N_2. As the trajectory at point M_0 has slope N_1, and the next isocline must intersect the trajectory with the slope N_2, the intersection point with isocline N_2 must be inside the angle of the lines N_1 and N_2. It can be approximately taken that the point B, the intersection of the isocline N_2 and the phase trajectory, is situated in the middle of the isocline section between the lines N_1 and N_2. An analogous procedure is valid for the points C and D, as well as for all other points. By choosing other initial points on the same isocline, the other phase trajectories are found. The isocline method is appropriate to find phase portraits when initial conditions for a given section are known.

The basis of the δ-method is to rearrange the second-order nonlinear differential equation to a specific form. Every second-order differential equation can be written as:

$$\ddot{x} + f(t, x, \dot{x}) = 0 \tag{5.32}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

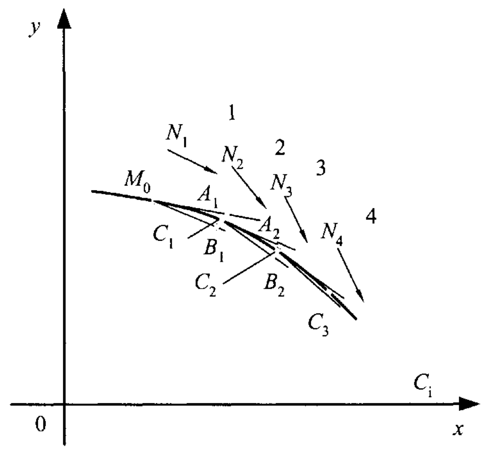


Figure 5.11: Construction of phase trajectories by isocline method.

The function f may not be a function of all variables x, $\dot{x}$ and t. Equation (5.32) is transformed to:

$$\ddot{x}+A^2\left[x+F\left(t,x,\dot{x}\right)\right] \tag{5.33}$$

where $A=\omega_n$ is a constant. Substituting the following into (5.33):

$$F\left(t,x,\dot{x}\right)=\delta \tag{5.34}$$

yields:

$$\ddot{x}+\omega_n^2\left(x+\delta\right)=0$$

or:

$$\ddot{x}=-\omega_n^2\left(x+\delta\right) \tag{5.35}$$

By dividing both sides of the expression (5.35) by $\dot{x}=\frac{dx}{dt}$, one gets:

$$\frac{d\dot{x}}{dx}=-\frac{\omega_n^2\left(x+\delta\right)}{\dot{x}} \tag{5.36}$$

By normalizing the ordinate of the phase the plane to $y=\frac{\dot{x}}{\omega_n}$ equation (5.36) is further transformed:

$$\frac{d\left(\frac{\dot{x}}{\omega_n}\right)}{dx}=-\frac{x+\delta}{\frac{\dot{x}}{\omega_n}} \tag{5.37}$$

With equation (5.37), the slope of the phase trajectory is defined in every point of the phase plane. The δ-*method* of constructing the phase trajectory boils down to

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

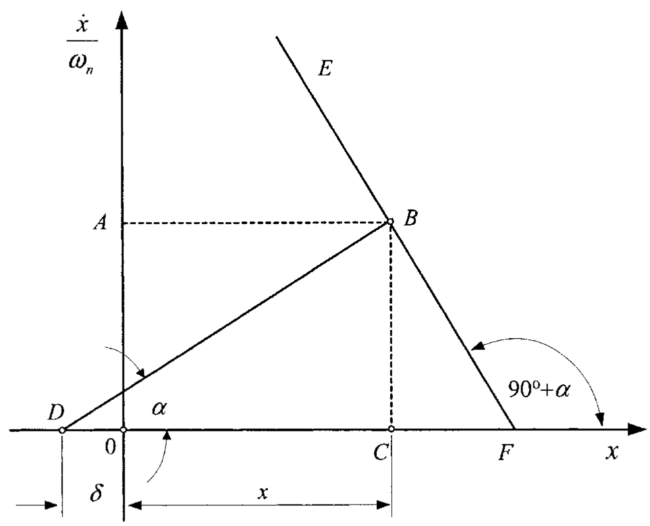


Figure 5.12: Construction of phase trajectory by δ-method.

determining (resolving) the value of δ_i and the slope of the tangent for the point B_i at the phase trajectory, Fig. 5.12.

The points on phase trajectory with the *δ-method* are determined by the following procedure: In the phase plane, the starting point B_0 is defined by initial conditions. Then from (5.34) the corresponding δ_0 is calculated and the point D_0 plotted on the abscissa axis. From this point the circle is plotted through the point B_0 on the phase trajectory. Close to that point on the circle, the next point B_1 is chosen with the coordinates $\left(\frac{\dot{x}_1}{\omega_n}, x_1\right)$. These coordinates put in (5.34) will result in δ_1 which defines the point D_1 on the x-axis. From that point the next circle is drawn through the point B_1. The procedure continues with defining the next point on the circle close to the point B_1 and so on until enough points $B_0, B_1, B_2, \ldots, B_i$ are found to plot the phase trajectory.

Method of Adjoint Solutions

When the systems contain elements with relay or piecewise linear characteristics in general, phase trajectories are constructed from trajectory parts which correspond to linear parts of nonlinear characteristic. Those individual parts of the nonlinear characteristics are described by linear equations and thus can be easily solved. The overall characteristic of a typical nonlinear element can be de-

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

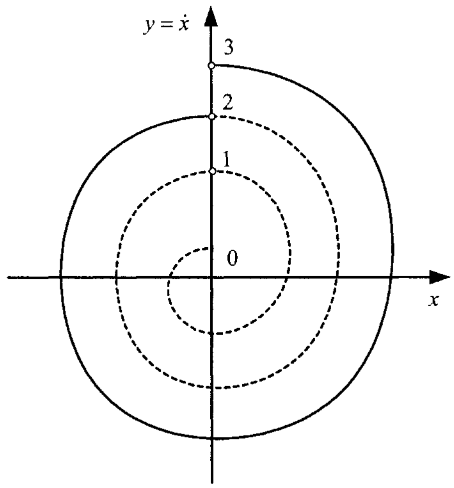


Figure 5.13: Phase trajectory of a nonlinear system.

scribed by a set of linear equations which are joined at certain points. Namely the construction of phase trajectories of nonlinear systems with typical nonlinearities consists of finding the trajectories of the corresponding linear systems and joining them at transition points. Such an approach is called the *adjoint solution method.* Each part of the phase plane which corresponds to one linear part of the nonlinear characteristic is called the *sheet* of this part of the characteristic. According to the number of linear parts of the nonlinear characteristic, the number of sheets of the phase plane is given. So we speak of *two-sheet*, *three-sheet* and *multi-sheet* phase planes. The transfer from one sheet to another can be continuous or discontinuous, which depends upon the differential equation of the system. Various sheets can be extended one to another (single-valued nonlinearities) or folded over from one to the next one (multi-valued nonlinearities).

For systems with smooth nonlinearities, graphical procedures are appropriate to determine the phase trajectories. However, for systems with relay and piecewise linear characteristics, the equation of the phase trajectory (5.27) can be determined for every sheet of the phase plane. The transient response of the system is found analytically by time intervals—in order to solve the differential equation of the next interval, the final values of the previous interval serve as initial conditions. Mapping phase trajectory points enables us to determine the periodic operating modes of systems with relay and piecewise linear characteristics without calculating integral curves—phase trajectories. In essence, difference equations are to be found which link the values of the nonlinear system variables at the boundaries of the linear parts of nonlinear characteristics.

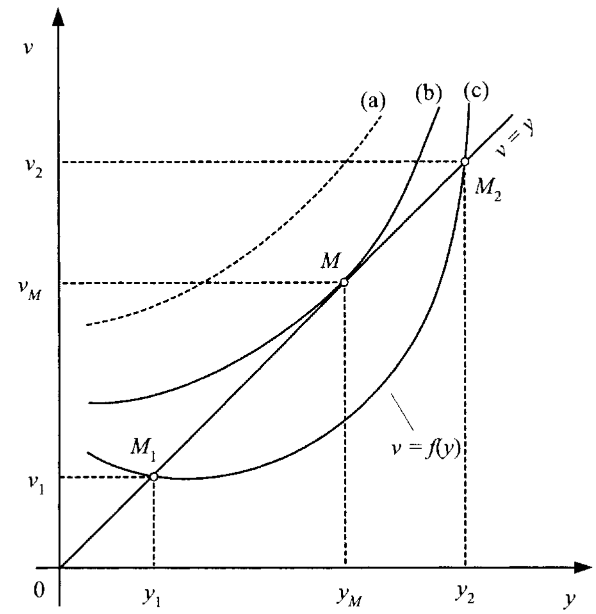


Figure 5.14: Possible form of function $v = f(y)$.

Let us see the dynamics of the nonlinear system given by the trajectory in Fig. 5.13. In a specific moment, the dynamics of the system are given by point 1 on the y-axis. After a certain time point 1 will reach point 2, then point 3 and so on. We can say that after some time all the points of the positive y-axis will be mapped to some other points on the same axis. If points circulate around the coordinate origin, the positive y-axis maps into itself. If the same point is copied into itself (*immobile point*), that means that the closed phase trajectory passes through this point, i.e. it is a singular point (*limit cycle*).

It can be concluded that for finding the parameters of self-oscillations, it is sufficient to determine mapped points of a half-line starting at the singular point of the phase plane.

The coordinates of point 2 on the phase trajectory Fig. 5.13 depend only on coordinates of point 1. Namely, the mapping of half-line $0y$ into itself, the coordinate v of the mapped point 2 depends on the coordinate y of point 1, $v = f(y)$. A possible form of this function is shown in Fig. 5.14.

The line $v = y$ in Fig. 5.14 is the *bisectrix* of the system. The intersection points of the line $v = y$ with the curve $v = f(y)$ are fixed points of the mapping, i.e. singular points which determine the limit cycle. The part of the curve $v = f(y)$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

lying below bisectrix $v = y$ is characterized by a damped transient response in the system, while the part of the curve above the bisectrix $v = y$ is characterized by an undamped transient response in the system. A stable limit cycle (stable equilibrium state) corresponds to the intersection point of the curve $v = f(y)$ and the bisectrix $v = y$, where the slope of the line $v = y$ is greater than the slope of the tangent to the curve $v = f(y)$. To the unstable limit cycle (unstable equilibrium state) corresponds the intersection points where the slope of the curve $v = f(y)$ is greater then the slope of bisectrix $v = y$. Coming back to Fig. 5.14, a limit cycle with oscillation amplitude y_1 represents stable self-oscillations, while the limit cycle with oscillation amplitude y_2 represents an unstable equilibrium state or unstable oscillations of the system, respectively.

In the case when bisectrix $v = y$ touches the curve $v = f(y)$ (point M on curve b, Fig. 5.14), the limit cycle is semistable. For initial values $v < v_M$, the system will sustain stable self-oscillations with the amplitude $y > y_M$, while for the case $v > v_M$, unstable self-oscillations follow with amplitudes $y > y_M$. The function $v = f(y)$ is found by integrating differential equations of the system, similarly to determining phase trajectories. With relay systems (on-off elements) the procedure is more appropriate if for the mapping line the line which determines the operating mode of the relay is taken.

5.2.1 Estimation of Stability and Performance by Means of Phase Trajectories

A phase portrait enables us to find the dynamics of the system for all possible[7] initial conditions. If for given initial conditions the phase trajectory of the system in consideration tends toward the coordinate origin, the equilibrium state is stable. On the other hand if the phase trajectory ends in a limit cycle or even at infinity, the equilibrium state of the system is unstable. The structure of the phase portrait accurately determines the region of initial deflection for which the system has different dynamic properties. The limit cycles and separatrices are the boundaries for the regions of different dynamic behavior.

Phase trajectories enable graphical displays of transient response which arise from initial conditions other than zero. The graphical procedure is based on the properties of phase trajectories. The phase trajectory is a graphic interpretation of:

$$y = \frac{dx}{dt} = f(x) \tag{5.38}$$

In order to find transient response from the graph in Fig. 5.15, it is necessary to find t. Fig. 5.15 illustrates the procedure.

[7] For a given system.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

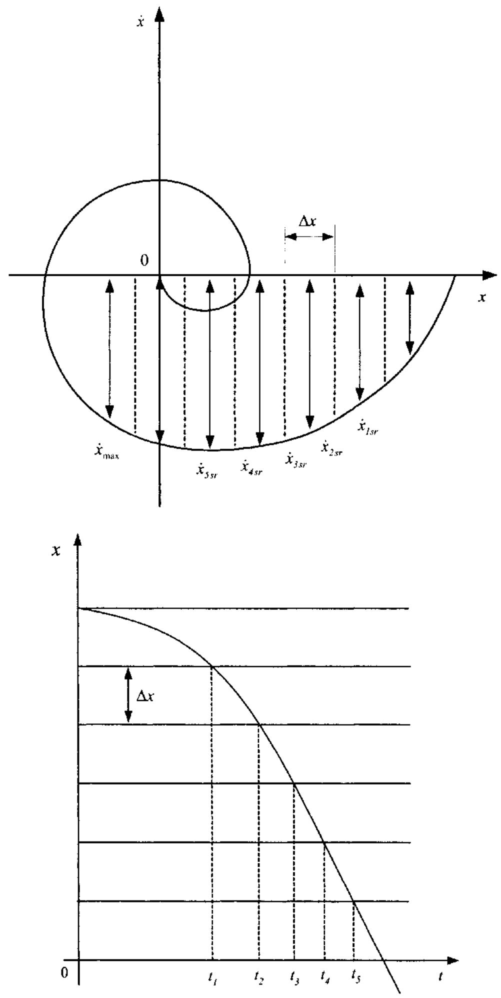


Figure 5.15: Graphical procedure to find a transient process from phase trajectory.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

The time t which corresponds to every point on the phase trajectory is obtained by the approximate expression:

$$\Delta t_i = \frac{\Delta x_i}{\Delta y_{i,av}} = \frac{\Delta x_i}{\Delta \dot{x}_{i,av}} \tag{5.39}$$

The region of the phase plane where this part of phase trajectory lies, and for which time must be found, is divided by vertical lines of width Δx. In the middle of every Δx, starting from the point $\Delta(x_0, y_0)$, $x_{i,av}$ and the corresponding value of $\dot{x}_{i,av}$, are found and by means of equation (5.39) Δt_i is calculated. On the t-axis (Fig. 5.15b) time $t_i = t_{i-1} + \Delta t_i$ is subsequently defined while on the x-axis the corresponding coordinate x is drawn (Fig. 5.15a).

The time of the transient response is found from the equation:

$$y = \dot{x} = \frac{dx}{dt} \tag{5.40}$$

or:

$$t = \int_{x_i}^{x_k} \frac{1}{y} dx \tag{5.41}$$

In order to find time t from the above expression, the graphical display of the function must be determined:

$$\frac{1}{y} = f(x) \tag{5.42}$$

The area below the curve $f(x)$ is equal to the time of the transient response for the region in consideration (Fig. 5.16). By means of the phase trajectories it is simple to find the maximum values of $x(t)$ and $\dot{x}(t)$, i.e. the values are directly read from graphical plots.

The method of phase trajectories is applied when the dynamic behavior of an unforced system with initial values other than zero is sought. By substituting the variables, it is possible to reduce to unforced systems the forced systems with a constant input variable, or with the input variable with constant velocity.

EXAMPLE 5.1

The system shown in Fig. 5.17 is first excited by a unit step function $r(t) = S(t) = 1, \forall t > 0$, and then by a ramp function $r(t) = t, \forall t = 0$.

Dynamic behavior of the linear system can be determined by variable $e(t)$. The differential equations of the system are:

$$\begin{aligned} T\ddot{y}(t) + \dot{y}(t) &= Ke(t)\,; \\ y(t) &= r(t) - e(t) \end{aligned} \tag{5.43}$$

$$T\ddot{e}(t) + \dot{e}(t) + Ke(t) = T\ddot{r}(t) + \dot{r}(t) \tag{5.44}$$

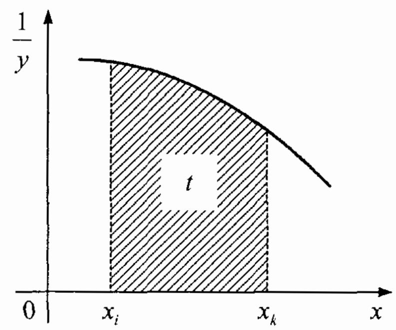


Figure 5.16: The duration of transient process.

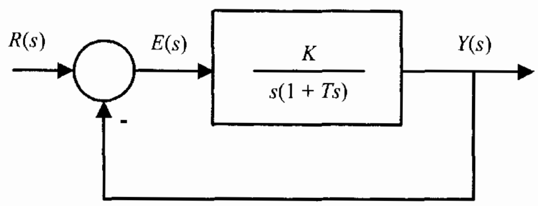


Figure 5.17: Block diagram of linear system.

If $r(t) = S(t)$, initial conditions relative to the variable $e(t)$ are:

$$e(0) = 1; \dot{e}(0) = 0 \tag{5.45}$$

The block diagram of the forced system in Fig. 5.17, with input $r(t) = S(t)$, is reduced to an unforced system with initial conditions (5.45) by choice of the variable $e(t)$. In the case of a ramp function input, $r(t) = t$, the equation (5.44) has the form:

$$T\ddot{e}(t) + \dot{e}(t) + Ke(t) = 1 \tag{5.46}$$

By introducing a new variable:

$$e_1(t) = e(t) - \frac{1}{K} \tag{5.47}$$

equation (5.46) becomes homogenous:

$$T\ddot{e}_1(t) + \dot{e}_1(t) + e_1(t) = 0 \tag{5.48}$$

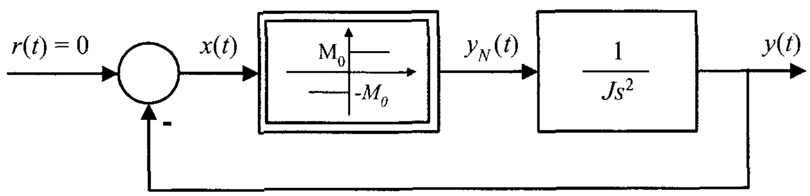


Figure 5.18: Block diagram of nonlinear system.

with initial conditions $e_1(0) = -1/K, e(0) = 0, \dot{e}_1(0) = \dot{e}(0) = 1$. *The analysis of the forced system can be carried out by determining the phase trajectory of the unforced system with the new variable* $e(t)$ *or* $e_1(t)$ *for unit step or ramp excitation, respectively.*

5.3 Examples of Application of Various Methods to Obtain Phase Trajectories

EXAMPLE 5.2 (PHASE TRAJECTORIES BY AN ANALYTICAL METHOD)
For the system shown in the block diagram (Fig. 5.18), the phase trajectory is to be found.

Solution. *The differential equation of the unforced system in the Laplace domain in relation to coordinate x is:*

$$y = \frac{1}{Js^2} \cdot y_N = \frac{1}{Js^2} \cdot M_0 \text{sign}\, x = -x \tag{5.49}$$

or in the time domain:

$$J\frac{d^2x}{dt^2} + M_0 \text{sign}\, x = 0 \tag{5.50}$$

By normalizing the above equation and introducing $\tau = t\sqrt{M_0/J}$, *we will get:*

$$\frac{d^2x}{d\tau^2} = \text{sign}\, x \tag{5.51}$$

By substituting $y = \frac{dx}{dt}$ *in (5.51), a system of two first-order differential equations is obtained:*

$$\frac{dx}{d\tau} = y$$

$$\frac{dy}{d\tau} = -\text{sign}\, x \tag{5.52}$$

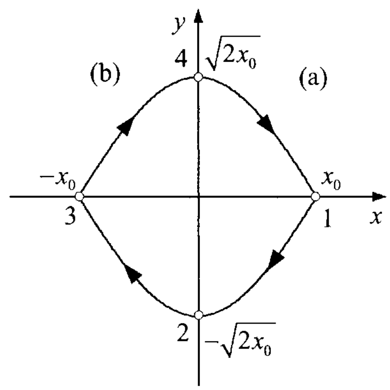


Figure 5.19: Phase trajectory.

Elimination of τ from (5.52) yields:

$$ydy = -\text{sign}x\, dx \tag{5.53}$$

Integration of (5.53) gives the resulting phase trajectories:

$$\frac{y^2}{2} = \mp x + c \tag{5.54}$$

The integration constant is determined by initial conditions $t = 0$, $y(0) = 0$, $\dot{y}(0) = 0$, $x(0) = x_0 = c$. The phase plane has two sheets, which reflect the state of the relay: sheet (a) with the equation $\dot{y} = -1$ and sheet (b) with the equation $\dot{y} = 1$.

The starting point of the trajectory in Fig. 5.19 is on the sheet (a) at the point $x = x_0$. Motion of the point in sheet (a) is given by the trajectory equation $y^2 = 2(x - x_0)$. Point 2 on the trajectory has coordinates $x = 0$, $y = -\sqrt{2x_0}$. The equation of the trajectory for sheet (b) is $y^2 = -2(x - x_0)$. By analogy, points 3 and 4 of the trajectory are found. Time for the motion along the trajectory from point 1 to point 2 is obtained by integrating (5.52) within the boundaries $y = 0$ and $y = -\sqrt{2x_0}$ and it is $\tau_a = \sqrt{2x_0}$.

The phase trajectory from Fig. 5.19 is a parabola symmetrical to the y-axis, i.e. a closed curve (limit cycle) of self-oscillations with amplitude $x_{max} = x_0$ and $y_{max} = \sqrt{2x_0}$ and oscillation period $T = 4\sqrt{2x_0}$. ■

EXAMPLE 5.3 (PHASE TRAJECTORIES BY A GRAPHO-ANALYTICAL METHOD) *For the system described by the equation:*

$$\ddot{x}(t) + x(t) = 0 \tag{5.55}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

the phase trajectory is to be found by a grapho-analytical method.

Solution. *By substituting $\dot{x} = y$ in (5.55), a system of two first-order differential equations is obtained:*

$$\frac{dx}{dt} = y$$
$$\frac{dy}{dt} = -x$$

By eliminating parameter t, the differential equation of the phase trajectories is obtained directly:

$$\frac{dy}{dx} = -\frac{x}{y}$$

If in the above equation $\frac{dy}{dx} = const = N$ is set, the equation of isoclines results:

$$y = -\frac{1}{N}x$$

By varying N from zero to infinity, a family of straight lines in the (x, y)-plane can be drawn, Fig. 5.20. The arrows which indicate the slope of the trajectory at any intersection point with the isoclines are perpendicular to the isoclines, i.e. the slope of the isoclines and of the trajectory are determined by N, so all the phase trajectories are circles in the (x, y) phase plane.

If the system is described by the equation:

$$\ddot{x}(t) + \omega_0^2 x(t) = 0$$

the equation of phase trajectories is:

$$\frac{dy}{dx} = -\frac{\omega_0^2 x}{y}$$

and the equation of isoclines is:

$$y = -\frac{\omega_0^2}{N}x$$

As is obvious from the last equation, in this case the family of isoclines is defined by straight lines through the coordinate origin, with slopes determined by ω_0^2. For $\omega_0^2 \neq 1$, the phase trajectory will take the form of an ellipse, with the major half-axis on the y-axis for $\omega_0^2 > 1$, and on the x-axis for $\omega_0^2 < 1$.

■

EXAMPLE 5.4
(PHASE TRAJECTORIES BY A METHOD OF ADJOINT SOLUTIONS)

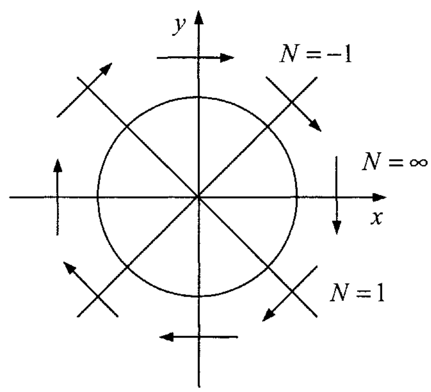


Figure 5.20: Phase trajectories by isocline method.

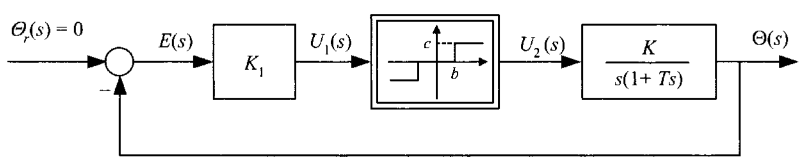


Figure 5.21: Block diagram of a servo system.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Find the phase trajectory of a relay servo system given by the block diagram in Fig. 5.21 in the plane $(\Theta, \dot{\Theta})$.

The system parameters from Fig. 5.21 are $K_1 = 0.065[V/^0]$; $K = 5.06[^0/Vs]$; $c = 6[V]$, $b = 0.1[V]$. *Initial conditions are* $\Theta(0) = 8[^0]$; $\dot{\Theta}(0) = 0\ [^0/s]$, *which means that the motor shaft is initialy turned for* $8[^0]$, *and is not rotating.*

Solution. *In order to solve this problem we have to link variables* Θ *and* $\dot{\Theta}$ *through the closed-loop system differential equation. The differential equation linking the coordinates* Θ *and* u_2 *is:*

$$T\frac{d^2\Theta}{dt^2} + \frac{d\Theta}{dt} = Ku_2 \tag{5.56}$$

or:

$$\frac{d^2\Theta}{dt^2} + \frac{1}{T}\frac{d\Theta}{dt} = \frac{K}{T}u_2 \tag{5.57}$$

By substituting $\Theta = x, \dot{\Theta} = \dot{x}$ *in equation (5.57), the equation in the phase plane results:*

$$\frac{d^2x}{dt^2} + \frac{1}{T}\frac{dx}{dt} = \frac{K}{T}u_2 \tag{5.58}$$

In the plane (u_1, u_2), *the lines of established relay contact are given by the straight lines* $u_1 = \pm b = 0.01[V]$. *In the plane* $(x, \dot{x})$ *such lines are determined by the function* $u_1 = f(\Theta) = f(x)$.

The voltage at the winding of the polarized relay of the system is:

$$u_1 = K_1 e = -K_1\Theta = -K_1 x = -0.065x \tag{5.59}$$

The relay is switched at the voltages $u_1 = 0.1[V]$ *and* $u_1 = -0.1[V]$. *From equation (5.59), the transformed expressions* u_1 *in the plane* $(x, \dot{x})$ *are derived:*

$$x = -\frac{0.1}{0.065} = -1.54[^0];\ x = 1.54[^0]$$

These are straight lines in the plane $(x, \dot{x})$, *parallel to the ordinate* $y = \dot{x}$. *With regard to the lines of established relay contact, the phase plane* $(x, \dot{x})$ *consists of three sheets (Fig. 5.22).*

Sheet 1: $u_1 \geq 0.1[V]$; $x = -1.54[^0]$; $u_2 = 6[V]$

Sheet 2: $|u_1| < 0.1[V]$; $|x| < 1.54[^0]$; $u_2 = 0[V]$

Sheet 3: $u_1 \geq 0.1[V]$; $x = 1.54[^0]$; $u_2 = -6[V]$

Equation (5.57) is solved by integrating along the linear parts, i.e. by sheets 1, 2 and 3 of the phase trajectory.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

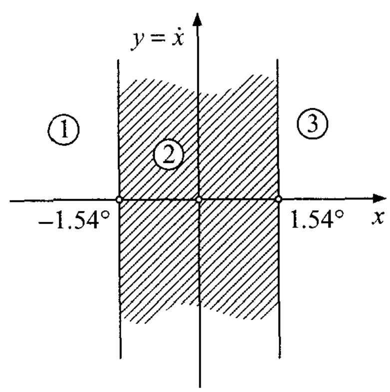


Figure 5.22: Sheets in phase plane.

For sheet 1 in the phase plane $u_2 = 6[V]$, *and by inserting K and T in (5.58) as well as* $y = \frac{dx}{dt}$, *the equation of the phase trajectory is:*

$$\frac{dy}{dx} = \frac{-2y+60.7}{y}$$

with the solution:

$$x = 15.1 - 0.5y - 15.1\ln(60.7 - 2y) + C_1 \tag{5.60}$$

For sheet 2, $u_2 = 0[V]$, *and equation (5.58) takes the form:*

$$\frac{dy}{dt} + \frac{1}{T}y = 0 \tag{5.61}$$

By substituting in (5.61) $y = \frac{dx}{dt}$ *and by eliminating the variable t, the equation of the phase trajectory becomes:*

$$\frac{dy}{dx} = -\frac{1}{T}$$

The solution of this equation is:

$$y = -\frac{1}{T}x + C_2 = -2x + C_2$$

For sheet 3, $u_2 = -6[V]$, *and by analogy with the expression (5.60), the equation of the phase trajectory is:*

$$\frac{dy}{dx} = \frac{-2y-60.7}{y}$$

with the solution:

$$x = -15.1 - 0.5y + 15.1\ln(60.7 + 2y) + C_3 \tag{5.62}$$

Integration constants C_1, C_2 and C_3 are determined by initial conditions, i.e. by coordinates of the starting points of the phase trajectories at each sheet of the phase plane.

The phase trajectory starts at sheet 3 in the point given by initial conditions $x(0) = 8[^0]$ and $\dot{x}(0) = 0[^0/s]$. By inserting these values in equation (5.62), $C_3 = -38.9$ is obtained. Thus, the equation of the phase trajectory for sheet 3 of the phase plane reads:

$$x = -54 - 0.5y + 15.1\ln(60.7 + 2y) \tag{5.63}$$

The boundary point of the trajectory, i.e. the point at the boundary of sheet 3 and sheet 2 is determined by using equation (5.63); the values are $x = 1.54$, $y = -21$. This point determines initial conditions of the trajectory in sheet 2 of the phase plane. By determining the constant $C_2 = -17.9$, the equation of the phase trajectory in sheet 2 is a straight line:

$$y = -2x - 17.9$$

The boundary point of the trajectory between sheet 2 and sheet 1 is found by equation (5.60): $x = -1.54$, $y = -14.8$. The equation of the phase trajectory in sheet 1 is:

$$x = 59.1 - 0.5y - 15\ln(60.7 - 2y) \tag{5.64}$$

The procedure continues until the phase trajectory is within sheet 2, i.e. inside the dead zone of the relay element (Fig. 5.23).

■

5.4 Conclusion

Phase trajectories have always been very instructive for anyone attempting to analyze more closely a particular system. This chapter has given various methods of constructing them. Analytical, grapho-analytical and the method of adjoint solutions were described. Those methods as well as the simulation method[8] are at our disposal for the purpose of obtaining deeper insight into dynamics of a nonlinear system. However, today we are witnessing that in practice phase trajectories are losing the utility which they previously had, when digital computers were not so common. Despite that fact their role at least for educational purposes is very important, because the qualitative structures (local or global) of a nonlinear system are easily obtained through the phase portrait.

[8]Today almost exclusively used for that purpose. The simulation method was not described here because we believe that the reader already knows how to use the simulation package.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

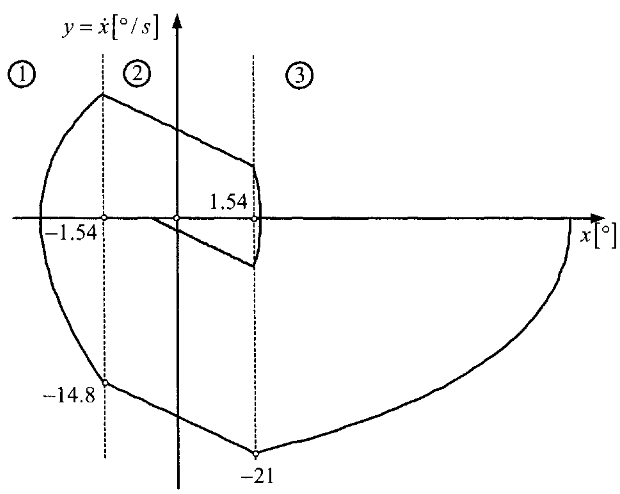


Figure 5.23: Phase trajectory obtained by method of adjoint solutions.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Chapter 6

Harmonic Linearization in Dynamic Analysis of Nonlinear Control Systems Operating in Stabilization Mode

In this chapter the harmonic linearization is used for dynamic analysis of nonlinear control systems operating in stabilization mode. This mode of operation is characterized by the fact that some desired operating point has to be maintained over a long period of time. Many nonlinear control systems, especially in process industries, are normally working in this mode of operation. Two cases are treated here: analysis of an unforced nonlinear control system experiencing self-oscillations and analysis of nonlinear control system forced to oscillate with desired frequency. It should be pointed out that it is very common for nonlinear control systems working in stabilization mode to oscillate around a desired operating point.

6.1 Describing Function in Dynamic Analysis of Unforced Nonlinear Control Systems

Analysis and synthesis of *linear* control systems by frequency methods have certain advantages which made them well accepted in engineering practice. Namely, the analysis of a control system such as stability analysis, filtering analysis, etc., as well as design of a controller (or compensator) by some design method is reduced to the relatively simple analysis of a group of *algebraic equations* in the frequency domain. Moreover, the possibility of graphical representation enables simplification of analysis and synthesis due to the fact that a clear relation between the frequency response and physical properties of the system can be established. No less important is the fact that the complexity of the method is not much greater with the increase of the system's order. However, the existing methods of analysis in the frequency domain are not directly applicable to a nonlinear system, since the frequency characteristic for a nonlinear system cannot be defined. For some nonlinear systems, which must fulfil certain conditions, an expanded version of the frequency method, the so-called *describing function method*, can be used for *approximate* analysis and prediction of the behavior of a nonlinear control system. This method is widespread in engineering practice as it possesses all the advantages of the frequency method for a linear system. In the absence of some general method for all nonlinear systems, this method has a general character, and if some necessary conditions are fulfilled the describing function method can give satisfying predictions of the system's behavior.

Necessary conditions for the application of the describing function method as an approximate method in the dynamic analysis of nonlinear control systems are:

1. The structure of the nonlinear system which consists of nonlinear and linear parts must be reduced to the structure such as in the block diagram in Fig. 6.1.

2. The linear part of the system G_L must have low-pass[1] filter properties or that of a resonant filter which is tuned to the first harmonic frequency of the input signal to the nonlinear part.

Besides these, one other condition exists which is not necessary, but is present in derivation of self-oscillations:

3. The nonlinear control system is unforced, and is in a stabilizing mode. This condition will be rejected later in the text[2].

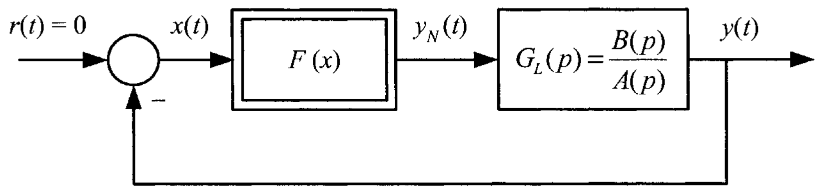
This method is worked out in Goldfarb (1947, 1965); Solodovnikov (1969); Popov and Pal'tov (1960); Petrov et al. (1967); Vidasyagar (1993); Slotine and Li (1991); and others.

An unforced nonlinear control system is seen in the situation when the regulator problem of the feedback system has to be solved, i.e. when the control system is in stabilizing mode. Namely, a large number of automatic control systems[3] work in a stabilizing mode, when it is desired that the system works in a nominal

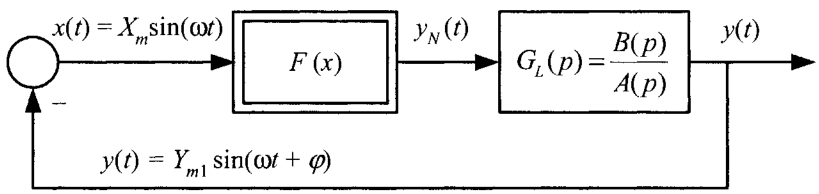
[1]The term low-pass filter must be taken conditionally. As it is quite possible for high-frequency oscillations to arise, the essential thing is that the linear part doesn't pass the harmonics higher than the first, nothwithstanding if the first harmonic has a low or high frequency.

[2]See Section 6.1.5 and 6.2.

[3]Especially in process control (chemical, petrochemical, pharmaceutical, etc.).



(a)



(b)

Figure 6.1: (a) Block diagram of an unforced nonlinear control system, (b) established monoharmonic (self-) oscillations.

operating mode and that the controller maintains the process at the nominal operating point $r(t) = 0$ eliminating the pulse-type disturbances. Also, for many nonlinear control systems the normal mode of operation is the self-oscillating mode. We can mention here two examples: control systems with two-term controllers (temperature processes) and a majority of extremal control systems work in extremal mode (static optimal control systems).

6.1.1 Analysis of Symmetrical Self-Oscillations

Symmetrical self-oscillations can appear in nonlinear systems which have a nonlinear part with symmetrical static characteristic. A periodic solution of nonlinear equation (4.5) by the describing function method is obtained from the condition of harmonic balance of the system whose structure is shown in Fig. 6.1.

The dynamics of such an unforced nonlinear control system (Fig. 6.1a) are described by the differential equation:

$$y(p) = G_L(p) \cdot F(x) = \frac{B(p)}{A(p)} \cdot F(x) = -x(p) \tag{6.1}$$

respectively:

$$x(p) + G_L(p)F(x) = 0 \tag{6.2}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

or:

$$A(p)x(p) + B(p)F(x) = 0 \tag{6.3}$$

The realization of monoharmonic undamped oscillations in the system is possible only when a periodic function—harmonic signal $x(t) = X_m sin\omega t$—acts at the input of nonlinear element. This is seen in Fig. 6.1b, when self-oscillations were established.

At the output of the nonlinear element, signal $y_N(t) = F[x(t)]$ contains the first harmonic $Y_N sin(\omega t + \varphi)$ and higher harmonics with amplitudes $Y_{m2}, Y_{m3}, \ldots, Y_{mk}$ and frequencies $2\omega, 3\omega, \ldots, k\omega$. Assuming that the linear part of the system filters out all the harmonics except the first harmonic, the output signal can be approximated by the equation:

$$y(t) \approx Y_{m1} \sin(\omega t + \varphi) = y_{N1}(t) \tag{6.4}$$

The assumption (6.4) is called the *filter hypothesis*. The condition that the linear part of the system $G_L(\omega)$ acts as a "low-pas" filter is necessary for the application of the harmonic linearization method.

The two block diagrams in Figs. 6.1a and 6.1b are equivalent if the equation:

$$x(t) + y(t) = 0 \tag{6.5}$$

holds. From (6.4) and (6.5) follows:

$$x(t) = -Y_{m1} \sin(\omega t + \varphi) \tag{6.6}$$

The condition (6.6) is fulfilled if:

$$X_m = Y_{m1} \tag{6.7}$$

$$\varphi = \pi \tag{6.8}$$

The equations (6.7) and (6.8) are called *harmonic balance equations*. Equation (6.7) indicates equality of amplitudes, and (6.8) phase shift of 180[deg].

If the filter hypothesis is to be satisfied, the frequency characteristic of the linear part of the system $G_L(j\omega)$ must pass the first harmonic and attenuate all the higher harmonics to a negligible degree:

$$\frac{y_{Nk}}{y_{N1}} \left| \frac{G_L(jk\omega)}{G_L(j\omega)} \right| << 1 \tag{6.9}$$

The condition (6.9) must be fulfilled for any integer $k \geq 2$. As with symmetrical oscillations in signal $y(t)$, the even harmonics are not present, the condition (6.9) is for symmetrical oscillation valid for $k \geq 3$.

By applying the describing function method, two main tasks are solved:

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

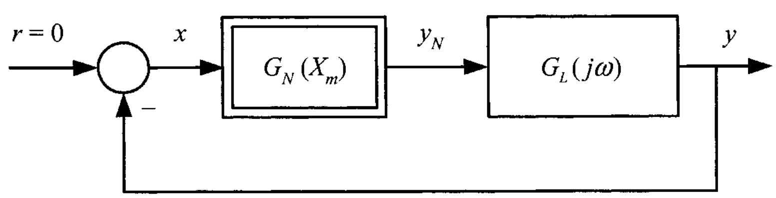


Figure 6.2: Block diagram of an unforced nonlinear control system, linearized by harmonic linearization.

1. Analysis of monoharmonic oscillations in closed-loop nonlinear control systems.
2. Determining conditions for which in the closed-loop nonlinear control system's self-oscillations cannot originate in the whole frequency band $0 \leq \omega \leq \infty$.

As $y_{Nk}/y_{NI} < 1$ for $k \geq 2$, the condition (6.9) can be relaxed and is practically fulfilled when the slope of the amplitude logarithmic frequency characteristic of the linear part $G_L(j\omega)$ is equal to $-20[dB/decade]$ or $-40[dB/decade]$.

With the slope $G_L(j\omega)$ equal to $-20[dB/decade]$, the condition (6.9) is fulfilled if:

$$\left|\frac{G_L(j2\omega)}{G_L(j\omega)}\right| = \frac{1}{2}; \left|\frac{G_L(j3\omega)}{G_L(j\omega)}\right| = \frac{1}{3} \tag{6.10}$$

and with the slope of $-40[dB/decade]$:

$$\left|\frac{G_L(j2\omega)}{G_L(j\omega)}\right| = \frac{1}{4}; \left|\frac{G_L(j3\omega)}{G_L(j\omega)}\right| = \frac{1}{9} \tag{6.11}$$

In the cases when $G_L(j\omega)$ doesn't satisfy equation (6.9), or (6.10) and (6.11) respectively, the results of the analysis by the describing function method must be tested experimentally or theoretically, by analyzing physical or mathematical models.

The structure of the control system in Fig. 6.1 which is linearized by harmonic linearization is shown in Fig. 6.2.

The replacement in (6.2) of the operator p by the variable $s = j\omega$ and $F(x) = G_N(X_m)x$, the equation of harmonic balance is obtained:

$$G_N(X_m)G_L(j\omega)x + x = 0 \tag{6.12}$$

The characteristic equation of the system (6.12) is:

$$G_N(X_m)G_L(j\omega) + 1 = 0 \tag{6.13}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

By solving the complex equation (6.13), the parameters X_m and ω of possible self-oscillations are determined. It must be emphasized that with this solution the amplitude and the frequency of the signal at the input to the nonlinear element $x(t)$ are obtained (Fig. 6.2). If the amplitude and the frequency of oscillations of the output signal are to be found, the solution will be identical ($X_m = Y_m$, $\omega_x = \omega_y$)[4] only if we are dealing with unity feedback. If this is not the case, the obtained solution $x(t)$ must be brought back through a feedback element in order to solve for $y(t)$. Equation (6.13) can be solved graphically or analytically. The graphical procedure for determining parameters X_m and ω is most appropriate when the control system contains a simple inertialess nonlinear element $y_N = F(x)$.

Graphical Procedure—Goldfarb Method (Goldfarb, 1947)

In order to apply this procedure, equation (6.13) is written in the form:

$$G_L(j\omega) = -\frac{1}{G_N(X_m)} = -G_N^{-1}(X_m) \tag{6.14}$$

or:

$$G_N(X_m) = -\frac{1}{G_L(j\omega)} = -G_L^{-1}(j\omega) \tag{6.15}$$

From equation (6.14) it is seen that the graphical solution requires the amplitude-phase frequency characteristic of the linear part $G_L(j\omega)$ of the system, as well as the negative inverse describing function amplitude-phase characteristic of the non-linear part. Obviously, each point on the frequency characteristic represents one frequency and each point on the negative describing function represents one amplitude of oscillation of x. If the plots $G_L(j\omega)$ and $-G_N^{-1}(X_m)$ have one or more intersection points, it means the existence of limit cycles, i.e. in the nonlinear system undamped periodic oscillations, self-oscillations (Fig. 6.3), are possible. When conditions in equations (6.14) and (6.15) are not fulfilled, i.e. the intersection points do not exist, the system cannot generate sustained periodic oscillations (self-oscillations).

With intersection points M and N (Fig. 6.3), the amplitude and the frequency of self-oscillations are determined. The amplitude X_m is determined by the points M and N of the negative inverse amplitude-phase characteristic of the nonlinear element $-G_N^{-1}(X_m)$, while the frequency ω is determined at the points M and N on the amplitude-phase frequency characteristic of the linear part $G_L(j\omega)$. In such a way, the intersection point M has the amplitude X_{mM} and the frequency ω_M, while the intersection point N has the amplitude X_{mN} and the frequency ω_N.

The graphical plot (Fig. 6.3) enables us to determine the stability of the self-oscillations of the input signal to the nonlinear element. One of the possible procedures for testing the stability of the self-oscillations is the application of Nyquist

[4] Index denotes frequencies of self-oscillations of signals x and y, respectively.

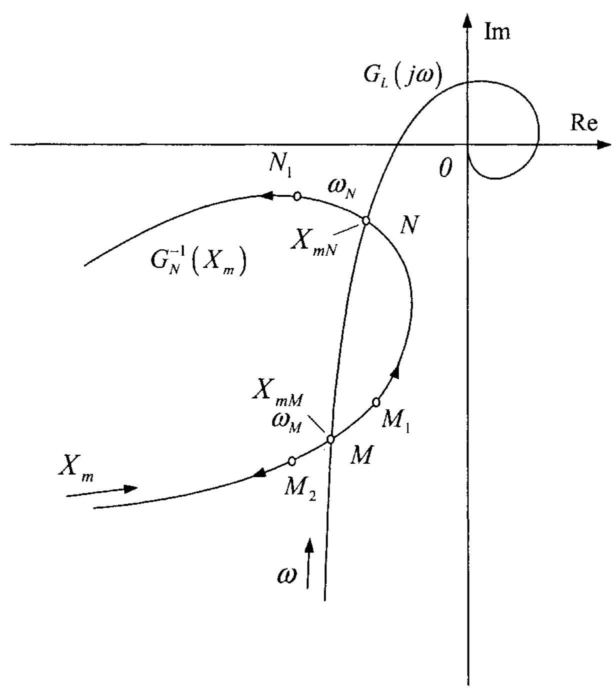


Figure 6.3: Goldfarb method of finding parameters of self-oscillations.

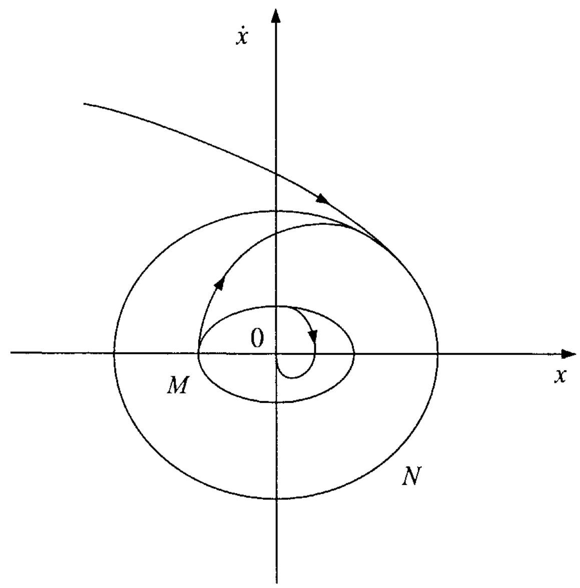


Figure 6.4: Phase trajectories of system which has unstable self-oscillations (limit cycle is smaller ellipse) and stable self-oscillations (limit cycle is larger ellipse).

stability criterion to the linearized system. If the equation of periodic oscillations for the linear system (the stability boundary of the system) has the form:

$$1 + G_0(j\omega) = 0$$

with critical point of the Nyquist plot $(-1, j0)$, the equation of self-oscillations for nonlinear systems is given by (6.13). The critical points for the Nyquist plot $G_L(j\omega)$ are the points M and N at the intersection with negative inverse characteristic $-G_N^{-1}(X_m)$ of the nonlinear element, Fig. 6.3. With a small increase of the amplitude of the input signal, the critical point M passes over along $-G_N^{-1}(X_m)$ to the point M_1, which is included by the plot $G_L(j\omega)$. According to the Nyquist criterion, the point M corresponds to the critical point $(-1, j0)$. Due to the fact that $G_L(j\omega)$ includes the point M with the increase of amplitude of the input signal $x(t)$, the periodic solution (or limit cycle) determined by the parameters of the point M is an unstable solution, i.e. self-oscillations with the amplitude X_{mM} and the frequency ω_M are unstable self-oscillations (the limit cycle M in Fig. 6.4).

However, with a small increase of the amplitude of the input signal, the critical point N goes over to the point N_1, which is not included by $G_L(j\omega)$. As the plot $G_L(j\omega)$ does not include the point N, with the increase of the input signal

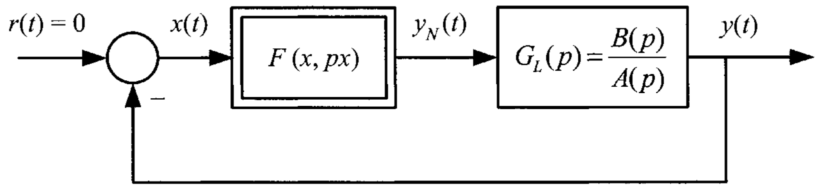


Figure 6.5: Block diagram of an unforced nonlinear control system.

amplitude X_m the periodic solution (or limit cycle) determined by the parameters of the point N is a stable solution. It means that the system will sustain stable self-oscillations with the amplitude X_{mN} and the frequency ω_N (the limit cycle N in Fig. 6.4).

By analogy to the Nyquist criterion, the stability of the periodic solution can be found by the Mikhailov criterion (Netushil, 1983). For instance, the dynamics of an unforced nonlinear control system of arbitrary structure with inertial nonlinear element of the form $y_N = F(x, px)$ (Fig. 6.5) is described by the equation:

$$G_L(p)F(x,px) + x = 0 \tag{6.16}$$

According to (3.51), the describing function of the nonlinearity $y_N = F(x, px)$ is:

$$F(x,px) = P(X_m,\omega)x + \frac{Q(X_m,\omega)}{\omega}px = G_N(X_m,\omega)x \tag{6.17}$$

From (6.16) and (6.17), the differential equation of the harmonic linearized system is obtained:

$$G_L(p)\left[P(X_m,\omega) + \frac{Q(X_m,\omega)}{\omega}p\right]x + x = 0 \tag{6.18}$$

By substituting p with s, the characteristic equation of the closed-loop linearized system is:

$$D(s) = G_L(s)G_N(X_m,\omega) + 1 = G_L(s)\left[P(X_m,\omega) + \frac{Q(X_m,\omega)}{\omega}s\right] + 1 = 0 \tag{6.19}$$

By inserting $G_L(s) = B(s)/A(s)$ in (6.19):

$$D(s) = A(s) + B(s)\left[P(X_m,\omega) + \frac{Q(X_m,\omega)}{\omega}s\right] = 0 \tag{6.20}$$

The equation (6.18) will have periodic solution $x(t) = X_m sin\omega t$ only in the case when the roots of the characteristic equation (6.20) are imaginary. Putting $s = j\omega$

in the equation (6.20) the equation will have real and imaginary parts:

$$D(j\omega) = A(j\omega) + B(j\omega)\left[P(X_m, \omega) + \frac{Q(X_m, \omega)}{\omega} j\omega\right] = 0 \tag{6.21}$$

and further:

$$D(j\omega) = Re(X_m, \omega) + jIm(X_m, \omega) \tag{6.22}$$

respectively:

$$\begin{aligned} Re(X_m, \omega) &= 0 \\ Im(X_m, \omega) &= 0 \end{aligned} \tag{6.23}$$

where $Re(X_m, \omega)$ is the real part of the function (6.21) and $Im(X_m, \omega)$ is the imaginary part of the function (6.21).

By solving equation (6.23) the parameters of self-oscillations X_m and ω are obtained analytically.

When nonlinearity $F(x, px)$ is replaced by the describing function $G_N(X_m, \omega)$, the closed-loop transfer function $G_{Ncl}(X_m, s, \omega)$ results:

$$\frac{Y(s)}{R(s)} = G_{Ncl}(X_m, s, \omega) = \frac{G_N(X_m, \omega) G_L(s)}{1 + G_N(X_m, \omega) G_L(s)} \tag{6.24}$$

The fact that the transfer function of a closed-loop linearized control system is identical in form to the linear control system closed by unity feedback enables us to find the parameters of oscillations by means of a Mikhailov plot, i.e. the graphical display of complex equation (6.22).

The graphical solution of the equation (6.22) corresponds to the Mikhailov plot which intersects coordinate origin (Fig. 6.6) with parameters X_m and ω.

It is possible to determine graphically the Mikhailov plot which passes through the coordinate origin of the polar plane (stability boundary of the linear system) from the inclined rectangle (the condition is $x_1 > x_2$), when the following expressions are applied:

$$\begin{aligned} X_m &= x_2 + \frac{\overline{bo}}{\overline{ab}}(x_1 - x_2) \\ \omega &= \omega_1 + \frac{\overline{co}}{\overline{cd}}(\omega_2 - \omega_1) \end{aligned} \tag{6.25}$$

Here x_1 and x_2 are the amplitude values for which the Mikhailov plots pass near the coordinate origin, while ω_1 and ω_2 are the frequencies on the plots x_1 and x_2 near the coordinate origin.

By means of Mikhailov plots the stability of the periodic solution can be estimated. By increasing the amplitude of the periodic solution to $X_m + \Delta x$, the Mikhailov plot will be either in position 1 or position 2 (Fig. 6.6b). The curve

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

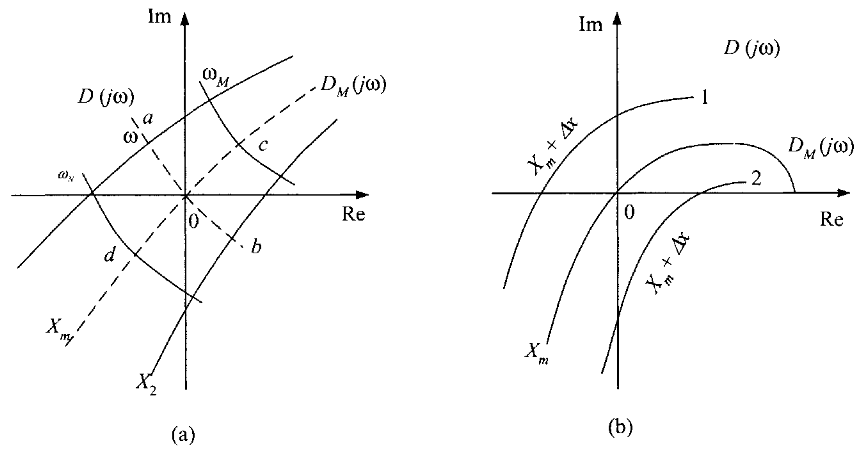


Figure 6.6: Mikhailov plots.

1 encloses the coordinate origin, i.e. the transient process is a stable damped oscillatory process. The curve 2 characterizes unstable oscillatory process with increasing amplitudes. If at $\Delta x > 0$, the Mikhailov plot has the position 1, and at $\Delta x < 0$ the position 2, the transient process will be such that oscillations with the amplitude $X > X_m$ are damped (will decrease to the value $X = X_m$), while oscillations with amplitude $X > X_m$ will increase to the value $X = X_m$. In the aforesaid case the transient process on both sides of the coordinate origin approaches the origin, i.e. the periodic solutions are stable self-oscillations. By analogy, if the Mikhailov plot at $\Delta x < 0$ has the position 2 and at $\Delta x < 0$ the position 1, the transient process on both sides of the coordinate origin has increasing amplitudes, and the periodic solution is unstable. That means that the system's equilibrium state is locally stable and globally unstable.

6.1.2 Analytical Stability Criterion of Self-Oscillations

With graphical methods of determining the parameters of self-oscillations, it is possible to formulate an analytical stability criterion. For instance, the stability of self-oscillations obtained by the Goldfarb method, the points M and N on the graphical display in Fig. 6.3 can be determined by using a damping factor σ of the oscillatory process which is close to self-oscillations. In other words, the frequency of the oscillatory process with damping factor σ is close to the frequencies of oscillatory processes in the points of harmonic balance M and N. The oscilla-

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

tory process with damping factor σ is described by the expression:

$$x(t) = X_m e^{\sigma t} \sin(\omega t + \varphi) \tag{6.26}$$

It follows that the process $x(t)$ will be damped for $\sigma < 0$, while for $\sigma > 0$ the process will increase unboundedly.

In the vicinity of the point which determines the periodic solution (the points M and N in Fig. 6.3), the equation (6.14) assumes the form:

$$-G_N^{-1}(X_m) = -V_N(X_m) = P_N(X_m) + jQ_N(X_m) = G_L(\sigma + j\omega) \tag{6.27}$$

respectively:

$$P_N(X_m) + jQ_N(X_m) = P_L(\sigma, \omega) + jQ_L(\sigma, \omega) \tag{6.28}$$

where:

$P_N(X_m)$ is the real part of $-G_N^{-1}(X_m)$,

$Q_N(X_m)$ is the imaginary part of $-G_N^{-1}(X_m)$,

$P_L(\sigma, \omega)$ is the real part of the frequency transfer function of the linear part of the closed-loop control system,

$Q_L(\sigma, \omega)$ is the imaginary part of the frequency transfer function of the nonlinear part of the closed-loop control system.

From (6.28) is obvious that the parameters σ and ω are functions of the oscillation amplitude X_m.

If small deflections (point M_1 in Fig.6.3) from the periodic solution (point M in Fig. 6.3) take place, with increasing X_m the coefficient $\sigma > 0$ will increase and consequently the amplitude of the process (6.26) $|x(t)| = X_m e^{\sigma t}$ will increase with time indicating that the periodic solution (point M) is unstable. On the contrary, if small deflections (point N_1 in Fig. 6.3) from the periodic solution (point N in Fig. 6.3) take place, with increasing X_m, the coefficient $\sigma < 0$ will increase and the amplitude of the process (6.26) will decrease with time till the original value X_{mN} is reached, i.e. the system sustains stable self-oscillations.

From this it is seen that the analytical stability of the periodic solution is determined by the change $d\sigma/dX_m$. If $d\sigma/dX_m < 0$ self-oscillations in the system are stable, and vice versa, for $d\sigma/dX_m > 0$, self-oscillations are unstable[5]. By determining $d\sigma/dX_m$from (6.28), the analytical expression for the stability of self-oscillations is obtained:

$$\frac{\partial P_N}{\partial X_m} = \frac{\partial P_L}{\partial \sigma}\frac{\partial \sigma}{\partial X_m} + \frac{\partial P_L}{\partial \omega}\frac{\partial \omega}{\partial X_m} \tag{6.29}$$

[5]By unstable self-oscillations we mean that oscillations with a particular amplitude and frequency will not be established, i.e. the state will move to another singular point (stable self-oscillation or other).

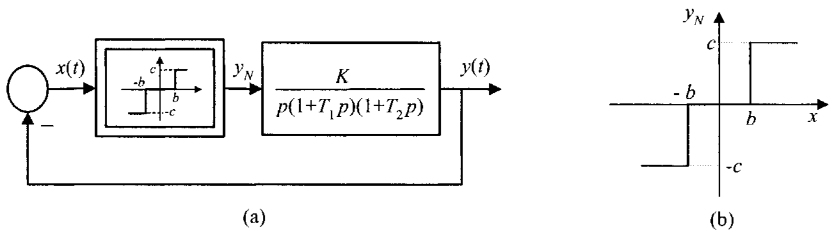


Figure 6.7: Block diagram of unforced closed-loop control system and nonlinear static characteristic of nonlinear element.

$$\frac{\partial Q_N}{\partial X_m} = \frac{\partial Q_L}{\partial \sigma}\frac{\partial \sigma}{\partial X_m} + \frac{\partial Q_L}{\partial \omega}\frac{\partial \omega}{\partial X_m} \tag{6.30}$$

For the complex function (6.28) to be analytic at the point (X_m, ω), the Cauchy-Riemann condition states:

$$\frac{\partial P_L}{\partial \sigma} = \frac{\partial Q_L}{\partial \omega}; \ \frac{\partial Q_L}{\partial \sigma} = -\frac{\partial P_L}{\partial \omega} \tag{6.31}$$

From (6.31), (6.30) and (6.29) follows:

$$\frac{d\sigma}{dX_m} = \frac{\frac{\partial Q_L}{\partial \omega}\frac{\partial P_N}{\partial X_m} - \frac{\partial P_L}{\partial \omega}\frac{\partial Q_N}{\partial X_m}}{\left(\frac{\partial P_L}{\partial \omega}\right)^2 + \left(\frac{\partial Q_L}{\partial \omega}\right)^2} \tag{6.32}$$

Since the denominator in (6.32) is always positive, the condition for the stability of self-oscillations will be fulfilled when:

$$\frac{\partial Q_L}{\partial \omega}\frac{\partial P_N}{\partial X_m} - \frac{\partial P_L}{\partial \omega}\frac{\partial Q_N}{\partial X_m} < 0 \tag{6.33}$$

The above expression is the *Mikhailov criterion of stability of self-oscillations.*

EXAMPLE 6.1

Find the stability of the equilibrium state of the nonlinear system with block diagram shown in Fig. 6.7. The following values are set: $K = 0.82[s^{-1}]$, $T_1 = T_2 = 0.05[s]$, $b = 0.25[V]$, $c = 110[V]$.

Solution. *For the Goldfarb method, the amplitude-phase frequency characteristic of the linear part* $G_L(j\omega)$ *as well as the harmonic linearized plot of the nonlinear part of the system* $-V_N(X_m) = -G_N^{-1}(X_m)$ *are to be determined.*

Frequency transfer function of $G_L(j\omega)$ *is:*

$$G_L(j\omega) = \frac{K}{j\omega(1+T_1 j\omega)(1+T_2 j\omega)} \tag{6.34}$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

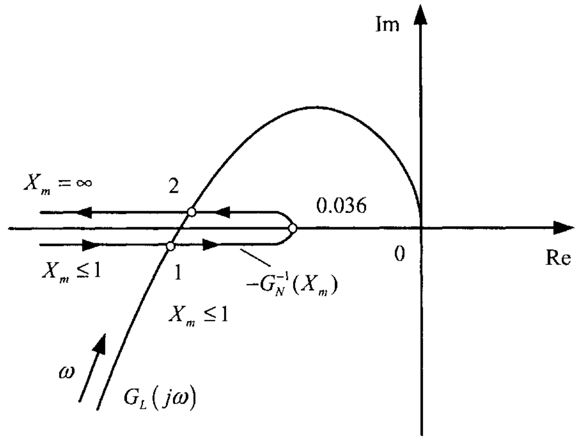


Figure 6.8: Plot of negative inverse describing function and frequency characteristic of linear part of closed-loop control system.

with modulus:

$$|G_L(j\omega)| = \frac{K}{\omega\sqrt{(1+T_1^2\omega^2)(1+T_2^2\omega^2)}} = \frac{0.82}{\omega(1+0.0025\omega^2)} \tag{6.35}$$

and argument:

$$\varphi(\omega) = -90\,\text{deg} - \tan^{-1}(\omega T_1) - \tan^{-1}(\omega T_2) = -90\,\text{deg} - 2\tan^{-1}(0.05\omega) \tag{6.36}$$

By varying ω from $\omega = 0$ to $\omega = \infty$, the last two equations define $G_L(j\omega)$ in Fig. 6.8.

The describing function of the nonlinear element (relay with a dead zone) is:

$$G_N(X_m) = P_N(X_m) = \frac{4c}{\pi X_m}\sqrt{1 - \frac{b^2}{X_m^2}};\ X_m > b \tag{6.37}$$

If the value X_m from $X_m = b = 0.25$ to $X_m = \infty$ is changing, the plot $-G_N^{-1}(X_m)$ of the nonlinear element can be constructed[6]*. From extremum of the function (6.37) results (for $X_m = b\sqrt{2}$):*

$$\left|G_N^{-1}(X_m)\right|_{min} = \frac{\pi b}{2c} = \frac{\pi \cdot 0.25}{2 \cdot 110} \approx 0.0036 \tag{6.38}$$

[6] The curve $-G_N^{-1}(Xm)$ is here a real quadratic function for this nonlinearity, see (6.36), and has double real negative solutions for every Xm. In Fig. 6.8 this is emphasized by plotting the curve as having an imaginary part, which is not true.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The curve (6.36) coincides with the negative real axis and has two branches which intersect the curve $G_L(j\omega)$ at two points. From (6.36) follows that at the crossing of $G_L(j\omega)$ with the real axis $\varphi = -180\,\mathrm{deg}$. From this follows:

$$-180\,\mathrm{deg} = -90\,\mathrm{deg} - 2\tan^{-1}(0.05\omega)$$

and also:

$$tan^{-1}(0.05\omega) = 45\,\mathrm{deg} \text{ or } 0.05\omega = 1$$

so $\omega = 20[s^{-1}]$. *Calculating the modulus of* $|G_L(j20)|$ *gives:*

$$|G_L(j20)| = \frac{0.82}{20(1+1)} = 0.0205 = \frac{1}{G_N(x)}$$

or:

$$G_N(x) = \frac{4 \cdot 110}{\pi}\sqrt{\frac{1}{x^2} - \frac{0.0625}{x^4}} = 48.75$$

The above equation gives:

$$x^4 - 8.24x^2 + 0.516 = 0$$

Two real solutions are:

$$x^2 = 8.176 \text{ and } X_{m_2} = 2.86$$
$$x^2 = 0.064 \text{ and } X_{m_1} = 0.257$$

The equation of harmonic balance:

$$G_L(j\omega) = -\frac{1}{G_N(X_m)} \tag{6.39}$$

has two solutions:

$$\omega = 20[s^{-1}] \text{ and } X_{m_1} = 0.257[V]$$
$$\omega = 20[s^{-1}] \text{ and } X_{m_2} = 2.86[V]$$

so that:

$$\begin{aligned} x_1(t) &= X_{m_1}\sin\omega t = 0.257\sin(20t) \text{ and} \\ x_2(t) &= X_{m_2}\sin\omega t = 2.86\sin(20t) \end{aligned} \tag{6.40}$$

The stable solution of the system (6.37) is the solution at the point 2, i.e. by increasing the amplitude X_m, the curve $G_L(j\omega)$ will not include the intersection point with parameters $X_m = 2.86[V]$ and $\omega = 20[s^{-1}]$. In this nonlinear closed-loop control system stable self-oscillations $x(t) = 2.86\sin 20t$ will be established.

■

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

6.1.3 Determination of Symmetrical Self-Oscillations

An analytical method to determine symmetrical self-oscillations originates from the analytical solution of complex equation (6.22), or with the solution of the system of equations (6.23). The analytical solution of the system (6.23) is quite often impossible to find. Because of that the synthesis of the parameters of linear or nonlinear compensator is done by use of either the analytical stability criterion and grapho-analytical method. Either complex equation (6.22) or the system of equations (6.23) enables us to formulate the analytical stability criterion (Petrov et al., 1967):

$$\frac{\partial R}{\partial X_m}\frac{\partial I}{\partial \omega} - \frac{\partial I}{\partial X_m}\frac{\partial R}{\partial \omega} > 0 \tag{6.41}$$

where:

$$D(j\omega) = Re(X_m, \omega) + jIm(X_m, \omega) = 0$$
$$R = Re(X_m, \omega);\; I = Im(X_m, \omega)$$

EXAMPLE 6.2
Find the parameters and the stability of the periodic solution of an electromechanical servo, described by block diagram in Fig. 6.7.

Solution. *The following differential equation describes the dynamics of the nonlinear system with nonlinear element $y_N = F(x)$:*

$$A(p)x + B(p)F(x) = 0 \tag{6.42}$$

Harmonic linearization of equation (6.42) gives:

$$A(p)x + B(p)G_N(X_m)x = 0 \tag{6.43}$$

and the characteristic equation of (6.43) is:

$$A(p) + B(p)G_N(X_m) = 0 \tag{6.44}$$

According to Fig. 6.7, the above equation becomes:

$$T_1T_2p^3 + (T_1 + T_2)p^2 + p + KG_N(X_m) = 0 \tag{6.45}$$

By replacing in (6.45) operator p with complex variable $s = j\omega$, the complex equation follows:

$$D(j\omega) = KG_N(X_m) - (T_1 + T_2)\omega^2 + j(\omega - T_1T_2\omega^3) = 0 \tag{6.46}$$

and further:

$$R(X_m, \omega) = KG_N(X_m) - (T_1 + T_2)\omega^2 = 0 \tag{6.47}$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

$$I(\omega) = \omega - T_1 T_2 \omega^3 = 0 \tag{6.48}$$

From equation (6.48), the frequency of the periodic solution can be found:

$$\begin{aligned} \omega^2 &= \frac{1}{T_1 T_2} = \frac{1}{0.05 \cdot 0.05} = 400[s^{-2}] \\ \omega &= 20[s^{-1}] \end{aligned} \tag{6.49}$$

The describing function of the nonlinear element is:

$$G_N(X_m) = \frac{4c}{\pi X_m}\sqrt{1 - \frac{b^2}{X_m^2}} \tag{6.50}$$

From (6.47), (6.48) and (6.49):

$$G_N(X_m) = \frac{T_1 + T_2}{K T_1 T_2} = \frac{4 y_N}{\pi X_m}\sqrt{1 - \frac{x_a^2}{X_m^2}} \tag{6.51}$$

Putting in equation (6.51) the values $T_1 = T_2 = 0.05[s]$, $K = 0.82[s^{-1}]$, $b = 0.25[V]$, $c = 110[V]$, *the equivalent gain of the nonlinear element is:*

$$G_N(X_m) = \frac{0.05 + 0.05}{0.82 \cdot 0.05 \cdot 0.05} = 48.75 \tag{6.52}$$

From (6.52) and (6.51) follows:

$$\frac{4 \cdot 110}{\pi X_m}\sqrt{1 - \frac{0.25^2}{X_m^2}} = 48.75 \tag{6.53}$$

so that two solutions for the amplitude of the input signal are obtained:

$$X_{m1} = 0.257[V]; \; X_{m2} = 2.86[V] \tag{6.54}$$

Stability of the periodic solution is determined by application of criterion (6.41) to the equations (6.47) and (6.48). The corresponding partial derivatives are:

$$\begin{aligned} \frac{\partial R}{\partial X_m} &= \frac{4KC}{\pi}\frac{2b^2 - X_m^2}{X_m^3\sqrt{X_m^2 - b^2}} \\ \frac{\partial R}{\partial \omega} &= -2(T_1 + T_2)\omega = -2(T_1 + T_2)\sqrt{\frac{1}{T_1 T_2}} \\ \frac{\partial I}{\partial X_m} &= 0 \\ \frac{\partial I}{\partial \omega} &= 1 - 3T_1 T_2 \omega^2 = 1 - 3T_1 T_2\left(\sqrt{\frac{1}{T_1 T_2}}\right)^2 = -2 \end{aligned}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

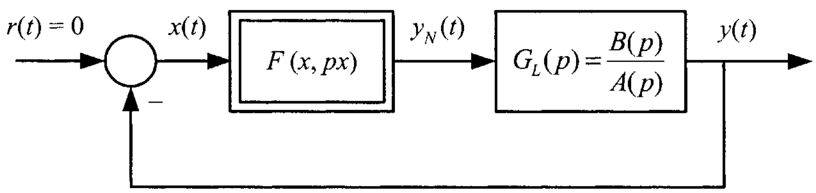


Figure 6.9: Block diagram of unforced nonlinear closed-loop control system.

By inserting the above expressions for partial derivatives in the criterion (6.41), the following inequality results:

$$\frac{4KC}{\pi}\frac{2b^2 - X_m^2}{X_m^3\sqrt{X_m^2 - b^2}} \cdot (-2) > 0 \tag{6.55}$$

The condition (6.55) will be fulfilled when $X_m^2 > 2b^2$, *or* $X_m > b\sqrt{2} = 0.3525[V]$. *The requirement* $X_m > 0.3525[V]$ *satisfies solution* $X_{2m} = 2.86[V]$. *We can conclude that in the described system stable self-oscillations* $x(t) = 2.86\sin 20t$ *will sustain.* ■

6.1.4 Asymmetrical Self-Oscillations—Systems with Asymmetrical Nonlinear Static Characteristic

Asymmetrical self-oscillations can be established in systems with an asymmetrical static nonlinear characteristic in the nonlinear element $F(x, px)$, in accordance with the transformation of an asymmetrical into a symmetrical characteristic (Fig. 1.15). The oscillations at the input of the nonlinear element, Fig. 6.9, are approximated by:

$$x(t) = x^0(t) + X_m \sin\omega t = x^0(t) + x^* \tag{6.56}$$

where:

$x^0(t)$ is the shift of the center of oscillations which is caused by the asymmetrical characteristic $F(x, px)$ and

$x^* = X_m \sin\omega t$ is the periodic component of the solution.

When the asymmetrical nonlinear characteristic $F(x, px)$ is expanded in a Fourier series, a nonoscillatory term $F^0(x^0, X_m, \omega)$ appears:

$$F(x, px) \approx F^0(x^0, X_m, \omega) + \left[P(x^0, X_m, \omega) + \frac{Q(x^0, X_m, \omega)}{\omega}p\right]x \tag{6.57}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

where:

$$F^0 = \frac{1}{2\pi}\int_0^{2\pi} F\left(x^0 + X_m \sin\psi, X_m \cos\psi\right) d\psi \tag{6.58}$$

$$\begin{aligned} P &= \frac{1}{\pi X_m}\int_0^{2\pi} F(x^0 + X_m \sin\psi, X_m \cos\psi)\sin\psi d\psi \\ Q &= \frac{1}{\pi X_m}\int_0^{2\pi} F(x^0 + X_m \sin\psi, X_m \cos\psi)\cos\psi d\psi \end{aligned} \tag{6.59}$$

and $\psi = \omega t$.

If the system contains a nonlinearity of zero order $y_N = F(x)$, the coefficients of harmonic linearization (F^0, P and Q) will not depend upon the frequency ω. For typical nonlinearities, coefficients F^0, P and Q are determined by analytical expressions or by graphical displays (Table A.2 in Appendix A).

The differential equation of the nonlinear control system given in block diagram (Fig. 6.9) is:

$$x(t) + y(t) = 0$$

or:

$$x(t) + F(x, px)\cdot G_L(p) = 0 \tag{6.60}$$

Combining equations (6.60) and (6.57), the harmonic linearized equation follows:

$$x(t) + \left(F^0 + Px^* + \frac{Q}{\omega}px^*\right) G_L(p) = 0 \tag{6.61}$$

Inserting in equation (6.61) $x(t) = x^0 + x^*$ and $G_L(p) = B(p)/A(p)$:

$$A(p)(x^0 + x^*) + B(p)\left(F^0 + Px^* + \frac{Q}{\omega}px^*\right) = 0 \tag{6.62}$$

or:

$$A(p)x^0 + A(p)x^* + B(p)F^0 + B(p)(P + \frac{Q}{\omega}p)x^* = 0 \tag{6.63}$$

Equation (6.63) can be satisfied only in the case when the terms for nonoscillatory component x^0 and the periodic component x^* are equal to zero. With $x^0(t) = x^0 = const.$, from (6.63) follows for the nonoscillatory component:

$$A(0)x^0 + B(0)F^0(x^0, X_m, \omega) = 0 \tag{6.64}$$

and for the periodic component:

$$A(p)x^* + B(p)\left[P\left(x^0, X_m, \omega\right) + \frac{Q(x^0, X_m, \omega)}{\omega}p\right]x^* = 0 \tag{6.65}$$

These two equations derived from equation (6.63) are nonlinearly dependent, i.e. the solution of the system of equations in the form (6.56) is not possible by superposition of the solution for x^0 from (6.64) and the solution for X_m from (6.65).

Solving the equation (6.65) is analogous to that of (6.18). Replacing p with $s = j\omega$ from the characteristic equation (6.65), the complex equation is obtained:

$$D(j\omega) = A(j\omega) + B(j\omega)G_N(x^0, X_m, \omega) \tag{6.66}$$

where:

$$G_N(x^0, X_m, \omega) = P(x^0, X_m, \omega) + jQ(x^0, X_m, \omega) \tag{6.67}$$

Combining (6.66), (6.67) and (6.64), three equations with three unknowns (x^0, X_m and ω) emerge:

$$A(0)x^0 + B(0)F^0(x^0, X_m, \omega) = 0 \tag{6.68}$$

$$R(x^0, X_m, \omega) = 0 \tag{6.69}$$

$$I(x^0, X_m, \omega) = 0 \tag{6.70}$$

From the last two equations, the amplitude X_m and frequency ω of self-oscillations are found as functions of the shift of the center of oscillations x^0:

$$X_m = X_m(x^0);\ \omega = \omega(x^0) \tag{6.71}$$

Putting (6.71) into the expression for $F^0(x^0, X_m, \omega)$, there follows:

$$F^0[x^0, X_m(x^0), \omega(x^0)] = \phi(x^0)$$

and expression (6.68) takes the form:

$$A(0)x^0 + B(0)\phi(x^0) = 0 \tag{6.72}$$

The function $F^0[x^0, X_m(x^0), \omega(x^0)] = \phi(x^0)$ is called the *shifting function* $\phi(x^0)$. It is most often a continuous function for all typical nonlinear elements. Determining x^0 from (6.72) and inserting it in (6.71) the solutions for parameters of asymmetrical self-oscillations x^0, X_m and ω can be found.

In a control system with an integral linear part (Type 1 control system) $G_L(p) = 1/pG(p)$, from (6.72), can be derived:

$$\phi(x^0) = -\frac{x^0}{G_L(0)} = 0 \tag{6.73}$$

Solution of equation (6.73) gives the shift of the center of oscillations. This is necessary for cancelling the asymmetrical self-oscillations caused by asymmetrical characteristics of the nonlinear part of the system (Figs. 6.10a and 6.10b).

Fig. 6.10b suggests that for $F^0 = \phi(x^0) = 0$, the shift of the self-oscillations' center x^0 cancels the asymmetry $F(x)$, so the self-oscillations become symmetrical in relation to the time axis. The self-oscillations at the input of the nonlinear part of the system $x = x^0 + X_m sin\omega t$ are obtained through the complex gain, by use of harmonic linearization. Self-oscillations of any signal in a closed-loop nonlinear control system $x_i(t) = x_i^0 + x_i^*$ can be determined by using the same technique.

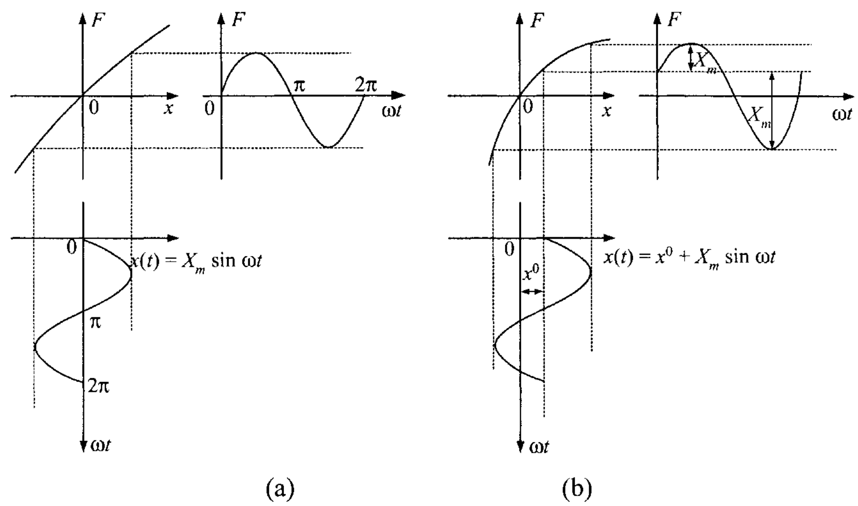


Figure 6.10: Effect of asymmetrical static characteristic of a nonlinear element.

EXAMPLE 6.3

For the system with block diagram in Fig. 6.9, the coefficients of harmonic linearization of the nonlinear element are to be determined, as well as analytical expressions for the parameters of the asymmetrical periodic solution x^0, X_m and ω.

Solution. *Supposing that the linear part of the system $G_L(p)$ satisfies equations (6.9), (6.10) and (6.11), the periodic solution, i.e. asymmetrical self-oscillations at the input of the nonlinear element, are approximated by:*

$$x(t) = x^0 + X_m \sin \omega t \tag{6.74}$$

Graphical display of the signal from the nonlinear element is given in Fig. 6.11b.

Coefficients of harmonic linearization of the asymmetrical static characteristic of the nonlinear element, i.e. of the signal $y_N(t)$, are:

$$y_N^0 = F^0 = \frac{1}{2\pi}\int_0^{2\pi} F(x^0 + X_m \sin \omega t)d(\omega t) = \frac{(\pi - 2\gamma)c - (\pi + 2\gamma)mc}{2\pi} \tag{6.75}$$

or:

$$F^0 = \frac{1-m}{2}c - \frac{1+m}{\pi}\gamma c \tag{6.76}$$

Coefficients of the periodic component are:

$$P(x^0, X_m) = \frac{1}{\pi X_m}\int_0^{2\pi} F(x^0 + X_m \sin \omega t)\sin \omega t d(\omega t) \tag{6.77}$$

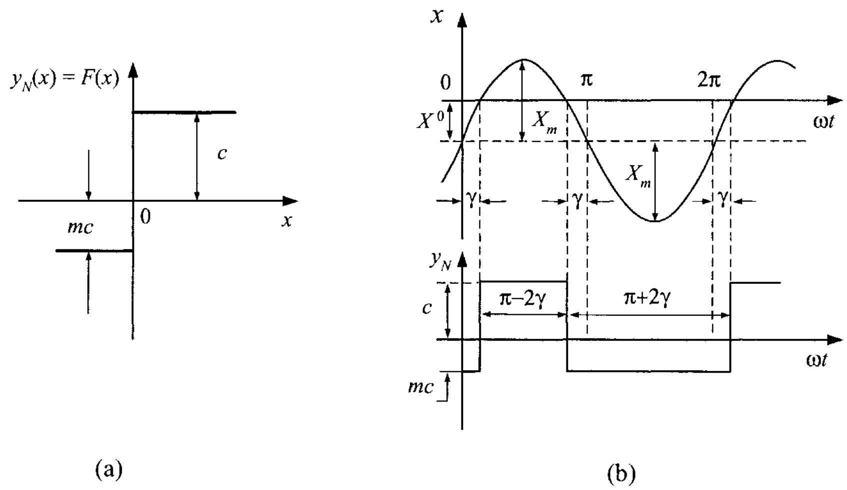


Figure 6.11: (a) Asymmetrical static characteristic of nonlinear element, (b) input and output signal of an asymmetrical nonlinearity.

From Fig. 6.11b follows:

$$\begin{aligned}
P(x^0,X_m) &= \frac{c}{\pi X_m}\int_{\gamma}^{\pi-\gamma}\sin\omega t d(\omega t) - \frac{mc}{\pi X_m}\int_{\pi-\gamma}^{2\pi+\gamma}\sin\omega t d(\omega t)\\
&= \frac{2(1+m)c}{\pi X_m}\cos\gamma \qquad (6.78)\\
Q(x^0,X_m) &= \frac{1}{\pi X_m}\int_0^{2\pi} F(x^0+X_m\sin\omega t)\cos\omega t d(\omega t) = 0
\end{aligned}$$

Equation (6.74) and the graphical display in Fig. 6.11b give:

$$x(t) = x^0 + X_m \sin\gamma = 0$$

or:

$$x^0 = -X_m\sin\gamma;\ \gamma = -\sin^{-1}\frac{x^0}{X_m} \qquad (6.79)$$

If (6.79) and (6.76) are combined:

$$F^0(x^0,X_m) = \frac{(1-m)c}{2} + \frac{(1+m)c}{\pi}\sin^{-1}\frac{x^0}{X_m} \qquad (6.80)$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

The describing function for the oscillatory component results from (6.78) and (6.79):

$$G_N(x^0, X_m) = P(x^0, X_m) = \frac{2(1+m)c}{\pi X_m}\sqrt{1-\left(\frac{x^0}{X_m}\right)^2} \tag{6.81}$$

Equations (6.80) and (6.64) give:

$$F^0(x^0, X_m)G_L(0) = -x^0 \tag{6.82}$$

For example, if the linear part of the system (Fig. 6.9) has the transfer function:

$$G_L(s) = \frac{K_2\left[K_1 + K_0(T_1 s+1)\right]}{s(T_2 s+1)(T_1 s+1)} = \frac{B(s)}{A(s)}$$
$$G_L(0) = 0$$

Equation (6.82) gives:

$$F^0(x^0, X_m) = -\frac{x^0}{G(0)} = 0 \tag{6.83}$$

From (6.83) and (6.80) the shift of the center of oscillations is found:

$$x^0 = -X_m \sin\frac{\pi}{2}\frac{1-m}{1+m} = X_m \cos\frac{\pi}{1+m} \tag{6.84}$$

The parameters of the periodic solution are obtained by solving the system of equations (6.69) and (6.72). The characteristic equation of the system is:

$$A(s) + B(s)G_N(x^0, X_m) = s(T_2 s+1)(T_1 s+1) + K_2\left[K_1 + K_0(T_1 s+1)\right]G_N(x^0, X_m) \tag{6.85}$$

Replacing s with $j\omega$ in (6.85) results in:

$$R(x^0, X_m, \omega) + jI(x^0, X_m, \omega) = 0 \tag{6.86}$$

respectively:

$$\begin{aligned} R(x^0, X_m, \omega) &= (K_1 + K_0)K_2 G_N(x^0, X_m) - (T_1 + T_2)\omega^2 = 0 \\ I(x^0, X_m, \omega) &= \left[1 + T_1 K_2 K_0 G_N(x^0, X_m)\right]\omega - T_1 T_2 \omega^3 = 0 \end{aligned} \tag{6.87}$$

From (6.87) follows:

$$\begin{aligned} \omega^2 &= \frac{K_1 + K_0}{T_1(T_2 K_1 - T_1 K_0)} \\ G_N(x^0, X_m) &= \frac{T_1 + T_2}{K_2 T_1 (T_2 K_1 - T_1 K_0)} \end{aligned} \tag{6.88}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

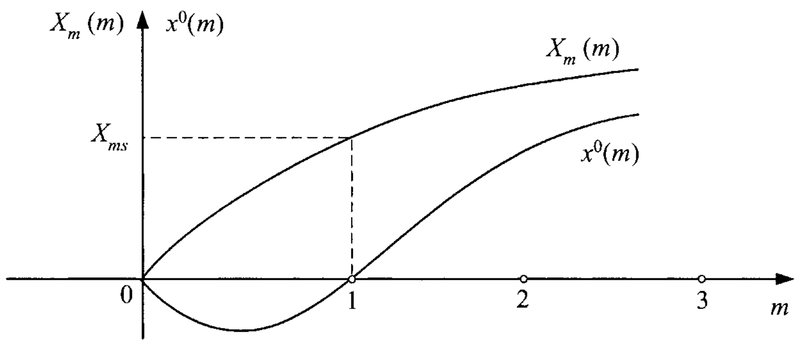


Figure 6.12: The relation between $X_m(m)$ and $x^0(m)$ according to expressions (6.89) and (6.84).

Combining (6.88), (6.81) and (6.84):

$$X_m = \frac{2c}{\pi} \frac{K_2 T_1 (T_2 K_1 - T_1 K_0)}{T_1 + T_2} (1+m) \sin \frac{\pi}{1+m} \tag{6.89}$$

The describing function of the symmetrical nonlinear element (Fig. 6.9) is:

$$G_N(X_{ms}) = \frac{4c}{\pi X_{ms}} \tag{6.90}$$

$$\omega_s^2 = \frac{K_1 + K_0}{T_1 (T_2 K_1 - T_1 K_0)} \tag{6.91}$$

$$X_{ms} = \frac{4c}{\pi} \frac{K_2 T_1 (T_2 K_1 - T_1 K_0)}{T_1 + T_2} \tag{6.92}$$

From (6.92) and (6.89), asymmetrical oscillations for the given example can be expressed by symmetrical oscillations:

$$X_m = \frac{X_{ms}}{2} (1+m) \sin \frac{\pi}{1+m} \tag{6.93}$$

The frequency ω of the periodic solution of the given unforced system is $\omega_{xm} = \omega_{xms}$. Namely, the frequency of the periodic solution with single-valued typical nonlinear elements doesn't depend on the symmetry of the static characteristic. The relation between $X_m(m)$ and $x^0(m)$ according to expressions (6.89) and (6.84) is shown in Fig. 6.12.

The coefficients of harmonic linearization $F^0(x^0, X_m)$, $Q(x^0, X_m)$ and $P(x^0, X_m)$ for asymmetrical static characteristics of basic nonlinear elements are given in table A.2 in Appendix A. ■

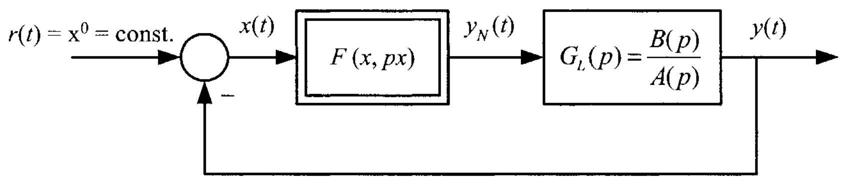


Figure 6.13: Block diagram of a forced nonlinear control system.

6.1.5 Asymmetrical Self-Oscillations—Systems with Symmetrical Nonlinear Characteristic

The determination of asymmetrical self-oscillations which can be established in the control systems which are subject to constant (or approximately constant) external action (Fig. 6.13) is analogous to finding the parameters of a periodic solution in unforced systems with an asymmetrical static characteristic of the nonlinear element.

The differential equation of the system in Fig. 6.13 is:

$$y(t) = x^0 - x(t) = G_L(p)F(x, px); \quad p = \frac{d}{dt}; \quad G_L(p) = \frac{B(p)}{A(p)} \tag{6.94}$$

or:

$$A(p)x + B(p)F(x, px) = A(p)r = M^0 \tag{6.95}$$

Because of the forced external signal $r(t) \approx const.$, the periodic solution can be asymmetrical. Asymmetrical self-oscillations at the input to nonlinear part of the system are approximated by:

$$x(t) = x^0 + x^* = x^0 + X_m \sin \omega t \tag{6.96}$$

where:

x^0 is the constant component, shift of the center of oscillation influenced by the constant reference input.

$x^* = X_m sin\omega t$ is the periodic component of the input signal to nonlinear element $x(t)$, obtained as the result of established self-oscillations in the nonlinear system, and determined by parameters X_m and ω which remain nearly constant within one period.

By expanding nonlinear functions $F(x)$ or $F(x, px)$ into Fourier series, a nearly constant component F^0 appears. The describing function of the nonlinear element $y_N = F(x, px)$ is:

$$F(x, px) = F^0(x^0, X_m, \omega) + P(x^0, X_m, \omega)x^* + Q(x^0, X_m, \omega)x^* \tag{6.97}$$

where:

$$F^0(x^0,X_m,\omega) = \frac{1}{2\pi}\int_0^{2\pi} F(x^0,X_m\sin\omega t,X_m\omega\cos\omega t)d(\omega t)$$
$$P(x^0,X_m,\omega) = \frac{1}{\pi X_m}\int_0^{2\pi} F(x^0,X_m\sin\omega t,X_m\omega\cos\omega t)\sin\omega t d(\omega t) \quad (6.98)$$
$$Q(x^0,X_m,\omega) = \frac{1}{\pi X_m}\int_0^{2\pi} F(x^0,X_m\sin\omega t,X_m\omega\cos\omega t)\cos\omega t d(\omega t)$$

For the nonlinear element of type $y_N = F(x)$, equation (6.97) gets the form:

$$F(x) = F^0(x^0,X_m) + P(x^0,X_m)x^* + Q(x^0,X_m)x^* \quad (6.99)$$

where:

$$F^0(x^0,X_m) = \frac{1}{2\pi}\int_0^{2\pi} F(x^0,X_m\sin\omega t)d(\omega t)$$
$$P(x^0,X_m) = \frac{1}{\pi X_m}\int_0^{2\pi} F(x^0,X_m\sin\omega t)\sin\omega t d(\omega t) \quad (6.100)$$
$$Q(x^0,X_m) = \frac{1}{\pi X_m}\int_0^{2\pi} F(x^0,X_m\sin\omega t)\cos\omega t d(\omega t)$$

From (6.100), (6.95) and (6.96) the linearized differential equation with typical nonlinear element $y_N = F(x)$ can be written:

$$A(p)(x^0+x^*) + B(p)\left\{F^0(x^0,X_m) + \left[P(x^0,X_m) + \frac{Q(x^0,X_m)}{\omega}p\right]x^*\right\} = M^0 \quad (6.101)$$

From (6.101) follow the equations:

$$A(0)x^0 + B(0)F^0(x^0,X_m) = M^0 \quad (6.102)$$
$$A(p)x^* + B(p)\left[P(x^0,X_m) + \frac{Q(x^0,X_m)}{\omega}p\right]x^* = 0 \quad (6.103)$$

In the system of equations (6.102) and (6.103) the nonlinear relation is preserved between x^0, X_m and ω. The value $F^0(x^0,X_m)$ depends upon solutions of X_m, ω which arise by solution of the equation (6.103), while the coefficients $P(x^0,X_m)$ and $Q(x^0,X_m)$ solely depend upon solving x_0 from the equation (6.102).

From (6.102) and (6.103) a set of three equations[7] with three unknowns is obtained. The solution of these three equations gives the parameters x_0, X_m and ω of the periodic solution.

[7]From the second equation (6.103) two equations (real and imaginary parts) are obtained by substituting $p = s$ and $s = j\omega$.

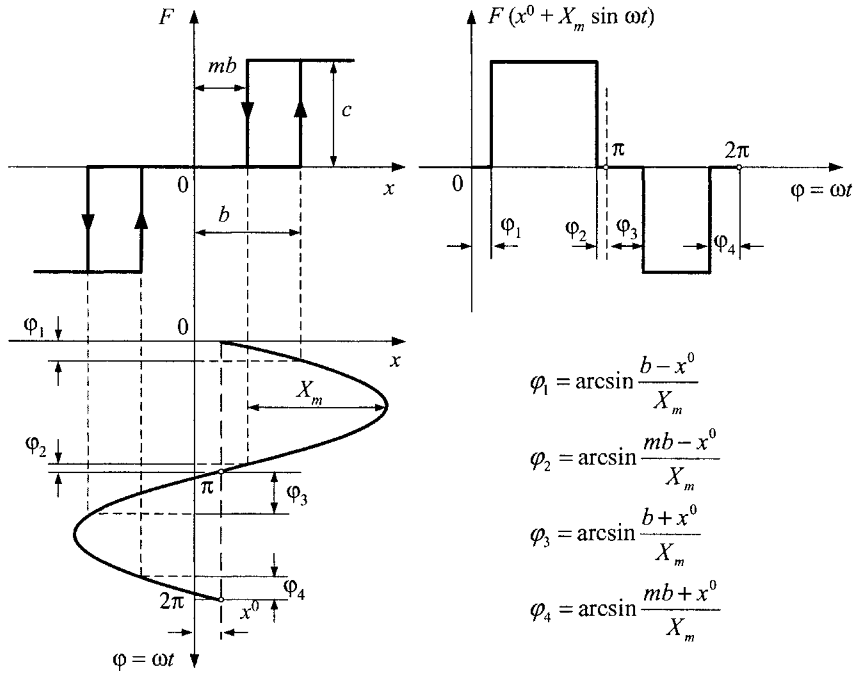


Figure 6.14: Input and output signals of a three-position relay with hysteresis.

EXAMPLE 6.4
Determine the coefficients of harmonic linearization of a general-type relay characteristic with asymmetrical harmonic input of the form $x(t) = x^0 + X_m sin\omega t$.

Solution. *A general-type relay characteristic with the input $x(t) = x^0 + X_m sin\omega t$ and the output $y_N = F(x) = F(x^0 + X_m sin\omega t)$ is shown in Fig. 6.14.*

Assuming that the amplitude of the harmonic component is $X_m \geq b + |x^0|$, it follows:

$$
\begin{aligned}
F^0(x^0, X_m) &= \frac{1}{2\pi} \int_0^{2\pi} F(x^0 + X_m \sin \omega t) d(\omega t) \\
&= \frac{c}{2\pi} \left[\int_{\varphi_1}^{\pi - \varphi_2} d(\omega t) - \int_{\pi + \varphi_3}^{2\pi - \varphi_4} d(\omega t) \right] \\
&= \frac{c}{2\pi} [\pi - \varphi_2 - \varphi_1 - (2\pi - \varphi_4 - \pi - \varphi_3)]
\end{aligned}
\tag{6.104}
$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

or:

$$F^0(x^0,X_m) = \frac{c}{2\pi}(\varphi_3 - \varphi_1 + \varphi_4 - \varphi_2) \tag{6.105}$$

Inserting in (6.105), the angles φ_1, φ_2, φ_3 and φ_4 (Fig. 6.14) results:

$$F^0(x^0,X_m) = \frac{c}{2\pi}\left[\sin^{-1}\frac{b+x^0}{X_m} - \sin^{-1}\frac{b-x^0}{X_m} + \sin^{-1}\frac{mb+x^0}{X_m} - \sin^{-1}\frac{mb-x^0}{X_m}\right] \tag{6.106}$$

Coefficients of the harmonic component are:

$$\begin{aligned}
P(x^0,X_m) &= \frac{1}{\pi X_m}\int_0^{2\pi} F(x^0 + X_m \sin\omega t)\sin\omega t d(\omega t)\\
&= \frac{c}{\pi X_m}\left[\int_{\varphi_1}^{\pi-\varphi_2}\sin\omega t d(\omega t) - \int_{\pi+\varphi_3}^{2\pi-\varphi_4}\sin\omega t d(\omega t)\right]\\
&= \frac{c}{\pi X_m}[-\cos(\pi-\varphi_2) + \cos\varphi_1 + \cos(2\pi-\varphi_4) - \cos(\pi+\varphi_3)]\\
&= \frac{c}{\pi X_m}(\cos\varphi_1 + \cos\varphi_2 + \cos s\varphi_3 + \cos\varphi_4)
\end{aligned}$$

and:

$$\begin{aligned}
Q(x^0,X_m) &= \frac{1}{\pi X_m}\int_0^{2\pi} F(x^0 + X_m \sin\omega t)\cos\omega t d(\omega t)\\
&= \frac{c}{\pi X_m}\left[\int_{\varphi_1}^{\pi-\varphi_2}\cos\omega t\cdot d(\omega t) - \int_{\pi+\varphi_3}^{2\pi-\varphi_4}\cos\omega t\cdot d(\omega t)\right]\\
&= \frac{c}{\pi X_m}[\sin(\pi-\varphi_2) - \sin\varphi_1 - \sin(2\pi-\varphi_4) + \sin(\pi+\varphi_3)]\\
&= \frac{c}{\pi X_m}(\sin\varphi_2 - \sin\varphi_1 + \sin\varphi_4 - \sin\varphi_3)
\end{aligned}$$

By substituting the expressions for angles (Fig. 6.14), it follows that:

$$\begin{aligned}
P(x^0,X_m) = \frac{c}{\pi X_m}&\left[\sqrt{1-\left(\frac{b-x^0}{X_m}\right)^2} + \sqrt{1-\left(\frac{mb-x^0}{X_m}\right)^2}\right.\\
&\left. + \sqrt{1-\left(\frac{b+x^0}{X_m}\right)^2} + \sqrt{1-\left(\frac{mb+x^0}{X_m}\right)^2}\right]
\end{aligned} \tag{6.107}$$

$$\begin{aligned}
Q(x^0,X_m) &= \frac{c}{\pi X_m}\left[\frac{mb-x^0}{X_m} - \frac{b-x^0}{X_m} + \frac{mb+x^0}{X_m} - \frac{b+x^0}{X_m}\right]\\
&= \frac{2cb}{\pi X_m^2}(m-1); \quad X_m \geq b + |x^0|
\end{aligned} \tag{6.108}$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

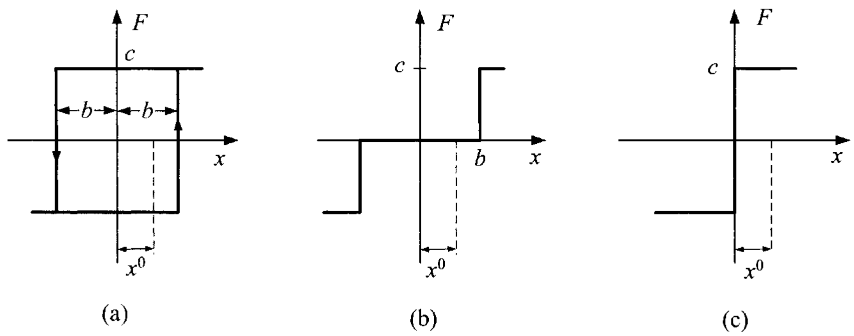


Figure 6.15: Relay characteristics: (a) with histeresis, (b) with dead zone, and (c) ideal two-position relay.

From the derived expressions for the coefficients of harmonic linearization of the general-type relay characteristic, the coefficients of harmonic linearization of relay characteristic with hysteresis, with dead zone and an ideal one (Fig. 6.15) can be found.

By inserting $m = -1$ in (6.106), (6.107) and (6.108), the coefficients of harmonic linearization of the relay characteristic with hysteresis (Fig. 6.15a) are obtained:

$$F^0(x^0, X_m) = \frac{c}{\pi}\left[\sin^{-1}\frac{b+x^0}{X_m} - \sin^{-1}\frac{b-x^0}{X_m}\right] \tag{6.109}$$

$$P(x^0, X_m) = \frac{2c}{\pi X_m}\left[\sqrt{1-\left(\frac{b-x^0}{X_m}\right)^2} + \sqrt{1-\left(\frac{b+x^0}{X_m}\right)^2}\right] \tag{6.110}$$

$$Q(x^0, X_m) = -\frac{4cb}{\pi X_m^2} \tag{6.111}$$

If $m = 1$ is put into (6.106), (6.107) and (6.108), the coefficients of harmonic linearization of relay characteristic with dead zone (Fig. 6.15b) are found:

$$F^0(x^0, X_m) = \frac{c}{\pi}\left[\sin^{-1}\frac{b+x^0}{X_m} - \sin^{-1}\frac{b-x^0}{X_m}\right] \tag{6.112}$$

$$P(x^0, X_m) = \frac{2c}{\pi X_m}\left[\sqrt{1-\left(\frac{b+x^0}{X_m}\right)^2} + \sqrt{1-\left(\frac{b-x^0}{X_m}\right)^2}\right] \tag{6.113}$$

$$Q(x^0, X_m) = 0 \tag{6.114}$$

Similarly with $b = 0$ the coefficients of harmonic linearization of the ideal relay characteristic (Fig. 6.15c) are:

$$F^0(x^0, X_m) = \frac{2c}{\pi} \sin^{-1} \frac{x^0}{X_m} \tag{6.115}$$

$$P(x^0, X_m) = \frac{4c}{\pi X_m} \sqrt{1 - \left(\frac{x^0}{X_m}\right)^2} \tag{6.116}$$

$$Q(x^0, X_m) = 0 \tag{6.117}$$

■

EXAMPLE 6.5

Determine the parameters of X_m, ω, x_0 of the periodic solution of the system as in Fig. 6.13, with ideal relay element (Fig. 6.15c) $y_N = F(x)$ and with coefficients of harmonic linearization from (6.115), (6.119) and (6.120). The linear part of the system is given by:

$$G_L(s) = \frac{K_2 [K_1 + K_0(T_1 s + 1)]}{s(T_2 s + 1)(T_1 s + 1)} = \frac{B(s)}{A(s)} \tag{6.118}$$

Solution. *The parameters of the periodic solution can be determined from the system of equations (6.102) and (6.103). From (6.115), (6.102) and (6.103) it follows that:*

$$A(0)x^0 + B(0)\frac{2c}{\pi} \sin^{-1} \frac{x^0}{X_m} = M^0 \tag{6.119}$$

$$A(p) + B(p)P(x^0, X_m) = 0 \tag{6.120}$$

From (6.119) we have:

$$\sin^{-1} \frac{x^0}{X_m} = \left[\frac{M^0 - A(0)x^0}{B(0)}\right] \frac{\pi}{2c} \tag{6.121}$$

i.e.:

$$\frac{x^0}{X_m} = \sin\left[\frac{\pi}{2c}\left(\frac{M^0 - A(0)x^0}{B(0)}\right)\right] \tag{6.122}$$

The next result comes from (6.119), (6.119) and (6.122):

$$P(x^0, X_m) = \frac{4c}{\pi X_m} \cos\left[\frac{\pi}{2c}\left(\frac{M^0 - A(0)x^0}{B(0)}\right)\right] \tag{6.123}$$

For:

$$G_L(s) = \frac{K_2 [K_1 + K_0(T_1 s + 1)]}{s(T_2 s + 1)(T_1 s + 1)} = \frac{B(s)}{A(s)} \tag{6.124}$$

is:

$$A(0) = 0;\ B(0) = K_2(K_1 + K_0) \tag{6.125}$$

Equations (6.122) and (6.123) give:

$$\frac{x^0}{X_m} = \sin\left[\frac{\pi}{2c}\left(\frac{M^0}{K_2(K_1+K_0)}\right)\right] \tag{6.126}$$

$$P(x^0, X_m) = \frac{4c}{\pi X_m}\cos\left[\frac{\pi}{2c}\left(\frac{M^0}{K_2(K_1+K_0)}\right)\right] \tag{6.127}$$

Replacing in (6.103) p with s, and s with $j\omega$, and introducing $A(j\omega)$ and $B(j\omega)$ from specified $G_L(j\omega)$ the complex equation is obtained:

$$R(x^0, X_m, \omega) + jI(x^0, X_m, \omega) = 0 \tag{6.128}$$

where:

$$\begin{aligned} R(x^0, X_m, \omega) &= (K_1 + K_0)K_2 P(x^0, X_m) - (T_1 + T_2)\omega^2 = 0 \\ I(x^0, X_m, \omega) &= \left[1 + T_1 K_2 K_0 P(x^0, X_m)\right]\omega - T_1 T_2 \omega^3 = 0 \end{aligned} \tag{6.129}$$

By eliminating $P(x^0, X_m)$ from the system of equations (6.129), the frequency of the periodic solution is obtained:

$$\omega^2 = \frac{K_1 + K_0}{T_1(T_2 K_1 - T_1 K_0)} \tag{6.130}$$

which is identical to (6.88). From (6.88) and (6.130) it is seen that the frequency of the periodic solution of the system with a single-valued nonlinear element depends exclusively on the parameters of the linear part of the system $G_L(s)$, i.e. on the gains and time constants. Equations (6.130) and (6.129) show that the frequency of self-oscillations at the input to the nonlinear element depends upon parameters of the linear part of the nonlinear control system. Including (6.130) into (6.129) the describing function is obtained:

$$P(x^0, X_m) = \frac{T_1 + T_2}{K_2 T_1(T_2 K_1 - T_1 K_0)} \tag{6.131}$$

From (6.88) and (6.131) it is also seen that the describing function $G_N(x^0, X_m) = P(x^0, X_m)$ of a single-valued nonlinear element can be expressed by the parameters of the linear part of the nonlinear control system. Equating (6.127) and (6.131) results in:

$$\frac{T_1 + T_2}{K_2 T_1(T_2 K_1 - T_1 K_0)} = \frac{4c}{\pi X_m}\cos\left[\frac{\pi}{2c}\left(\frac{M^0}{K_2(K_1 - K_0)}\right)\right] \tag{6.132}$$

respectively:

$$X_m = \frac{4cK_2T_1(T_2K_1 - T_1K_0)}{\pi(T_1+T_2)}\cos\left[\frac{\pi}{2c}\left(\frac{M^0}{K_2(K_1-K_0)}\right)\right] \tag{6.133}$$

From (6.126) by use of trigonometric relation $\sin x\cos x = \frac{1}{2}\sin 2x$ *follows:*

$$\begin{aligned} x^0 &= X_m \sin\left[\frac{\pi}{2c}\left(\frac{M^0}{K_2(K_1-K_0)}\right)\right] \\ &= \frac{4cK_2T_1(T_2K_1 - T_1K_0)}{\pi(T_1+T_2)}\cdot\frac{1}{2}\sin\left[\frac{\pi}{c}\left(\frac{M^0}{K_2(K_1-K_0)}\right)\right] \end{aligned} \tag{6.134}$$

and from (6.134) and (6.93) also follows:

$$x^0 = \frac{X_{ms}}{2}\sin\left[\frac{\pi}{c}\left(\frac{M^0}{K_2(K_1-K_0)}\right)\right] \tag{6.135}$$

Due to the fact that asymmetrical self-oscillations at the input of the nonlinear element have the form $x(t) = x^0 + X_m \sin\omega t$, *combining (6.133) and (6.135) the asymmetrical oscillations at the input of the nonlinear element are:*

$$x(t) = \frac{X_{ms}}{2}\sin\left[\frac{\pi}{c}\left(\frac{M^0}{K_2(K_1-K_0)}\right)\right] + X_{ms}\cos\left[\frac{\pi}{2c}\left(\frac{M^0}{K_2(K_1-K_0)}\right)\right]\sin\omega t \tag{6.136}$$

The conclusion from (6.129) and (6.136) is that the action of a nearly constant $r(t) = x^0 t \approx$ *const. on the nonlinear system with established self-oscillations leads to:*

- *Shift of the oscillation's center in relation to time axis and*
- *Influence on the amplitude of the oscillatory component at the input to the nonlinear part of the system.*

It may be concluded that the action of a nearly constant reference signal has an essential influence on the static accuracy of the system.

A possible periodic solution for the input signal to the nonlinear element is seen in Fig. 6.16. ■

6.1.6 Reliability of the Describing Function Method

The describing function method has had a successful application in analysis of nonlinear control systems for more than fifty years. Many practical problems in self-oscillating systems are solved by this method. As it is not an exact method, it is necessary to have this in mind, and the results are to be taken with caution. Three inaccuracies most often appear:

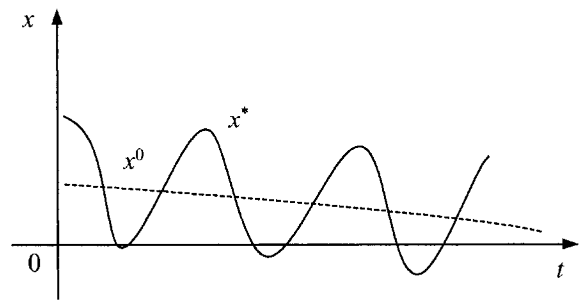


Figure 6.16: A possible periodic solution under the action of a nearly constant reference signal.

1. The amplitude and the frequency of self-oscillations are not evaluated accurately.
2. Self-oscillations do not exist in a real system, although they are predicted by the describing function method.
3. Self-oscillations exist in a real system, although they were not predicted by the describing function method.

The first inaccuracy appears quite often, because the method describes the oscillations as a pure sinusoidal signal with constant amplitude and frequency, which assumes that all higher harmonics are completely eliminated. How well the results of the analysis (frequency and amplitude of self-oscillations) coincide to accurate values depends on fulfilling the necessary conditions (filter hypothesis) to apply this method. In order to get accurate results, simulation of the nonlinear control system must take place.

The other two inaccuracies are rare, but they have grave consequences. For this reason, simulation must be carried out, too. Nevertheless, such inaccuracy can be foreseen if the frequency characteristic of the linear part of the system $G_L(j\omega)$ and relative position of the inverse describing function $-G_N^{-1}(X_M)$ are analyzed.

Filter hypothesis is not satisfied. The accuracy of the describing function method greatly depends on the filter hypothesis (6.4) being satisfied. Otherwise, no correct answers can be expected. It was shown that the describing function method gave wrong results with systems whose linear part has resonant overshoots in the frequency characteristic.

Graphical conditions. If the frequency characteristic $G_L(j\omega)$ is tangential or nearly tangential to $-G_N^{-1}(X_M)$, the conclusions of the describing function method can be wrong. The reason is that with neglected high-frequency harmonics or the inaccuracy of the mathematical model at high frequencies, the intersection point of $G_L(j\omega)$ and $G_N^{-1}(X_M)$ cannot be properly determined, especially in the case of bad filtering of the linear part. However, if the curves $G_L(j\omega)$ and $G_N^{-1}(X_M)$ intersect perpendicularly, the results of the describing function method are normally reliable.

As an example, the analysis of self-oscillations with the course-keeping control system for a ship when a nonlinear model of the rudder is used can be cited (Kuljača et al., 1983 and Reid et al., 1982). In this case the describing function method doesn't foresee the existence of self-oscillations, but the simulation and tests during the navigation point out the existence of self-oscillations. The reason for such a bad estimate is that the filter hypothesis was not satisfied in this particular case.

6.2 Forced Oscillations of Nonlinear Systems

When a nonlinear system is intentionally forced to oscillate with a desired frequency, such oscillations are called *forced oscillations*. When the system sustains self-oscillations with unfavorable frequency, there are certain methods to eliminate or to diminish undesirable effects of self-oscillations. One such method is the method of forced oscillations. It is based on the excitation of the closed-loop nonlinear control system by a forced harmonic signal $f(t) = F_v sin\omega_v t$. The frequency of this signal must be below the bandpass frequencies of the linear part of the system, so it can pass from input to the output of the control system[8]. The frequency of the input signal is either a desired one or some frequency more favorable than the existing frequency of self-oscillations. In this chapter will be discussed symmetrical forced oscillations (6.2.1) as well as asymmetrical forced oscillations (6.2.2), which appear in the systems with asymmetrical static characteristics of the nonlinear element.

6.2.1 Symmetrical Forced Oscillations

Contrary to self-oscillations, forced oscillations which can be established in nonlinear control systems can be caused by different properties of the system as well as by external signals which act on the system. Therefore, the methods to determine these processes are generally complex and cumbersome.

The superposition principle is not applicable, the solution cannot be achieved by superposition of individual solutions from various external actions. By anal-

[8] It must be noticed that—contrary to the dither signal whose frequency must be above the bandpass frequencies of linear system—here is expected that the system will pass an external harmonic signal. Such a signal is denoted by $f(t)$, so it can be distinguished from the dither signal, which is usually denoted by $d(t)$.

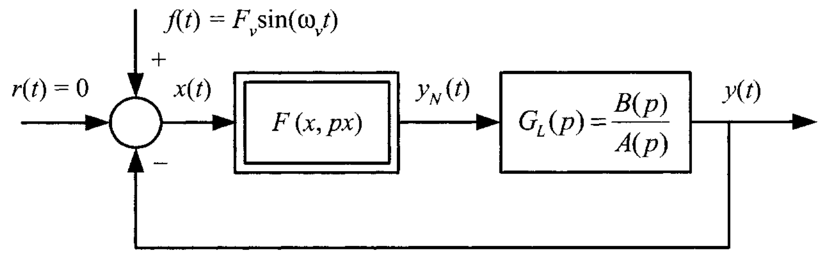


Figure 6.17: Block diagram of a nonlinear closed-loop control system subject to harmonic signal.

ogy, it is not possible to determine periodic solutions by superposition of inherent self-oscillations and established forced oscillations (Petrov et al., 1967; Solodovnikov, 1969; Besekerskii and Popov, 1975).

For a class of nonlinear systems with the structure as in Fig. 6.17, the monoharmonic forced oscillations can be determined by a mathematical procedure similar to the case of self-oscillations. Forced oscillations are established at the frequency ω_v of the external periodic signal.

Establishing monoharmonic forced oscillations with frequency ω_v is conditioned by the system's structure and parameters, as well as with limitations on the amplitude and the frequency of the external periodic signal. The area of a system's parameters which enables establishing monoharmonic forced oscillations is called *quenching area*, i.e. within this area self-oscillations are exterminated, while forced oscillations are established with the frequency ω_v of the known external signal $f(t)$. Symmetrical monoharmonic forced oscillations are determined by the solution of the differential equation of the closed-loop control system in Fig. 6.17.

The dynamics of the nonlinear control system in stabilization mode ($r(t) = 0$), which is excited by a harmonic signal $f(t) = F_v sin\omega_v t$ (Fig. 6.17), is described by the closed-loop differential equation. First, the series connection of linear and nonlinear parts are described by:

$$y(t) = G_L(p)F(x, px) \tag{6.137}$$

Second, by inserting $y(t) = f(t) - x(t)$ the closed-loop differential equation can be obtained:

$$x(t) + G_L(p) \cdot F(x, px) = f(t) \tag{6.138}$$

where: $G_L(p) = B(p)/A(p)$, $f(t) = F_v sin\omega_v t$:

$$A(p)x(t) + B(p)F(x, px) = A(p)f(t) \tag{6.139}$$

A periodic solution of the closed-loop differential equation (6.139) is sought in the form:

$$x(t) = X_m \sin(\omega_v t + \varphi) = X_m \sin\psi \tag{6.140}$$

where unknown quantities are the amplitude X_m and the phase φ of forced oscillations. In equation (6.139) the variables $x(t)$ and $f(t)$ are the functions of independent variable time t, so $f(t)$ can be expressed in terms of $x(t)$:

$$\begin{aligned} f(t) &= F_v \sin\omega_v t = F_v \sin[(\omega_v t + \varphi) - \varphi] \\ &= F_v \cos\varphi \sin(\omega_v t + \varphi) - F_v \sin\varphi \cos(\omega_v t + \varphi) \\ &= \frac{F_v}{X_m}\cos\varphi x(t) - \frac{F_v}{X_m}\cdot\frac{\sin\varphi}{\omega}px(t) \end{aligned} \tag{6.141}$$

respectively:

$$f(t) = \frac{F_v}{X_m}\left(\cos\varphi - \frac{\sin\varphi}{\omega}p\right)x(t) \tag{6.142}$$

Harmonic linearization of a symmetric nonlinearity $F(x,px)$ is according to (6.17) with known frequency $\omega = \omega_v$:

$$F(x,px) = \left[P(X_m,\omega_v) + \frac{Q(X_m,\omega_v)}{\omega_v}p\right]x(t) \tag{6.143}$$

Combining (6.143), (6.142) and (6.139), the harmonic linearized differential equation of the nonlinear closed-loop control system follows:

$$\left[A(p) - A(p)\frac{F_v}{X_m}\left(\cos\varphi - \frac{\sin\varphi}{\omega_v}p\right)\right]x + B(p)\left[P(X_m,\omega_v) + \frac{Q(X_m,\omega_v)}{\omega_v}p\right]x = 0 \tag{6.144}$$

This is a homogenous differential equation with the characteristic equation of first approximation:

$$A(s) - A(s)\frac{F_v}{X_m}\left(\cos\varphi - \frac{\sin\varphi}{\omega_v}s\right) + B(s)\left[P(X_m,\omega_v) + \frac{Q(X_m,\omega_v)}{\omega_v}s\right] = 0 \tag{6.145}$$

For $s = j\omega_V$, from (6.145) the equation for determining the parameters of forced oscillations X_m and φ is derived:

$$X_m \cdot \frac{A(j\omega_v) + B(j\omega_v)\left[P(X_m,\omega_v) + jQ(X_m,\omega_v)\right]}{A(j\omega_v)} = F_v e^{-j\varphi} \tag{6.146}$$

The solution of (6.146) can be found either graphically or analytically. For the analytical solution the equation (6.146) is transformed into the form:

$$\begin{aligned} &X_m \cdot \frac{R(X_m,\omega_v) + jI(X_m,\omega_v)}{R_A(\omega_v) + jI_A(\omega_v)} \\ &= X_m \cdot \frac{\sqrt{R^2(X_m,\omega_v) + I^2(X_m,\omega_v)}}{\sqrt{R_A^2(\omega_v) + I_A^2(\omega_v)}} \cdot \frac{e^{j\tan^{-1}\frac{I}{R}}}{e^{j\tan^{-1}\frac{I_A}{R_A}}} = F_v e^{-j\varphi} \end{aligned} \tag{6.147}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

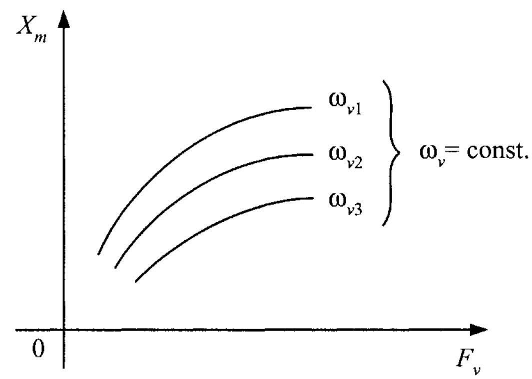


Figure 6.18: The amplitudes of oscillations at the input to the nonlinear element depend on the amplitude and frequency of the exciting harmonic signal to the nonlinear closed-loop control system.

where:

$R(X_m, \omega_v)$ is the real part of the nominator in (6.146),

$I(X_m, \omega_v)$ is the imaginary part of the denominator in (6.146),

$R_A(\omega_v)$ is the real part of the denominator $A(j\omega)$ in (6.146),

$I_A(\omega_v)$ is the imaginary part of the denominator $A(j\omega)$ in (6.146).

From (6.147) follow the equations for determining X_m and φ:

$$X_m = F_v \frac{\sqrt{R_A^2(\omega_v) + I_A^2(\omega_v)}}{\sqrt{R^2(X_m, \omega_v) + I^2(X_m, \omega_v)}} \tag{6.148}$$

$$\varphi = \tan^{-1}\frac{I_A}{R_A} - \tan^{-1}\frac{I}{R} \tag{6.149}$$

From (6.148) we can find $X_m(F_v)$ for $\omega_v = const.$ and $X_m(\omega_v)$ for $F_v = const.$, or generally the functional relation of X_m with some system parameter K, $X_m(K)$ for a given F_v and ω_v. Afterwards, the phase angle φ is found from (6.149). A possible functional relation of $X_m(F_v)$ with ω_v=const. is graphically presented in Fig. 6.18.

By implementing the graphical procedure for a given value of the frequency $\omega = \omega_v$, the curve $Z(X_m)$ is plotted in the polar plane:

$$Z(X_m) = X_m \frac{A(j\omega_v) + B(j\omega_v)\left[P(X_m, \omega_v) + jQ(X_m, \omega_v)\right]}{A(j\omega_v)} \tag{6.150}$$

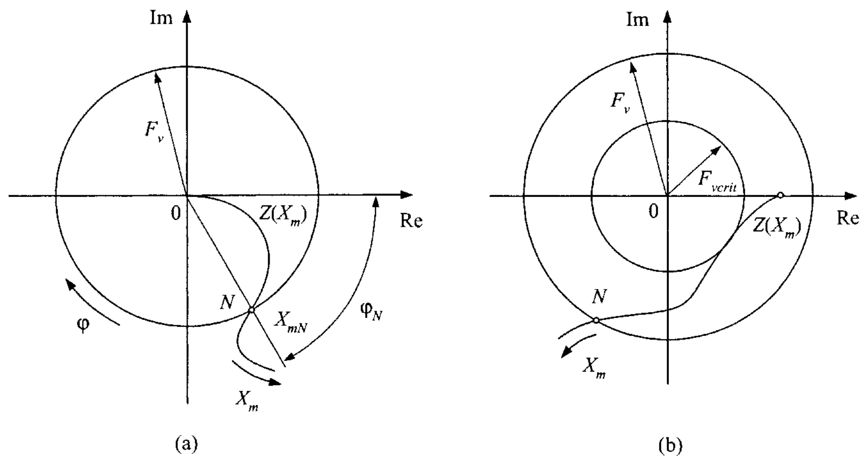


Figure 6.19: Graphical procedure of determining the parameters (amplitude and frequency) of forced oscillations, (a) for $F_{vcrit} = 0$, (b) for $F_{vcrit} > 0$.

At the intersection point N of the circle of the radius F_v and the curve $Z(X_m)$, the amplitude $X_m = X_{mN}$ and phase angle $\varphi = \varphi_N$ of forced oscillations can be found. The amplitude is read along the curve $Z(X_m)$ and the phase angle φ along the circle.

In the case when the curve $Z(X_m)$ is as in Fig. 6.19a, monoharmonic self-oscillations with the frequency ω_v exist in the system at any amplitude of the external periodic value of F_v. If the curve $Z(X_m)$ has the form as in Fig. 6.19b, forced oscillations will emerge only if the amplitude of the external periodic signal is greater than some boundary value, i.e. $F_v > F_{vcrit}$. If the condition $F_v > F_{vcrit}$ is not satisfied, forced periodic oscillations with the frequency ω_v will not exist in the control system, but instead (in general) some complex periodic process can be established.

The dependence of the parameters of forced oscillations upon the frequency ω_v or upon any parameter K of the control system is determined by a graphical procedure (Fig. 6.20).

In order to find the relation between the amplitude X_m and the amplitude of any external action F_v, a graphical plot is constructed (Fig. 6.21).

From the intersection points of the curve $Z(X_m)$ with concentric circles of the radius F_v, two cases for the dependence of $X_m(F_v)$ are possible, as shown in Fig. 6.21b and 6.21c. For the case (b) for every F_v the forced oscillatons will be established in the control system (Fig. 6.21b). However, for the case (c) this will depend on F_{vcrit} (Fig. 6.21c). If the amplitude of the external periodic signal is

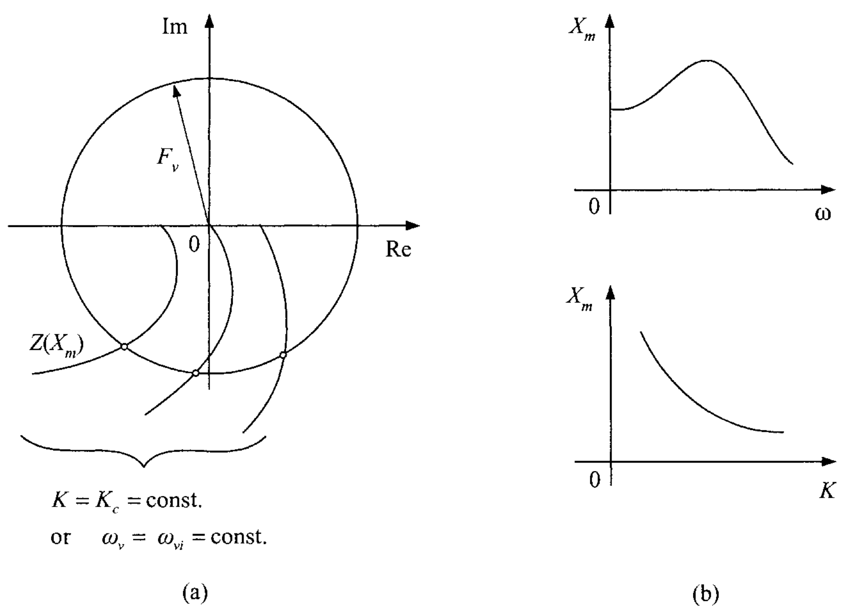


Figure 6.20: Graphical procedure to determine how the parameters of forced oscillations depend on exciting frequency or some other parameter of control system: (a) $Z(X_m)$ for various $K = const.$ or $\omega_v = const.$ (b) $X_m = X_m(\omega_v)$ and $X_m = X_m(K)$ according to graphical solution of (6.150).

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

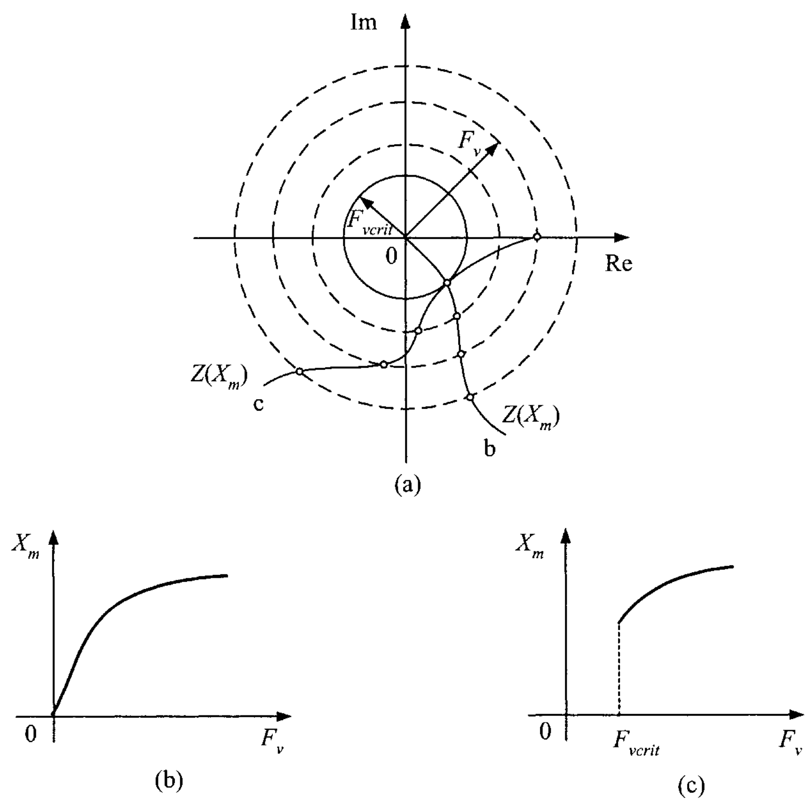


Figure 6.21: Graphical procedure of finding how the amplitudes of forced oscillations depend on excitation amplitude, F_v: (a) $Z(X_m)$ for $F_{vcrit} = 0$ and $F_v > 0$, (b) $X_m = X_m(F_v)$ for $F_{vcrit} = 0$, and (c) $X_m = X_m(F_v)$ for $F_{vcrit} > 0$.

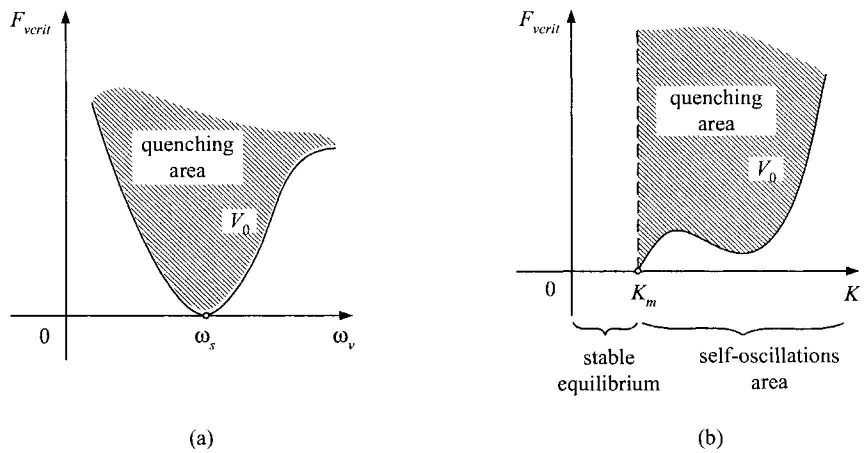


Figure 6.22: Typical forms of quenching areas.

smaller than the boundary value, i.e. $F_v < F_{vcrit}$, forced oscillations will not exist in the system, and the existing self-oscillations will not be *quenched by forced oscillations*. A typical form of quenching areas in the plane (ω_v, F_v) is shown in Fig. 6.22.

In Fig. 6.22. the diagram shows the relation between the external periodic signal F_v and the frequency ω_v with given parameters of the system. From the display it is visible that when the frequency of the external signal ω_v is equal to the frequency of self-oscillations ω_{s0}, the amplitude of the external signal is $F_{vcrit} = 0$. Fig 6.22b shows how F_{vcrit} depends upon the variable parameter of the system K, where ω_v is given.

The quenching conditions (Fig. 6.22a and 6.22b) enable us to determine the amplitudes F_v and frequencies ω_v of the external periodic signal, which make possible the elimination of unwanted self-oscillations and the restoration of forced oscillations with permissible amplitude and frequency.

Generally, at the intersection points of the curve $Z(X_m, \omega_v, K)$ and the circles with various radii F_v, the dependence of parameters of forced oscillations X_m and ω_v upon the amplitude of the external signal F_v and a certain parameter K of the control system are determined; thus, the synthesis of the corresponding nonlinear compensator is made possible.

EXAMPLE 6.6

For the nonlinear closed-loop control system with block diagram as in Fig. 6.23, parameters of forced oscillations X_m and φ are to be found if $K = 10[V]$, $b = 4[V]$,

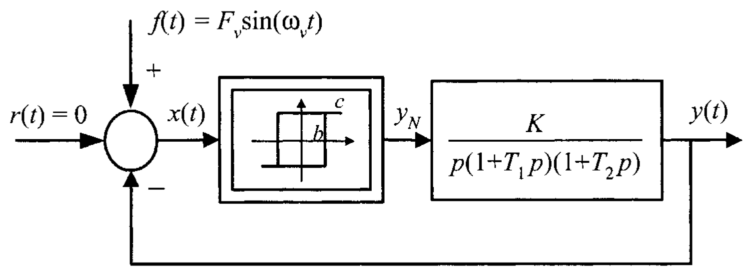


Figure 6.23: Block diagram of a nonlinear closed-loop control system in stabilization mode.

$T_1 = 0.01[s]$, $T_2 = 0.02[s]$, $F_v = 20[V]$, $\omega_v = 10[s^{-1}]$ *and* $c = 10[V]$.

Solution. *The differential equation of the series connection of the nonlinear and linear part in Fig. 6.23 is:*

$$y(t) = G_L(p) \cdot F(x) = \frac{K}{p(T_1 p + 1)(T_2 p + 1)} \cdot F(x) \tag{6.151}$$

By putting $y(t) = f(t) - x(t)$ *in equation (6.151) the closed-loop differential equation can be obtained:*

$$x(t) + \frac{K}{p(T_1 p + 1)(T_2 p + 1)} \cdot F(x) = f(t) \tag{6.152}$$

respectively:

$$(T_1 p + 1)(T_2 p + 1) p x(t) + K \cdot F(x) = (T_1 p + 1)(T_2 p + 1) p \cdot f(t) \tag{6.153}$$

From (6.139) and (6.153) follows:

$$A(p) = (T_1 p + 1)(T_2 p + 1) p; \; B(p) = K \tag{6.154}$$

Substituting for $f(t)$ *the expression (6.142) in the equation (6.153), the homogenous differential equation of the closed-loop control system is obtained:*

$$(T_1 p + 1)(T_2 p + 1) p \left[1 - \frac{F_v}{X_m} \left(\cos\varphi - \frac{\sin\varphi}{\omega_v} p \right) \right] x(t) + KF(x) = 0 \tag{6.155}$$

By inserting in (6.155) the expression for $F(x)$*:*

$$F(x) = \left[P(X_m) + \frac{Q(X_m)}{\omega_v} p \right] x \tag{6.156}$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

the characteristic equation of the linearized closed-loop control system is written down:

$$(T_1p+1)(T_2p+1)p\left[1-\frac{F_v}{X_m}\left(\cos\varphi-\frac{\sin\varphi}{\omega_v}p\right)\right]+K\left[P(X_m)+\frac{Q(X_m)}{\omega_v}p\right]=0 \tag{6.157}$$

When the derivative operator p is replaced by $s=j\omega_v$ *in the above equation and the equation is rearranged, the complex characteristic equation ensues:*

$$X_m\left[1-\frac{K\left[P(X_m)+jQ(X_m)\right]}{(T_1+T_2)\omega_v^2-j(1-T_1T_2\omega_v^2)\omega_v}\right]=F_v(\cos\varphi-j\sin\varphi)=F_ve^{-j\varphi} \tag{6.158}$$

The coefficients of harmonic linearization of given nonlinearity $F(x)$ *are* [9]*:*

$$P(X_m)=\frac{4c}{\pi X_m}\sqrt{1-\frac{b^2}{X_m^2}};\ Q(X_m)=-\frac{4bc}{\pi X_m^2};\ X_m\geq b$$

These values are inserted in the equation (6.158) and this is rearranged:

$$Z(X_m)=X_m-\frac{3.63\sqrt{X_m^2-16}+47,5+j\left(11,8\sqrt{X_m^2-16}-14,5\right)}{X_m}=20e^{-j\varphi} \tag{6.159}$$

The graphical procedures of finding the amplitude X_m *and phase angle* φ *of forced oscillations are applied as follows:*

In the polar plane the circle of radius $F_v=20$ *is drawn, which corresponds to the right side of equation (6.159) (Fig. 6.24a). By plotting various* X_m *values, the function* $Z(X_m)$ *is drawn in the polar plane too. The intersection point of these two plots offers the solution of the equation (6.159)* $X_m=21[V]$ *and* $\varphi=35[\mathrm{deg}]$.

Various values of F_v *in the polar plane have corresponding circles of radius* F_v*. At the intersection points with* $Z(X_m)$*, the parameters* X_m *and* φ *as functions of the amplitude of the external periodic signal are found* [10] *(Fig. 6.24b)—by analogy, the relations between any quantities of the system and the system parameters can be determined.* ■

6.2.2 Asymmetrical Forced Oscillations

In a nonlinear control system with the structure as in Fig. 6.23, asymmetrical forced oscillations are possible, caused by an asymmetrical static characteristic of the nonlinear element $F(x)$ or $F(x,px)$. The differential equation of the closed-loop control system is like equation (6.139) (Popov and Pal'tov, 1960; Petrov et al., 1967; Netushil, 1983; Cook, 1986):

$$A(p)x(t)+B(p)F(x,px)=A(p)\cdot f(t) \tag{6.160}$$

[9] See equation (6.109) for $x^0=0$.

[10] Note the existence of $F_{vcrit}>0$ in this example even though it seems from Fig. 6.24 as $F_{vcrit}=0$.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

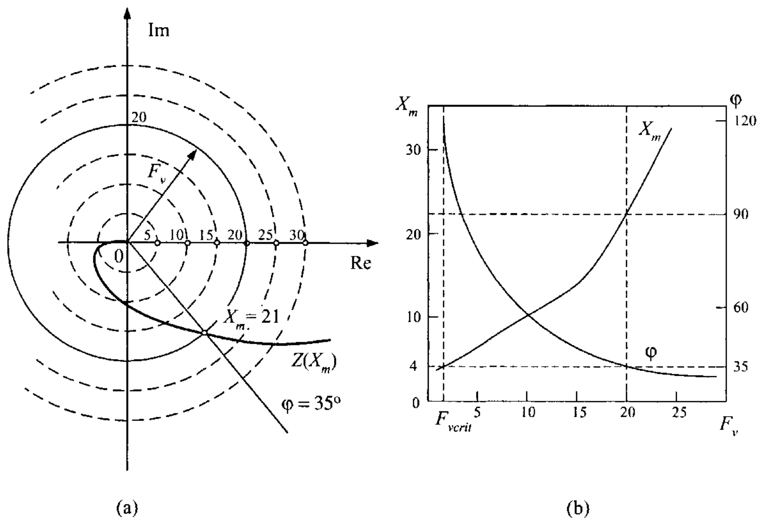


Figure 6.24: Graphical display of solving the problem of forced oscillations for Example 6.6.

The solution of the above differential equation for nonlinear forced oscillations is sought in the form:

$$x(t) = x^0(t) + x^*(t) = x^0(t) + X_m \sin(\omega_v t + \varphi) \tag{6.161}$$

Harmonic linearization of the asymmetrical nonlinearity $F(x, px)$ is determined by expressions (6.57), (6.58) and (6.59) for $\omega = \omega_v$:

$$F(x, px) \approx F^0(x^0, X_m, \omega_v) + \left[P(x^0, X_m, \omega_v) + \frac{Q(x^0, X_m, \omega_v)}{\omega_v} p\right] x \tag{6.162}$$

where:

$$F^0 = \frac{1}{2\pi} \int_0^{2\pi} (x^0 + X_m \sin\psi, X_m \omega_v \cos\psi) d\psi \tag{6.163}$$

$$P = \frac{1}{\pi X_m} \int_0^{2\pi} (x^0 + X_m \sin\psi, X_m \omega_v \cos\psi) \sin\psi d\psi \tag{6.164}$$

$$Q = \frac{1}{\pi X_m} \int_0^{2\pi} (x^0 + X_m \sin\psi, X_m \omega_v \cos\psi) \cos\psi d\psi \tag{6.165}$$

with $\psi = \omega t + \varphi$.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

x^0 is the nonoscillatory component of the periodic solution—the shift of the center of forced oscillations caused by the asymmetrical nonlinear characteristic $y_N = F(x, px)$.

From (6.142), (6.160) and (6.162) follows the harmonically linearized differential equation of the system:

$$A(p)(x^0 + x^*) + B(p)\left[F^0 + Px^* + \frac{Q}{\omega_v} px^*\right] = A(p)\frac{F_v}{X_m}\left[\cos\varphi - \frac{\sin\varphi}{\omega_v} p\right] x^* \tag{6.166}$$

As with the procedure for determining parameters of asymmetrical self-oscillations, the solution of the system (6.63), (6.64) and (6.166) gives the equations for determining the nonoscillatory $x^0(t)$ and periodic components (X_m and φ) of asymmetrical forced oscillations (6.161):

$$A(p)x^0 + B(p)F^0(x^0, X_m, \omega_v) = 0 \tag{6.167}$$

$$\left[A(p) - A(p)\frac{F_v}{X_m}\left(\cos\varphi - \frac{\sin\varphi}{\omega_v} p\right) + B(p)\left(P + \frac{Q}{\omega_v} p\right)\right] x^* = 0 \tag{6.168}$$

respectively:

$$A(0)x^0 + B(0)F^0(x^0, X_m, \omega_v) = 0 \tag{6.169}$$

$$X_m \frac{A(j\omega_v) + B(j\omega_v)\left[P(x^0, X_m, \omega_v) + jQ(x^0, X_m, \omega_v)\right]}{A(j\omega_v)} = F_v e^{-j\varphi} \tag{6.170}$$

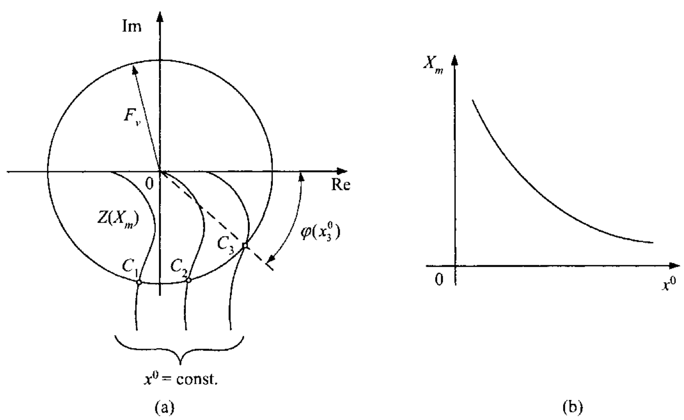


Figure 6.25: Graphical procedure to determine the parameters of asymmetrical forced oscillations: (a) graphical solution for X_m given by various $x^0 = const.$, (b) plot of $X_m = X_m(x^0)$ obtained from procedure in (a).

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

From the last equation, the functions:

$$X_m(x^0, \omega_v, F_v); \; \varphi(x^0, \omega_v, F_v) \tag{6.171}$$

are found with x^0 as the unknown quantity.

For the graphical method of solving (6.170), we need the graphical display (Fig. 6.19) of a series of curves $Z(X_m)$ drawn for various values of x^0 (Fig. 6.25).

6.2.3 Resonance Jump

The resonance jump, as described in Section 1.3, Figure 1.6, can be observed through analysis of frequency characteristics of nonlinear system in Fig. 6.26.

The quantities from Fig. 6.26 are:

$f(t) \;= F_v \sin(\omega_v t)$ is the harmonic input to the system

$x(t) \;= X_m \sin(\omega_v t + \varphi_x)$

$F(x, px)$ is the nonlinear part of the system

$G_L(p) = B(p)/A(p)$ is the linear part of the system

The most suitable method to determine the resonance jump of the nonlinear system shown in Fig. 6.26 is the describing function method. The process of determining nonlinear resonance is based on equation (6.146), which is for this purpose rewritten as:

$$X_m\left[1 + G_L(j\omega_v)G_N(X_m, \omega_v)\right] = F_v e^{-j\varphi_x} \tag{6.172}$$

Frequency properties of the system in Fig. 6.26 can be determined by simulation or by grapho-analytical and analytical solutions of equation (6.172). Considering that the linear part of the system $G_L(j\omega)$ often contains in the denominator the polynomial of higher order $A(j\omega)$, it is best to determine the frequency characteristics of the difference signal $X_m(\omega_v)$ by grapho-analytical procedure. In cases when the frequency characteristic $X_m(\omega_v)$ has the properties as the function depicted in Fig. 1.6, it is possible to observe the resonance jump in the system. For

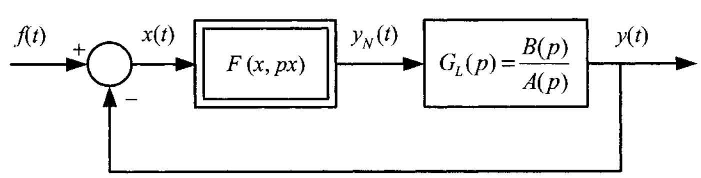


Figure 6.26: Block diagram of a forced system.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

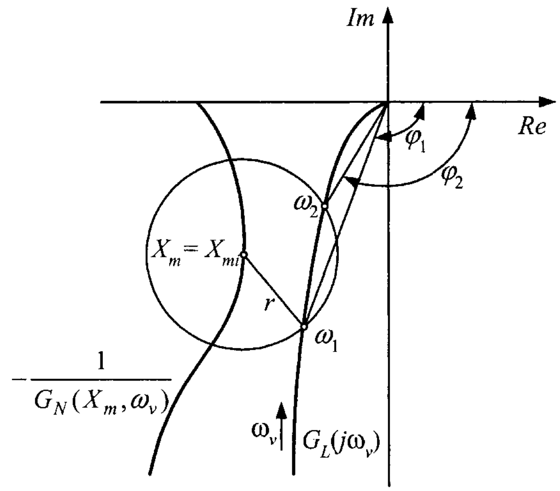


Figure 6.27: The determination of $X_m(\omega_v)$.

grapho-analytical determination of resonance frequencies ω_{v1} and ω_{v2}, the equation (6.172) is rewritten as (Tsypkin, 1977):

$$\frac{F_v e^{-j\varphi_x}}{X_m G_N(X_m, \omega_v)} - \frac{1}{G_N(X_m, \omega_v)} = G_L(j\omega_v); \;\; \varphi = \varphi_x \tag{6.173}$$

For grapho-analytical solution of (6.173), the frequency characteristic $G_L(j\omega_v)$ and $-G_N^{-1}(X_m, \omega_v)$ have to be drawn in the complex plane $[U, jV]$, see Fig. 6.27.

On the graphical display of $-G_N^{-1}(X_m, \omega_v)$ the point $X = X_m$ is determined. This point is the center of the circle with radius:

$$r = \left| \frac{F_v e^{-j\varphi_x}}{X_m G_N(X_m, \omega_v)} \right| = \frac{F_v}{X_m \left| G_N(X_m, \omega_v) \right|} \tag{6.174}$$

For different values $X_m = X_{mi}$, respective circle radii are determined using (6.174). A circle with radius r_i and center at $X_m = X_{mi}$ is drawn on $-G_N^{-1}(X_m, \omega_v)$. At the points of intersection of the circle and the frequency characteristic $G_L(j\omega_v)$ the frequencies $\omega_v = \omega_1$ and $\omega_v = \omega_2$ are read out, with corresponding amplitude $X_m = X_{mi}$. If the obtained frequency characteristic has properties as shown in Fig. 1.6, the appearance of resonance jump is possible in the system. By increasing the frequency $\omega_v > \omega_{v2}$, the amplitude X_m very rapidly increases, and by decreasing the frequency $\omega_v < \omega_{v1}$, the amplitude X_m suddenly decreases[11] Analogous to

[11]The frequency characteristic $X_m(\omega_v)$ can also be determined with the procedure described in Section 6.2.1, Fig. 6.20.

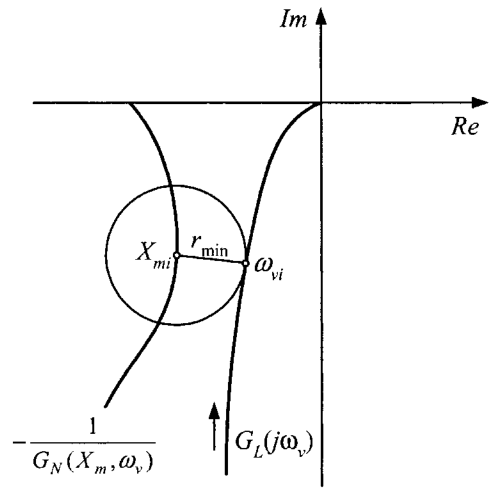


Figure 6.28: The determination of F_{vcrit} for given value of ω_v.

graphical display in Fig. 6.19, the critical value of the periodic exciting signal can be determined grapho-analytically (Fig. 6.28).

At the point of $G_L(j\omega_v)$ for specific frequency of harmonic input $\omega_v = \omega_{vi}$, the circle that tangentially touches the graph $-G_N(X_m, \omega_v)$ is drawn. The radius of this circle $r_{\min}$ is determined by the expression:

$$r_{\min}(\omega_v) = \left| \frac{F_v e^{j\varphi}}{X_m G_N(X_m, \omega_v)} \right| \tag{6.175}$$

The equation (6.173) will be satisfied for values of ω, X_m and F_v for which the following is true:

$$r(\omega_v) \geq r_{\min}(\omega_v) \tag{6.176}$$

From (6.175) the expression for the critical value of the harmonic input's amplitude F_v is obtained:

$$F_{vcrit} = r_{\min}(\omega_v) G_N(X_m, \omega_v) \tag{6.177}$$

The area of possible appearance of resonance jump can also be determined from the dependency $F_v = f(X_m)$ for $\omega_v = const$. The shape of $F_v = f(X_m)$ significantly depends on nonlinear element properties, Fig. 6.29.

The resonance jump can appear only in the case when the characteristic $F_v = f(X_m)$ is non-unique (Fig. 6.29).

It can be seen from Fig. 6.29 that for $F_v \geq F_{v1}$ the amplitude of forced oscillations increases, while for $F_v \leq F_{v1}$ the amplitude decreases. From this fact it

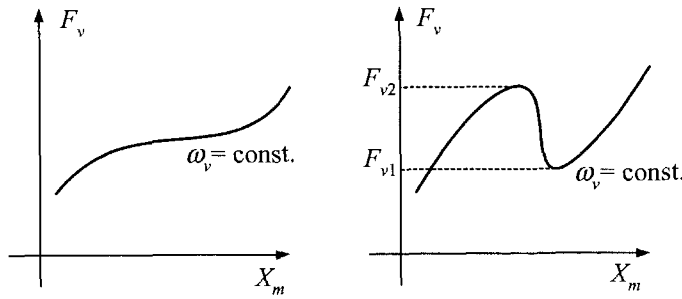


Figure 6.29: Qualitative display of $F_v = f(X_m)$.

follows that the condition of appearance of resonance jump is determined with the following relation:

$$\left(\frac{dF_v}{dX_m}\right)_{\omega=\text{const.}} \leq 0 \tag{6.178}$$

For the class of nonlinear systems in Fig. 6.26, for unique nonlinearity $y_N = F(x)$, the condition (6.178) follows from (6.172). In (6.172), the nonlinearity $y_N = F(x)$ is substituted with describing function $F(x) \approx G_N(X_m) = P(X_m)$. When we substitute $G_L(j\omega_v) = U(j\omega_v) + jV(\omega_v)$ into (6.172), the following is obtained:

$$F_v e^{j\varphi_x} = X_m P(X_m)\left[\frac{1}{P(X_m)} + U(\omega_v) + jV(\omega_v)\right] \tag{6.179}$$

The modulo of expression (6.179) is:

$$F_v = X_m P(X_m)\sqrt{\left(\frac{1}{P(X_m)} + U(\omega_v)\right)^2 + V^2(\omega_v)} \tag{6.180}$$

From (6.180) and (6.178) the condition for existence of resonance jump is obtained:

$$N(X_m, \omega_v) = \left[U(\omega_v) + \frac{1}{P(X_m)}\right]\left[U(\omega_v) + \frac{1}{P(X_m) + X_m \frac{dP}{dX_m}}\right] + V^2(\omega_v) \leq 0 \tag{6.181}$$

The equation $N(X_m, \omega_v)$ is the equation of the circle in complex plane with radius:

$$r = \frac{1}{2}\left[\frac{1}{P(X_m) + X_m \frac{dP}{dX_m}} - \frac{1}{P(X_m)}\right] \leq 0 \tag{6.182}$$

and center at:

$$\left[-\frac{1}{2}\left(\frac{1}{P(X_m) + X_m \frac{dP}{dX_m}} + \frac{1}{P(X_m)}\right), 0\right] \tag{6.183}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

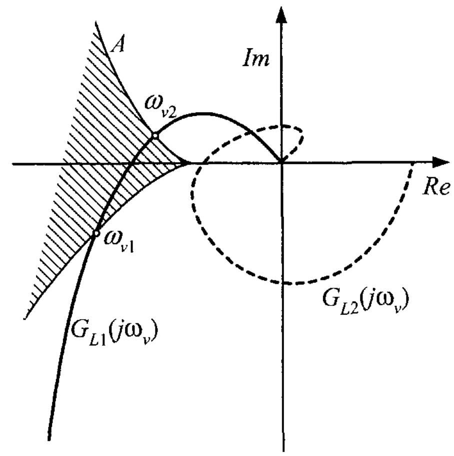


Figure 6.30: Qualitative display of frequency area where resonance jump is possible.

For different values of X_m = const., the family of curves in complex plane $[U(\omega_v), V(\omega_v)]$ is obtained. The area in which resonance jumps of amplitude $X_m(\omega_v)$ can occur is determined by the envelope of the family of circles. Parametric equation of the envelope is obtained by solving the system of equations:

$$\begin{aligned} N(X_m, \omega_v) &= 0 \\ \frac{dN(X_m, \omega_v)}{dX_m} &= 0 \end{aligned} \tag{6.184}$$

From (6.181) and (6.184) it follows:

$$\begin{aligned} U(\omega_v) &= -\frac{P\left(2\frac{dP}{dX_m} + X_m\frac{d^2P}{dX_m^2}\right) + \frac{dP}{dX_m}\left(P + X_m\frac{dP}{dX_m}\right)}{P^2\left(2\frac{dP}{dX_m} + X_m\frac{d^2P}{dX_m^2}\right) + \frac{dP}{dX_m}\left(P + X_m\frac{dP}{dX_m}\right)^2} \\ V(\omega_v) &= \pm\frac{-X_m\frac{dP}{dX_m}\sqrt{\frac{dP}{dX_m}\left(2\frac{dP}{dX_m} + X_m\frac{d^2P}{dX_m^2}\right)}}{P^2\left(2\frac{dP}{dX_m} + X_m\frac{d^2P}{dX_m^2}\right) + \frac{dP}{dX_m}\left(P + X_m\frac{dP}{dX_m}\right)^2} \end{aligned} \tag{6.185}$$

The qualitative graphical display of the envelope (6.185) is shown in Fig. 6.30.

The condition of occurrence of resonance jump of forced oscillations $X_m(\omega_v)$ is that the envelope A intersects with the characteristic $G_L(j\omega)$. From Fig. 6.30 it can be concluded that the resonance jump is only possible for the part of linear

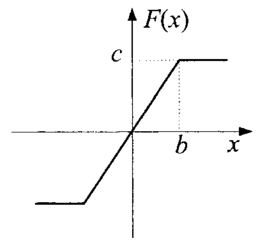


Figure 6.31: Static characteristic of the nonlinearity in Example 6.7.

characteristic G_{L1} in the area $(\omega_{v1}, \omega_{v2})$, while for the part G_{L2} the conditions for occurrence of resonance jump are not fulfilled.

Analogous to presented procedures for the analysis of the resonance jump, different analytical and simulation procedures are also used.

EXAMPLE 6.7

For the system on Fig. 6.26, determine the frequency characteristic $X_m(\omega_v)$ and conditions for the occurrence of resonance jump. The static characteristic of the nonlinear element is shown in Fig. 6.31, and the linear part of the system is given by following transfer function:

$$G_L(s) = \frac{K}{s(s+1)} = \frac{200}{s(s+1)}$$

The system is subjected to harmonic input signal $F_v \sin \omega_v t = 20 \sin \omega_v t$.

Solution. *The system equation, according to (6.172), is:*

$$X_m(1 + G_L(j\omega_v)G_N(X_m)) = F_v e^{-j\varphi} \tag{6.186}$$

$$G_L(j\omega_v) = \frac{K}{j\omega_v(j\omega_v+1)} = -\frac{K}{1+\omega_v^2} - j\frac{K}{\omega_v(1+\omega_v^2)} = U(\omega_v) + jV(\omega_v) \tag{6.187}$$

Substituting (6.187) into (6.186) yields:

$$X_m\left(1 - \frac{K\omega_v + jK}{\omega_v(1+\omega_v^2)}G_N(X_m)\right) = F_v e^{-j\varphi_x}$$

i.e.:

$$X_m\left(1 - \frac{K}{1+\omega_v^2}G_N(X_m)\right) = F_v \cos\varphi_x \tag{6.188}$$

$$\frac{X_m K}{\omega_v(1+\omega_v^2)}G_N(X_m) = F_v \sin\varphi_x \tag{6.189}$$

From (6.188) and (6.189) the expression which determines the frequency characteristic of difference signal $X_m(\omega_v)$ is obtained:

$$X_m(\omega_v) = \frac{F_v}{\sqrt{\left(1 - \frac{K}{1+\omega_v^2} G_N(X_m)\right)^2 + \frac{K^2 G_N^2(X_m)}{\omega_v^2(1+\omega_v^2)^2}}} \tag{6.190}$$

$$\varphi_x = \arctan \frac{K G_N(X_m)}{\omega_v(1+\omega_v^2) - K G_N(X_m)} \tag{6.191}$$

The describing function of nonlinearity in Fig. 6.31 is:

$$G_N(X_m) = P(X_m) = \frac{2}{\pi}\left(\arcsin\frac{b}{X_m} + \frac{b}{X_m}\sqrt{1 - \frac{b^2}{X_m^2}}\right); X_m \geq b \tag{6.192}$$

To determine the frequency area where the occurrence of resonance jump is possible, the members of parametric equation (6.185) have to be determined. From (6.182) it follows:

$$\frac{dP}{dX_m} = \frac{4b}{\pi} \frac{b^2 - X_m^2}{X_m^3\sqrt{X_m^2 - b^2}} \tag{6.193}$$

$$\frac{d^2P}{dX_m^2} = \frac{4b}{\pi} \frac{2X_m^2 - 3b^2}{X_m^4\sqrt{X_m^2 - b^2}} \tag{6.194}$$

From (6.192), (6.193) and (6.194) for $b = 2$ we obtain:

$$P(X_m) = \frac{2}{\pi}\left(\arcsin\frac{2}{X_m} + \frac{2\sqrt{X_m^2 - 4}}{X_m^2}\right)$$

$$\frac{dP}{dX_m} = \frac{8}{\pi} \frac{4 - X_m^2}{X_m^3\sqrt{X_m^2 - 4}}$$

$$\frac{d^2P}{dX_m^2} = \frac{8}{\pi} \frac{2X_m^2 - 12}{X_m^4\sqrt{X_m^2 - 4}},$$

i.e.:

$$2\frac{dP}{dX_m} + X_m \frac{d^2P}{dX_m^2} = \frac{8}{\pi} \frac{(-4)}{X_m^3\sqrt{X_m^2 - 4}}$$

$$P + X_m \frac{dP}{dX_m} = \frac{2}{\pi}\arcsin\frac{2}{X_m} + \frac{4}{\pi} \frac{4 - X_m^2}{\sqrt{X_m^2 - 4}}$$

From (6.187) for $K = 200$ we obtain:

$$U = -\frac{200}{1+\omega_v^2};\ V = -\frac{200}{\omega_v(1+\omega_v)};\ V^2 = -\frac{U^3}{U+200};\ U > -200$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

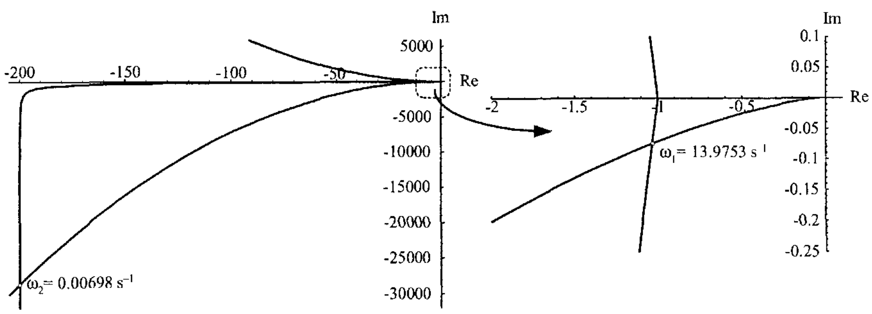
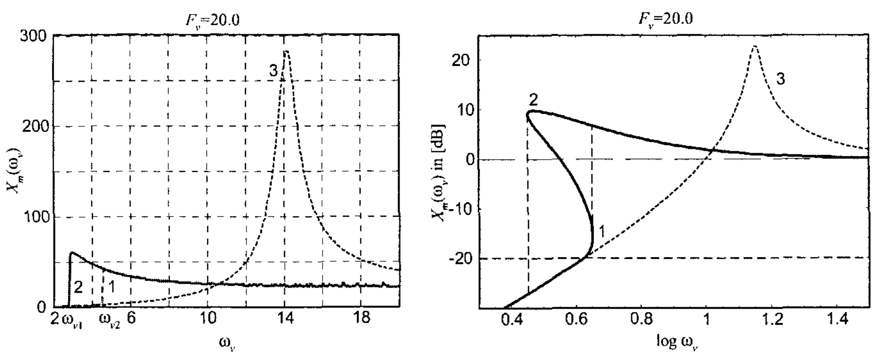


Figure 6.32: Frequency characteristic of linear part of the system $G_L(j\omega_v)$ and envelope A.



(a) (b)

Figure 6.33: The frequency characteristic $X_m(\omega_v)$: (a) obtained by simulation and (b) determined analytically. Legend: 1 – frequency increases, 2 – frequency decreases, 3 – linear system frequency response.

By substituting the obtained relations into (6.185), the parametric equation of the envelope is determined. The area of resonance jump is determined by frequency interval $(\omega_{v1}, \omega_{v2})$, *i.e. by the points of intersection of the characteristic* $G_L(j\omega_v)$ *and the envelope A, see Fig. 6.32.*

The frequency characteristic (6.190) can conveniently be determined by simulation, see Fig. 6.33a. The graphical representation of $X_m(\omega_v)$ *obtained analytically is given in Fig. 6.33b.*

Resonant frequencies in Fig. 6.33a are $\omega_{v1} = 4.55$ *and* $\omega_{v2} = 2.85$, *and resonant amplitudes are* $X_m(\omega_{v1}) = 41.481$ *and* $X_m(\omega_{v2}) = 56.318$. *In Fig. 6.33b resonant frequencies are* $\omega_{v1} = 4.475$ *and* $\omega_{v2} = 2.829$, *and resonant amplitudes are* $X_m(\omega_{v1}) = 43.4691$ *and* $X_m(\omega_{v2}) = 56.5685$. *By comparing the results of sim-*

Table 6.1: Resonant frequencies and resonant magnitudes for different values of amplitude F_v.

F_v	ω_{v1}	ω_{v2}	X_{m1a}	X_{m1b}	X_{m2a}	X_{m2b}
2.5	5.82115	9.51569	0.517129	14.5501	2.23818	7.93584
3.5	5.19282	8.64646	0.555204	18.1721	2.33936	10.0918
5	4.59681	7.71242	0.604289	22.9813	2.47718	13.2690
10	3.61617	5.97736	0.725694	36.1589	2.84856	23.5776
15	3.13488	5.06450	0.813521	47.0203	3.13843	33.6180
20	2.82857	4.47490	0.883877	56.5685	3.38131	43.4691

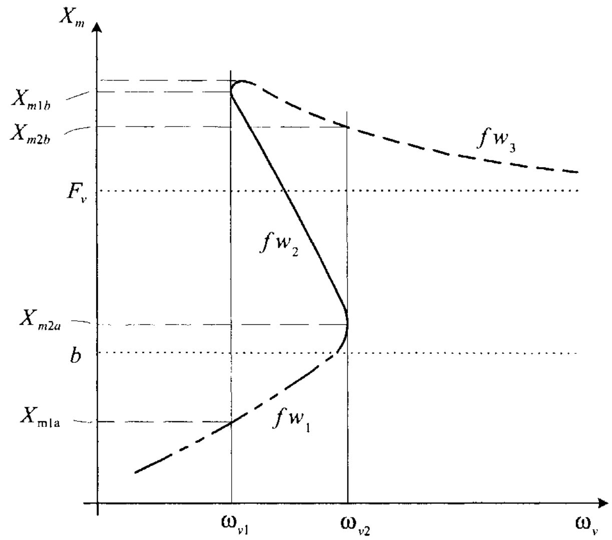


Figure 6.34: Graphical display of quantities in Table 6.1.

ulation and analytical results we can conclude that simulation results are more precise, since analytical results depend on accuracy of approximation of nonlinear element with its describing function.

The results of computation for different amplitudes F_v are shown in Table 6.1. The quantities from Table 6.1 are displayed in Fig. 6.34. As can be seen for this example, increasing the amplitude F_v lowers the resonant frequencies.

■

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

6.3 Conclusion

Harmonic linearization (describing function method) is a powerful classical tool in the hands of control engineers. Many nonlinear control systems in operation today used in one design stage or another the describing function method. Even though this is the classical method it still has attractive features for any control engineer because of the insight it gives in the frequency domain.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Chapter 7

Harmonic Linearization in Dynamic Analysis of Nonlinear Control Systems in Tracking Mode of Operation

Harmonic linearization used for dynamic analysis of nonlinear control systems in tracking mode of operation will be covered in this chapter. This mode of operation is common for electromechanical servo systems, guidance and control systems of moving objects (airplanes, ships, underwater vehicles, etc.) and in any control task where change from one to another operating point has to be made in a desired way. Contrary to control systems working in stabilization mode, in tracking tasks oscillations are not allowed and should be eliminated if possible. If this is not possible, as is often the case, then they should be used for improving the behavior of a control system. This is treated in the first part of the chapter, where vibrational linearization is used for this purpose. However, there are situations where external disturbance is a harmonic signal or can be approximated by the harmonic signal. How to analyze those systems is dealt with in the second part of the chapter.

7.1 Vibrational Linearization with Self-Oscillations

Established self-oscillations in nonlinear control systems can be used to change the static properties of the nonlinear element and thus to obtain the desired behavior of the control system. Such an application takes place in control systems when the reference signal is changing slowly in relation to the frequency of established

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

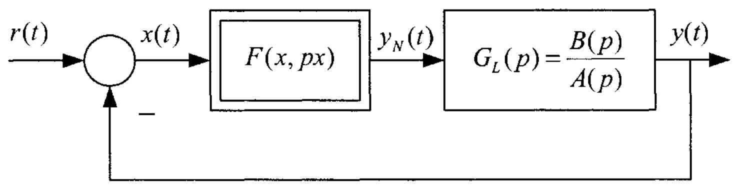


Figure 7.1: Block diagram of nonlinear control system with nearly constant reference signal $r(t)$.

self-oscillations. The procedure is applicable in situations when the tracking problem of the nonlinear closed-loop control systems has to be solved, and the established self-oscillations are to be used for this useful purpose. The same effect can be achieved with a high-frequency dither signal[1]. The basic condition is that the frequency of self-oscillations is much higher than the frequency of the reference signal. If this condition is fulfilled the necessary information can be brought to the control signal in order to move the process in the right direction and obtain the control goal.

In this chapter the dynamic behavior of the nonlinear closed-loop control system with established self-oscillations will be analyzed when subject to reference signals of certain properties. The block diagram of the control system is shown in Fig. 7.1 (Solodovnikov, 1969; Petrov et al., 1967; Csáki, 1972; Mohler, 1991).

The dynamics of the nonlinear closed-loop control system is described by nonlinear differential equation of the form (6.139):

$$A(p)x(t) + B(p)F(x, px) = A(p)r(t) \tag{7.1}$$

The reference signal has an arbitrary form, in contrast to the case in Section 6.2, when the reference signal was zero ($r(t) = 0$, a regulator problem). The frequency of the reference signal is much lower than the frequency of self-oscillations. In a first approximation, the reference signal $r(t)$ which acts upon the system with established self-oscillations can be regarded as constant or a slowly changing quantity within every period of self-oscillations. When the linear part of a control system contains ν pure integrators (type ν) then $A(p)$ will contain multiplier p^ν, i.e. $A(p) = p^\nu A_1(p)$, and the right side of the equation (7.1) will be $p^\nu A_1(p)r(t)$ or $A_1(p)p^\nu r(t)$, respectively. Either velocity ($\nu = 1$) or acceleration ($\nu = 2$) of the reference signal can be regarded as the "reference signal".

With the cited properties of the reference signal and assuming the filtering properties of the linear part of the system $G_L(p)$, the solution of the equation (7.1)

[1]Here, instead of injecting the external "high" frequency dither signal, the established self-oscillations inside the control system can be used for the same purpose.

is sought in the form (6.56). Hereby, the non-oscillatory component in the system (Fig. 7.1) is caused by a nearly constant value of $r(t)$.

Under the influence of nearly constant reference (or disturbance) signals, periodic oscillations in the system will be asymmetric to the time axis and will take the form:

$$x(t) = x^0(t) + x^* = x^0(t) + X_m \sin \omega t \tag{7.2}$$

where $x^0(t)$ is nearly constant component at the input to nonlinear element—shift of the center of self-oscillations under the action of nearly constant reference signal $r(t) \approx x^0$ and $x^* = X_m sin\omega t$ are self-oscillations which are established in the nonlinear control system with a slowly changing amplitude $X_m(t)$ and the frequency $\omega(t)$ within every period.

In analogy with the discussion in Section 6.1.4, by expanding the nonlinear function into a Fourier series, the describing function of the nonlinear element is:

$$F(x, px) \approx F^0(x^0, X_m, \omega) + P\left(x^0, X_m, \omega\right) x^* + Q(x^0, X_m, \omega) x^* \tag{7.3}$$

Inserting (7.2) and (7.3) in (7.1), the harmonically linearized equation of the system is obtained:

$$\begin{aligned} A(p)(x^0 + x^*) + B(p)\left[F^0(x^0, X_m, \omega) + P\left(x^0, X_m, \omega\right) x^* + Q(x^0, X_m, \omega) x^*\right] \\ = A(p) r(t) \end{aligned} \tag{7.4}$$

Similarly to the decomposition of (6.62), a system of equations can be written down:

$$A(p)x^0 + B(p)F^0\left(x^0, X_m, \omega\right) = A(p)r(t) \tag{7.5}$$

$$A(p)x^* + B(p)\left[P\left(x^0, X_m, \omega\right) + \frac{Q(x^0, X_m, \omega)}{\omega} p\right] x^* = 0 \tag{7.6}$$

In the above system of equations, the nonlinear dependence of the system parameters x^0, X_m, ω is preserved, i.e. the solution cannot be found by superposition of the solutions of each equation separately. By substituting in (7.6) derivative operator p with the variable $s = j\omega$, from (7.5) and (7.6) a system of three algebraic equations with three unknowns x^0, X_m and ω follows:

$$A(0)x^0 + B(0)F^0(x^0, X_m, \omega) = A(0)r(0) = M^0 \tag{7.7}$$

$$Re(x^0, X_m, \omega) = 0 \tag{7.8}$$

$$Im(x^0, X_m, \omega) = 0 \tag{7.9}$$

where $Re(x^0, X_m, \omega)$ and $Im(x_0, X_m, \omega)$ are real and imaginary part of the characteristic equation of the expression (7.6); while $M^0 = A(0)r(0) = A(0)x^0$.

By graphical or analytical means, the amplitude X_m and the frequency ω of self-oscillations can be found as functions of the non-oscillatory (nearly constant) component $x^0(t)$:

$$X_m = X_m(x^0);\ \omega = \omega(x^0) \tag{7.10}$$

By inserting (7.10) in the expression for the coefficient of harmonic linearization for the non-oscillatory component of the Fourier series $F^0(x^0,\ X_m,\ \omega)$, the smoothed static characteristic can be written:

$$F^0(x^0, X_m, \omega) = \Phi(x^0, X_m, \omega) \tag{7.11}$$

When (7.11) is put into (7.1) and the derivative operator p is substituted by complex variable s where s is fixed to $s = 0$, the equation for determining the non-oscillatory component of the shift of the center of self-oscillations $x^0(t)$ is obtained:

$$A(0)x^0 + B(0)\Phi(x^0) = M^0 \tag{7.12}$$

where $A(0)r(0) \approx M^0 \approx const.$

We can say that if the nonlinear control system has self-oscillations, then a slowly changing component of the signal at the input to the nonlinear element $x^0(t)$ represents the reference signal of the nonlinear closed-loop control system ($r(t) \approx x^0$) with established self-oscillations $x = X_m \sin \omega t$. From (7.11) can be concluded that the nonlinear characteristic $y_N = F(x, px)$ is transformed due to existing self-oscillations into a new characteristic, $\Phi(x^0, X_m, \omega)$. According to (7.11), $\Phi(x^0, X_m, \omega)$ is dependent on the reference signal $r(t) \approx x^0$ and also on the amplitude X_m and the frequency of self-oscillations ω. This can have a strong influence on the dynamic properties of the system. The nonlinear element with the nonlinear static characteristic $y_N = F(x)$ will be transformed due to established self-oscillations into a new static characteristic $\Phi(x^0, X_m)$ which is not dependent on the frequency of self-oscillations.

Thus, for example, if at the input to the nonlinear element of the type saturation (Fig. 7.2) acts a signal $x(t) = x^0 + x^*$, the static characteristic of the type saturation will be dependent on the amplitude X_m of self-oscillations. By expansion of the nonlinear function $F(x)$ in a Fourier series (see (7.3)), the following is evident:

1. The constant component[2] in the series $F^0 = \Phi(x^0, X_m, \omega)$ is not a function of frequency, as the nonlinear element is without inertia (saturation), so we have $F^0 = \Phi(x^0, X_m)$;

2. The constant component F_0 is a consequence of the existence of dc component $x^0(t)$ at the input of the nonlinear element;

[2] A slowly changing component.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

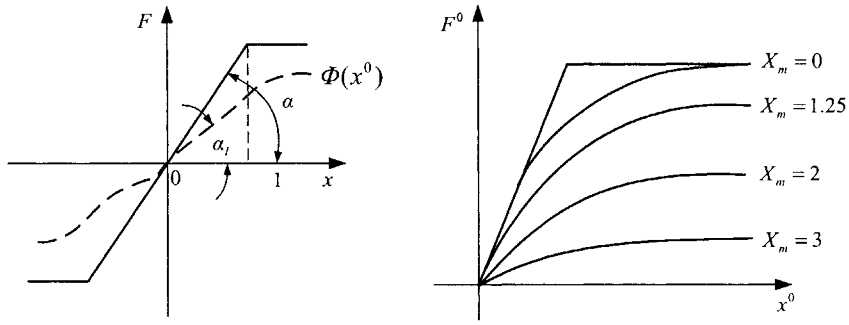


Figure 7.2: (a) Nonlinear static characteristic and (b) smoothed static characteristic of saturation.

3. Due to the self-oscillations $x^* = X_m \sin \omega t$ at the input of the nonlinear element, the constant component F^0 will depend upon the amplitude of self-oscillations X_m;

4. The constant component can be written as:

$$F^0 = \frac{1}{2\pi} \int_0^{2\pi} F\left(x^0 + X_m \sin\psi, X_m \cos\psi\right) d\psi$$

where $\psi = \omega t$. With increasing amplitude of self-oscillations X_m, F^0 will deform in the direction of decreasing gain in relation to the gain of the linear part (see static characteristic of saturation in Fig. 7.2a). It holds $\tan\alpha_1 < \tan\alpha$ (where α is the slope of the straight line of nonlinear static characteristic of saturation in the linear part around the origin).

The smoothing effect of the nonlinear static characteristic is achieved because the input signal to the nonlinear element contains also the harmonic component x^* (self-oscillations). This is called vibrational smoothing (or lubrication) of the nonlinearity. Vibrational smoothing here is the consequence of self-oscillations[3].

The function $\Phi(x^0, X_m)$ is called the smoothed static characteristic of the nonlinear element (Fig. 7.2b). It is obvious that the existence of an oscillatory signal is essential to achieve this effect. Whether vibrational smoothing of a nonlinear static characteristic of nonlinear element is achieved by injecting a dither signal or by already established self-oscillations is not important. It is essential to get the oscillations at the input to the nonlinear element, since only then such effect

[3]In this case, the role of a dither signal now has the input signal x^* at the input to the nonlinear element, which oscillates with frequency (ω) of self-oscillations, so the same effect as with a dither signal is obtained, namely dynamic lubrication (see Section 4.1.3).

can be expected. A dc component of a signal at the input to a nonlinear element cannot cause this effect. For nonlinear control systems, it is essential that the dc component which is present in the reference signal of the closed-loop control system is transferred to the control signal in order to lead the process in the right direction, i.e. to improve the accuracy of tracking. The next example illustrates that this effect can be achieved with vibrational smoothing.

EXAMPLE 7.1

As an example a nonlinear element with the static characteristic of the type dead zone (Fig. 7.3a) or backlash (Fig. 7.3b) is shown.

If at the input to the nonlinear element the signal is without an oscillatory component x^* *(self-oscillations), this signal* $x(t) = x^0$ *will not propagate through the nonlinear element if* $x^0 < b$. *Namely, the input is under the sensitivity threshold, and* $F(x, px) = 0$. *The control signal will be zero, and the process will not progress in the right direction.*

With existing self-oscillations $x(t) = x^0 + x^*$, *the non-oscillatory component of the input signal* $x^0 < b$ *propagates through the nonlinear element in the form of a constant component* F^0. *In Fig. 7.3c the smoothed static characteristic* $\Phi(x^0)$ *is shown. The smoothed characteristic does not contain an insensitivity zone for the component of the input signal* $x^0(t)$.

In order that the nonlinear characteristics $F(x, px)$ *are smoothed, the self-oscillations with sufficiently high frequency* ω *and with the amplitude* $X_m > b - x^0$ *or* $X_m > b + x^0$ *must be established. An additional requirement is that the linear part of the system* $G_L(j\omega)$ *practically suppresses the oscillatory component* $x^* = X_m sin \omega t$.

The effect of vibrational smoothing is positive in a large number of problems since it causes the annulment of dead zone. Nevertheless, there are other examples when it has a negative influence on the dynamic properties of the system, as in the case of nonlinearities of the type saturation (Fig. 7.2). Here is $F^0 > \Phi(x^0)$, and the component x^0 passes through the nonlinear element with smaller gain.

The possibility to simplify the determination of parameters x^0, X_m and ω of the input signal to the nonlinear element originates from two basic properties of the function $\Phi(x^0)$:

- Smoothed static characteristic depends upon the properties of the nonlinear element, the structure and parameters of the control system, and
- It is independent of the form, number and place of action of nearly constant external signals.

Independently of the form of the nonlinear static characteristic $F(x)$ or $F(x, px)$, the smoothed static characteristic $\Phi(x^0)$ becomes continuous, and—contrary to a nonlinearity $F(x, px)$—can be linearized with any linearization method, for example Fig. 7.4.

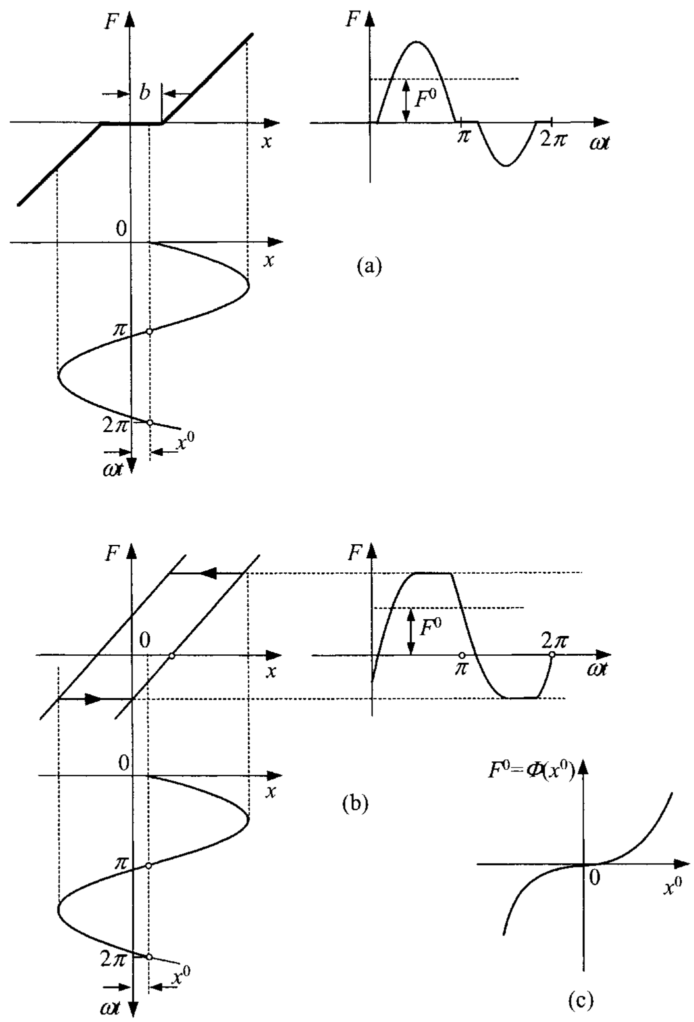


Figure 7.3: Effects which harmonic inputs produce on nonlinear element of type dead zone (a), of type backlash (b) and smoothed static characteristic of nonlinearity (c).

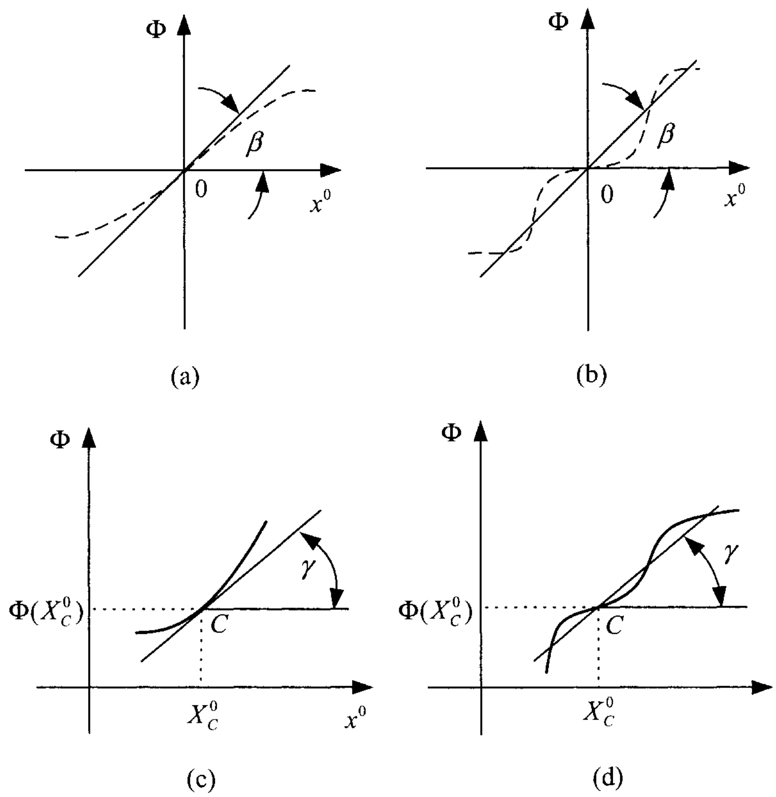


Figure 7.4: Linearization of various smoothed static characteristics.

Linearization of the static characteristic of the nonlinear element with the purpose to obtain a vibrationally smooth static nonlinear characteristic $\Phi(x^0, X_m)$ has the name *vibrational linearization*.

By linearizing the function $\Phi(x^0)$ (Fig. 7.4a and 7.4b) in a definite region, we obtain:

$$F^0 = K_N \cdot x^0 \tag{7.13}$$

where:

$$K_N = (\frac{d\Phi}{dx^0})_{x^0=0} = \tan\beta \tag{7.14}$$

By analogy, the linearization (Fig. 7.4c and 7.4d) gives:

$$F^0 = F_C^0 + K_N(x^0 - x_c^0) \tag{7.15}$$

where:

$$F_C^0 = \Phi(x_c^0)$$

$$K_N = (\frac{d\Phi}{dx^0})_{x^0=x_c^0} = \tan\gamma$$

For the smoothed static characteristics in Fig. 7.4b and Fig. 7.4d, K_N was obtained using the secant method of linearization (see Section 3.1). Since the function $\Phi(x^0)$ is determined by function $F^0(x^0, X_m, \omega)$ for $X_m = X_m(x^0)$ and $\omega = \omega(x^0)$, equation (7.14) can be written in the form:

$$K_N = \left(\frac{d\Phi}{dx^0}\right)_{x^0=0} = \left(\frac{\partial F^0}{\partial x^0} + \frac{\partial F^0}{\partial X_m}\frac{dX_m}{dx^0} + \frac{\partial F^0}{\partial \omega}\frac{d\omega}{dx^0}\right) \tag{7.16}$$

For odd symmetrical nonlinear characteristics $y_N = F(x)$, the function $F^0 = F^0(x^0, X_m)$ is not dependent on the frequency of self-oscillations ω. Therefore, the coefficient $K_N(X_m)$ can be found from the function F^0 without determining $\Phi(x^0)$:

$$F^0(x^0, X_m) = \frac{1}{2\pi}\int_0^{2\pi} F(x^0 + X_m \sin\psi)d\psi \tag{7.17}$$

where $\psi = \omega t$. Also,

$$\left(\frac{\partial F^0}{\partial X_m}\right) = \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{\partial F}{\partial x}\right)_{x=X_m\sin\psi} \sin\psi d\psi = 0 \tag{7.18}$$

where $\partial F/\partial x$ is an even function. Combining (7.18) and (7.16) gives:

$$K_N = \left(\frac{\partial F^0}{\partial x^0}\right)_{x^0=0} \tag{7.19}$$

In all other cases the coefficient K_N can be found by applying the expression (7.16). By inserting $F^0(x^0, X_m, \omega) = \Phi(x^0)$ in (7.16), the general nonlinear differential equation of the closed-loop control system with established self-oscillations is written down:

$$A(p)x^0 + B(p)\Phi(x^0) = A(p) \cdot r(t) \tag{7.20}$$

By linearizing the smoothed static characteristic $\Phi(x^0)$ by means of the coefficient $K_N(X_m)$, from the equations (7.19) and (7.16) follows:

$$F^0\left[x^0, X_m(x^0), \omega\left(x^0\right)\right] = \Phi(x^0) = K_N \cdot x^0 \tag{7.21}$$

If (7.21) is inserted in (7.20) the linearized differential equation of the nonlinear closed-loop control system is obtained:

$$[A(p) + B(p)K_N]x^0 = A\left(p\right) \cdot r(t) \tag{7.22}$$

Based on (7.22), the nearly constant component of the input to nonlinear element $x^0(t)$ in nonlinear control systems with self-oscillations can be found by linear methods. The synthesis of the nonlinear control system by changing of parameters or structure requires one to determine the associated X_m and $K_N(X_m)$. By applying the equations (7.5) and (7.6), it is possible to work out the engineering methods for determining the control algorithms for nonlinear control systems with self-oscillations.

The described properties of nonlinear control systems where self-oscillations occur have the following advantages and drawbacks:

1. In the case when self-oscillations in the system are allowed, they can be used for smoothing the nonlinear characteristic, i.e. for eliminating jumps, hysteresis, insensitivities, etc.

2. In the case when the self-oscillations in the system are unavoidable but are harmful to the quality of the control process, the basic drawback is the dependence of the control signal gain $K_N(X_m)$ on the amplitude X_m. This can greatly deteriorate accuracy, stability and other dynamic properties of the nonlinear control system during tracking mode of operation.

The magnitude of external actions (reference or disturbance signals) which bring about the shift of the center of self-oscillations $x^0(t)$ can also establish conditions for self-oscillations and influence the amplitude of self-oscillations X_m, i.e. $X_m = X_m(x^0)$. The coefficients $P(x^0, X_m, \omega)$ and $Q(x^0, X_m, \omega)$ depend on the shift value $x^0(t)$ in (7.5). Namely, a constant reference signal $r(t) = const.$ or some other constant disturbance can change the conditions for establishing self-oscillations as well as the stability region of the system. As the amplitude of self-oscillations at the input to the nonlinear element X_m and the shift of the center of these self-oscillations x^0 influence all dynamic properties of the system $K_N(X_m), P(x^0, X_m, \omega)$, $Q(x^0, X_m, \omega)$, the presented general approach must be worked out in detail for the specific methods of analysis and synthesis of the considered class of systems.

It is also necessary to emphasize that the setting of K_N by partial differentiation of the function F^0 with respect to x^0 is not always allowed. Namely, the function F^0 depends on three variables x^0, X_m and ω. In their part, the variables X_m and ω depend on the shift of the center of self-oscillations, i.e. $X_m = X_m(x^0)$ and $\omega = \omega(x^0)$. However, the gain K_N of linearized smoothed static characteristic obtained by self-oscillations at the input to the nonlinear element is found by expression (7.16) when the nonlinear static characteristic is not odd symmetrical.

As an example, for smoothed static characteristic $\Phi(x^0)$ in Fig. 7.4c and 7.4d, the linear differential equation (7.22) is:

$$A(p)\Delta x^0 + B(p)K_N\Delta x^0 = A(p)\Delta r(t) \tag{7.23}$$

where:

$$\Delta x^0 = x^0 - x_c^0; \;\; \Delta r(t) = r(t) - r_c^0 \tag{7.24}$$

From (7.23) and (7.24), the determining of $x_c^0$and r_c^0 results in:

$$A(0)x_C^0 + B(0)K_N x_C^0 = A(0)r_C^0 \tag{7.25}$$

where $K_N x_C^0 = F_C^0 = \Phi(x_C^0)$.

Table A.3 in Appendix A includes expressions for gains K_N for some typical nonlinear elements.

The phenomenon of vibrational linearization (smoothing of nonlinearities), or linearization of the smoothed static characteristic $\Phi(x^0)$, to a large degree facilitates the dynamic analysis of nonlinear control systems with complex structures. For example, the simplified block diagram of airplane course control in Fig. 7.5 comprises the nonlinear servo system for the vertical rudder located in the airplane's tail. This servo is realized with a relay, power amplifier, the driving motor and additional feedback loop.

In the system in Fig. 7.5, the nonlinear rudder servo subsystem for controlling the vertical[4] rudder at the airplane tail can be analyzed by vibrational linearization, i.e. by the method of smoothing the relay characteristic by self-oscillations.

A sufficiently high frequency of self-oscillations can be achieved by the corresponding choice of a system's parameters or by introducing a compensator. The intention is that the amplitude of self-oscillations of signal $u_1(t)$ is sufficiently small to prevent the oscillations of the rudder as well as the oscillations of the airplane around its center of gravity.

[4]A similar control system is also used for ailerons (horizontal rudders in airplane's tail, to control the pitch of the aircraft).

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

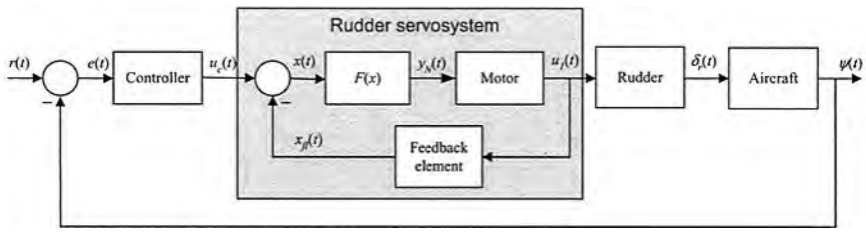


Figure 7.5: Block diagram of an airplane course control system.

The system can be designed in steps. First, self-oscillations are determined for the rudder servo system (internal feedback loop) with the input u_c and output u_1, where u_c is a slowly changing (nonoscillatory with regard to the frequency of self-oscillations) reference signal. Next, for the rudder servo system, the smoothed static characteristic $\Phi(x^0)$, frequency $\omega = \omega(u_c)$ and amplitude $X_m = X_m(u_c)$ of self-oscillations are determined, but it must remain outside of possible oscillations of the airplane. After that the nonlinear differential equation (7.20) is determined for the rudder servo system. The linearization of smoothed static characteristic $\Phi(u^c)$ by a straight line $F^0 = K_N u_c$ results in a linear differential equation of the rudder servo system (7.22). The differential equation of the closed-loop control system as a whole includes the vibrationally linearized differential equation of the rudder servo system.

For the simplified block diagram of the course-keeping control system for the aircraft in Fig. 7.5 the vibrations of the airplane around its center of gravity can be viewed as a slowly changing action of the external signal on the rudder servo system in which much higher frequency self-oscillations than the frequency of vibrations of the airplane around the center of gravity are established. Thus, the two processes can be well analyzed as conditionally separate processes. In this example it is possible after determining $\Phi(u_c)$ of the rudder servo system to analyze the servo system with a new nonlinear element $P(u_c)$ and to repeat the dynamic analysis for the whole system by making the use of harmonic linearization.

7.2 Dynamic Analysis of Nonlinear Control Systems in Tracking Mode of Operation with Forced Oscillations

A specific type of behavior in nonlinear control systems is evident in situations when the system is subject to the action of reference (set-point) and external periodic excitation signals (forcing harmonic signals). As is already said, control

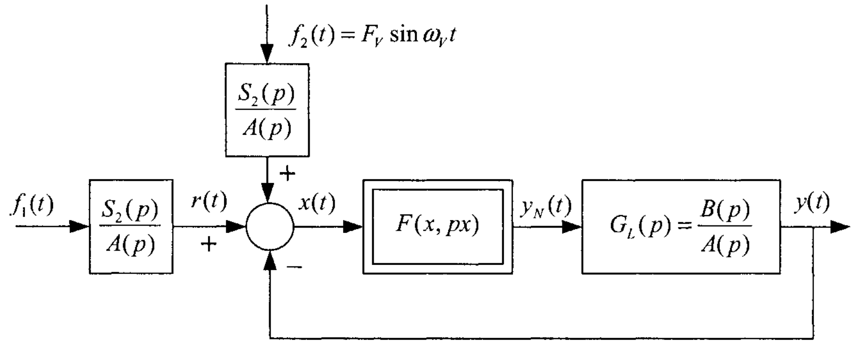


Figure 7.6: Block diagram of a nonlinear closed-loop control system in tracking mode of operation.

systems in tracking mode of operation should not oscillate. However, for nonlinear control systems this is often their natural behavior and self-oscillations are common. What remains at our disposal to counteract that is to force the control system toward oscillations of desired frequency (forced oscillations). Analyzing nonlinear control systems in tracking mode of operation, which is at the same time forced to oscillate with tolerable frequency of forced oscillations, will be the prime topic in the following text. The block diagram of a closed-loop nonlinear control system in tracking mode of operation with applied forcing harmonic signal $f_2(t)$ is given in Fig. 7.6.

The dynamic behavior of a closed-loop nonlinear control system in tracking mode of operation with established forced oscillations is described by the following differential equation (Bessekerskii and Popov, 1975; Solodovnikov, 1969; Petrov et al., 1967):

$$A(p)x(t) + B(p)F(x, px) = S_1(p)f_1(t) + S_2(p)f_2(t) \tag{7.26}$$

where:

$f_1(t)$ is the slowly varying reference (or disturbance) signal with respect to the frequency of the forced oscillations ω_v,

$f_2(t)$ is the external periodic excitation signal (forcing harmonic signal) of the form:

$$f_2(t) = F_v \sin \omega_v t \tag{7.27}$$

$S_1(p), S_2(p)$ are the signal conditioning polynomials.

The control process in which there exists the relation between established forced oscillations w_v and a nearly constant[5] reference signal do not undergo the superposition principle, i.e. both signals $f_1(t)$ and $f_2(t)$ are nonlinearly coupled and influence the dynamics of the nonlinear control system.

Solving equation (7.26) is simpler than solving equation (7.1), since the frequency ω_v is known. Provided that equation (7.26) satisfies the filter hypothesis, the solution is sought in the form:

$$x_v = x_v^0 + x_v^*;\ x^* = X_m \sin(\omega_v t + \varphi) \tag{7.28}$$

where $x^0(t)$ is a slowly varying component (signal) at the input of the nonlinear part of the closed-loop control system $F(x, px)$ and X_m, ϕ are amplitude and phase of forced oscillations; they are slowly varying quantities when the control signal changes.

By solving equation (7.26) in the form of (7.28), the unknowns $x^0(t)$, $X_m(t)$ and $\phi(t)$ are found.

Harmonic linearization of nonlinearity $F(x, px)$ is carried out by expressions (6.162) to (6.165). By means of equation (6.142), the external periodic signal $f_2(t)$ is expressed by the periodic component of the solution $x^*(t)$. By inserting the obtained expressions for $F(x, px)$ and $f_2(t)$ and by decomposing, the system of equations is written down:

$$A(p)x^0 + B(p)F^0 = S_1(p)f_1(t) \tag{7.29}$$

$$\left[A(p) - S_2(p)\frac{F_v}{X_m}\left(\cos\varphi - \frac{\sin\varphi}{\omega_v}p\right) + B(p)\left(P + \frac{Q}{\omega_v}p\right)\right]x^* = 0 \tag{7.30}$$

where:

$$F^0 = F^0(x^0, X_m, \omega_v);\ P = P(x^0, X_m, \omega_v);\ Q = Q(x^0, X_m, \omega_v) \tag{7.31}$$

Again the operator p is replaced by the variable $s = j\omega_v$ in the characteristic equation of (7.30). By rearranging (7.30) the following equation for determining the parameters of forced oscillations is obtained:

$$X_m \frac{A(j\omega_v) + B(j\omega_v)\,[P + jQ]}{S_2(j\omega_v)} = F_v e^{-j\varphi} \tag{7.32}$$

From this equation, either by the graphical or analytical procedures presented in Chapter 6, the parameters of the established forced oscillations (vibrations) are determined in the form:

$$X_m = X_m(x^0);\ \varphi = \varphi(x^0) \tag{7.33}$$

[5]With respect to the frequency ω_v of the forced oscillations.

where x^0 is the unknown parameter.

Amplitude $X_m(x^0)$ from (7.33) substituted into (7.29) gives:

$$A(p)x^0 + B(p)\Phi(x^0) = S_1(p)f_1(t) \tag{7.34}$$

where:

$$F^0 = F^0(x^0, X_m(x^0), \omega_v) = \Phi(x^0)$$

is the smoothed static characteristic of the nonlinear element. The smoothed static characteristic of nonlinear part of the system $\Phi(x^0)$ is a new nonlinear characteristic for the nonlinear control system which sustains forced oscillations (vibrations) with the frequency ω_v.

Since any nonlinearity (dead zone, backlash, hysteresis, etc.) is smoothed by established forced oscillations, the new nonlinear characteristic $F(x^0)$ as a smooth function can be linearized by the tangent or secant linearization methods (see Section 3.1):

$$\Phi(x^0) = K_N x^0; \; K_N = \left(\frac{d\Phi}{dx^0}\right)_{x^0=0} \tag{7.35}$$

Similarly to the presentation in Section 7.1 for uneven symmetrical nonlinearities $y_N = F(x)$, a simpler linearization method is applied:

$$K_N = \left(\frac{\partial F^0}{\partial x^0}\right)_{x^0=0} \tag{7.36}$$

i.e. in the above-mentioned cases the signal gain at the input to nonlinear element K_N can be found directly by partial differentiation of the first coefficient of the Fourier series $F^0 = (x^0, X_m)$:

$$K_N = \left(\frac{\partial F^0}{\partial x^0}\right)_{x^0=0} = K_N(X_m)$$

without determining smoothed static characteristic $\Phi(x^0)$, i.e. it is sufficient to find the amplitude of forced oscillations for $x^0 = 0$.

The general procedure to find smoothed static characteristic $\Phi(x^0)$ and $X_m(x^0)$, i.e. the dependence of the amplitude of forced oscillations (vibrations) at the input to the nonlinear element upon a slowly varying signal at the input of the nonlinear part of the closed-loop control system, is of a great practical interest and characterizes only nonlinear control systems.

Linearization of the smoothed static characteristic $\Phi(x^0)$, i.e. determining $K_N(X_m)$ from (7.35) or (7.36), the nonlinear differential equation of a closed-loop control system with forced oscillations present becomes a linear differential equation:

$$[A(p) + B(p)K_N]x^0 = S_1(p)f_1(t) \tag{7.37}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Analysis and synthesis of the control system described by (7.37) is realized by applying the linear theory of automatic control. In the analysis, and especially synthesis, of the control system, it must be taken into account that the gain $K_N(X_m)$ depends generally upon the structure and parameters of the system. Namely, after the change of the structure or some parameter of the system, the quantities X_m and K_N must again be calculated, i.e. X_m and K_N are essential for the stability and the dynamic and static properties of the control process. There exist a lot of systems where vibrations negatively influence their dynamical properties. As an example ships or airplanes always have inherent vibrations. However, such inevitable vibrations can be sometimes useful in stabilizing the control system, i.e. if they are treated as external periodic functions. The described procedure is generally applied to analyze the influence of vibrational or random disturbances on the quality of control task in nonlinear control systems.

The characteristic equation of the linearized control system (7.37) is:

$$A(p) + B(p)K_N = 0 \tag{7.38}$$

where $K_N = K_N(X_m, F_V, \omega_V)$ is the equivalent gain of nonlinear element with slowly varying signal $x^0(t)$ at its input. The equivalent gain K_N characterizes significant nonlinear properties of the control system. $K_N(X_m)$ is a function of the amplitude X_m of established forced oscillations; of slowly varying input function $f_1(t)$; of the control system's parameters; and of parameters of external vibrations F_v and ω_v. It is evident that the equivalent gain K_N has a great impact on every dynamic property of the control system. The "linear" characteristic equation (7.38) determines the stability of linearized control system while solution of (7.37) determines static and dynamic accuracy of the linearized control system.

A basic requirement in determining the coefficient K_N is to find the amplitude of forced oscillations X_m for $x^0 = 0$ at the input of nonlinear element. This amplitude can be put into an expression for K_N, which for typical nonlinear elements is given in Table A.3 in the Appendix A.

The problem of finding the equation (7.29) is substantially simplified in the following cases:

1. Linear part of the system (7.37) described by the transfer function:

$$G_L(p) = \frac{B(p)}{A(p)}$$

   prevents the vibrations with the frequency ω_v to pass through it, but they pass through $S_2(p)/A(p)$.

2. Vibrations with the frequency ω_v are damped by one part of the linear system (e.g. controlled object), while some other part (e.g. inner feedback loop) allows them to pass. In other words, the linear part as a whole doesn't sufficiently damp outside vibrations.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

3. High-frequency vibrations are forced directly to the nonlinear element and are filtered out by the linear part of the system (vibrational smoothing of nonlinearity).

Case 1. The equation for finding the amplitude of the forced oscillations results from (7.32):

$$X_m = \frac{\sqrt{R_{S_2}^2(\omega_v) + I_{S_2}^2(\omega_v)}}{\sqrt{R_A^2(\omega_v) + I_A^2(\omega_v t)}} F_v \tag{7.39}$$

where R_{S_2}, R_A, I_{S_2}, I_A are real and imaginary parts of $S_2(j\omega_v)$ and $A(j\omega_v)$, respectively. The value of X_m is determined from (7.39); it depends upon the system parameters, the form of nonlinear static characteristic and the parameters of forced oscillations, and is not dependent upon x^0. Coefficient K_N is sought in the tables, in such a way that the calculated value of X_m is inserted into expression for K_N of a given nonlinear element.

Case 2. In accordance with equations (7.29) and (7.30), the dynamics of nonlinear system (7.26) can be described by equations:

$$A(p)x^0 + B(p)F^0 = S_1(p)f_1(t) \tag{7.40}$$

$$A(p)x^* + B(p)\left(P + \frac{Q}{\omega_v}p\right)x^* = S_2(p)f_2(t) \tag{7.41}$$

The block of that part of the control system which damps the forced oscillations is excluded from (7.41), so the equation for the oscillatory component of the system has the form:

$$A_1(p)x^* + B_1(p)\left(P + \frac{Q}{\omega_v}p\right)x^* = S_{2_1}(p)f_2(t) \tag{7.42}$$

From (7.42) and (7.32) results:

$$X_m = \frac{\sqrt{R_1^2(X_m, \omega_v) + I_1^2(X_m, \omega_v)}}{\sqrt{R_{S_{2_1}}A^2(\omega_v) + I_{S_{2_1}}^2(\omega_v t)}} F_v \tag{7.43}$$

Coefficient $K_N(X_m)$ is found in the same way as in the case 1.

Case 3. An external vibrational function (forced harmonic signal) acts at the input to the nonlinear element:

$$\begin{aligned} f_2(t) &= F_v \sin \omega_v t; \\ X_m &= F_v;\ \varphi = 0, \end{aligned} \tag{7.44}$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

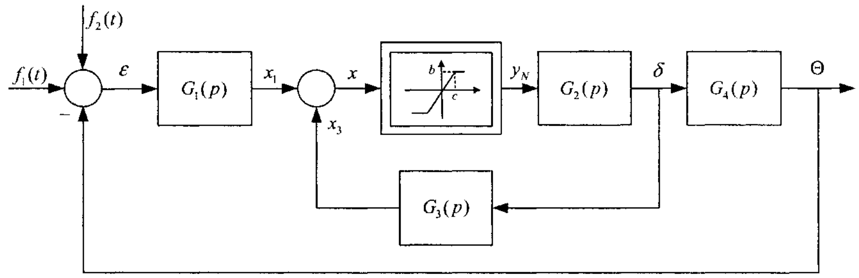


Figure 7.7: Block diagram of the aircraft pitch control system.

For the specific nonlinearity, K_N is determined by setting $K_N(X_m)$, $X_m = F_v$ into the expression for $K_N(X_m)$.

EXAMPLE 7.2 (AIRCRAFT PITCH CONTROL SYSTEM[6])

For the control of pitch of an airplane, the block diagram of the system "autopilot-elevators[7]-airplane" in Fig. 7.7. It is necessary to determine the amplitude of outside vibrations—the disturbances that act on the pitch of the airplane. These disturbances can be elastic vibrations of the airplane registered by the gyroscope together with the useful control signal, electrical fluctuations in the control circuits, vibrations of mechanical components of the control system, etc.

The elements of the block diagram are:

$f_1(t)$ – *reference signal (pitch),*

$f_2(t) = F_v sin w_v t$ – *external vibrational disturbance,*

$G_1(p) = \frac{x_1(p)}{e(p)} = K_1 + K_2 s$ – *transfer function of autopilot,*

$G_2(p) = \frac{\delta_e(p)}{y_N(p)} = \frac{1}{p(T_2 p+1)}$ – *transfer function of driving servomotor for the elevators,*

$G_3(p) = \frac{x_3(p)}{\delta_e(p)} = K_3$ – *transfer function of proportional internal feedback,*

$G_4(p) = \frac{\Theta(p)}{\delta_e(p)} = \frac{K_4(T_4 p+1)}{p^3+a_2p^2+a_1p+a_0}$ – *transfer function of the airplane on the pitch angle* Θ,

$\delta_e(p)$ – *deflection angle of the elevator.*

The operator p *in Fig. 7.7 is replaced with complex variable* $s = \sigma \pm j\omega$; *given parameters are* $\omega_v = 100[s^{-1}]$; $T_2 = 0.08[s]$; $K_1 = 0.9$; $K_2 = 0.4$.

Solution. *Differential equations of the nonlinear control system are:*

$$(p^3 + a_2 p^2 + a_1 p + a_0)\Theta = K_4(T_4 p + 1)\delta_e \tag{7.45}$$

[6] Adopted and verified from Petrov, Solodovnikov and Topcheev (1967).

[7] The change of the pitch moment of the airplane is accomplished by horizontal elevators (horizontal rudders) situated in the airplane's tail.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

$$x_1 = (K_1 + K_2 p)[f_1(t) + f_2(t) - \Theta] \tag{7.46}$$

$$(T_2 p + 1)p\delta_e = y_N = F(x);\ x = x_1 - x_3 \tag{7.47}$$

$$x_3 = K_3 \delta_e \tag{7.48}$$

The described control system can be analyzed by a simplified procedure, as the linear part of the system practically damps the oscillatory component of external disturbance with frequency ω_v. *So, the high-frequency component* $Y_n \sin 100t$ *at the output from* $G_2(p)$ *component will be attenuated by a factor of 800:*

$$F_v \delta_e = |G_2(j\omega_v)| y_N = \frac{y_N}{\omega_v \sqrt{\omega_v^2 T_2^2 + 1}} \approx \frac{y_N}{100\sqrt{64}} \approx \frac{y_N}{800} \tag{7.49}$$

and has practically no influence on the input signal x of the nonlinear element.

The amplitude of vibrations at the input to the nonlinear element $X_m \approx X_{1m}$; $x_3 \approx 0$:

$$X_m = F_v \sqrt{K_1^2 + K_2^2 \omega_v^2} \tag{7.50}$$

The equation for the subsystem (7.47) in the presence of external vibrations is:

$$(T_2 p + 1)p\delta_e = K_N x^0 \tag{7.51}$$

where (Table A.3, Appendix A):

$$K_N = \frac{2K}{\pi} \sin^{-1} \frac{c}{X_m} \tag{7.52}$$

The differential equation of the system for the reference signal—slowly varying component at the input to nonlinear element is:

$$[A(p) + B(p)K_N]x^0 = S_1(p) f_1(t) \tag{7.53}$$

The characteristic equation of the system (7.52) is:

$$A_5 s^5 + A_4 s^4 + A_3 s^3 + A_2 s^2 + A_1 s + A_0 = 0 \tag{7.54}$$

The stability conditions from the Hurwitz criterion are:

$$\begin{gathered} A_5 > 0;\ A_4 > 0;\ A_3 > 0;\ A_2 > 0;\ A_1 > 0;\ A_0 > 0 \\ A_4 A_3 - A_5 A_2 > 0 \\ (A_4 A_3 - A_5 A_2)(A_2 A_1 - A_3 A_0) - (A_4 A_1 - A_5 A_0)^2 > 0 \end{gathered} \tag{7.55}$$

In the third equation in (7.55), numerical values of the system's parameters will be inserted and the condition for the stability of the system $H(K_N)$ *follows:*

$$H(K_N) = K_N^3 - 14{,}2K_N^2 + 6{,}51 K_N > 0 \tag{7.56}$$

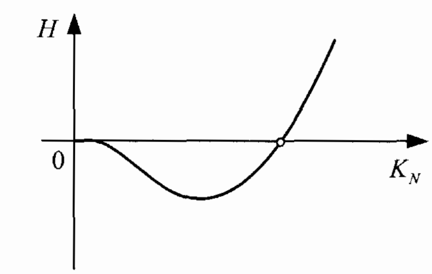


Figure 7.8: The stability boundary of a nonlinear system—Example 7.2.

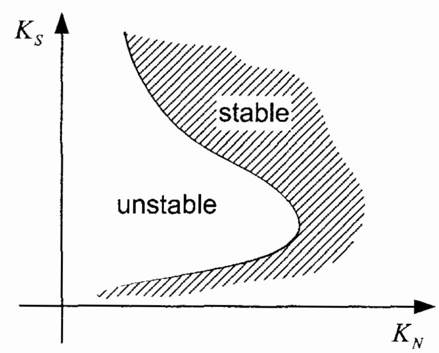


Figure 7.9: Stability region in parametric plane.

A possible graphical plot of $H(K_N)$ is given in Fig. 7.8.

All the stability conditions (7.55) follow from (7.56), and are satisfied for $K_N >$ 13.7. The maximum allowed value of forced oscillations (vibrations) at the input to the nonlinear element is obtained from (7.52):

$$X_m = \frac{b}{sin\frac{\pi K_{N\min}}{2K}} \tag{7.57}$$

for $K = tan^{-1}(c/b) = 80$; $b = 0.5$; $K_{Nmin} = 13.7$ follows that $X_m = 1.87$.

The maximum allowed value of external vibrations $f_2(t)$ is:

$$F_{V\max} = \frac{X_m}{\sqrt{K_1^2 + K_2^2\omega_v^2}} = 0,047 \tag{7.58}$$

The system stability region in the plane K_3, $K_N(X_m)$ is qualitatively shown in Fig. 7.9. ■

In general, from the characteristic equation of the linearized closed-loop control system (e.g. equation (7.54)), and with help of the Hurtwitz (or similar) stability criterion, it is possible to design the regulator. For the presented example it follows that knowing the stability conditions (7.55) and the stability region in the parameter plane (Fig. 7.9) it is possible to define K_N and K_3. Parameter K_N can be expressed as $K_N = K_N(F_v)$. Knowing F_v allows us to find the corresponding autopilot parameters K_1 and K_2.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Chapter 8

Performance Estimation of Nonlinear Control System Transient Responses

Estimating the performance indices ('damping' and 'frequency') of time-invariant nonlinear control systems, operating in stabilization mode, is not easy because performance indices are time-variant, contrary to time-invariant linear control systems. Performance diagrams of nonlinear control system transient responses, plotted in a system's parameter plane and presented in this chapter, can significantly improve our capability to estimate nonlinear control system performance and design nonlinear control systems with desired performances.

Performance of a control system can be analyzed through performance indices which are various indicators of the quality of dynamic response of the control system. The zero-input response (transient response[1] in the following text) is usually used for evaluating the dynamic performances of a control system in the stabilizing mode of operation. However, for control systems in tracking mode of operation the zero-input response is not adequate for that purpose because the transfer function zeros[2] are also very important, because they define not only dynamic behavior of control system, but also the tracking capability of the control system (accuracy of tracking the reference signal and filtering capability of disturbances). Conventional approaches to control system analysis and design use the following

[1] The term transient response will be used instead of the zero-input response. Note however that in linear system theory definitions of zero-input and transient response differ, i.e. the transient response is part of the response which decays with the time, while steady-state response is part of a response which persists indefinitely. On the contrary the zero-input response is a response to initial conditions only (no excitation or external signals are acting on a system).

[2] For linear SISO control systems.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

performance indices:

- From the time domain: percentage overshoot σ_m [%], settling time t_s [s], rise time t_r [s], error constants C_i, etc.
- From the frequency domain: bandwidth $BW\left[s^{-1}\right]$, resonant frequency $\omega_m\left[s^{-1}\right]$, natural frequency $\omega_n\left[s^{-1}\right]$, resonant peak M_m, damping factor ζ, gain and phase margin, etc.

In this chapter we are primarily concerned with the nonlinear control system in stabilizing mode of operation, so its zero-input response (transient response) is the focus of our interest. We will concentrate only on two performance indices, i.e. 'damping' and 'frequency' of the nonlinear control system experiencing damped oscillatory behavior in stabilization mode of operation.

The estimation of the characteristics of transient processes in nonlinear systems differs from that in linear systems, as the performance indicators behave differently in the two cases.

For stable linear systems with conjugate complex poles, the transient response has exponentially damped oscillations. It is said that the system behavior is underdamped. This can be written as:

$$x(t) = a_0 e^{-\zeta\omega_n t}\sin(\omega_n\sqrt{1-\zeta^2}+\varphi) = a_0 e^{\zeta_0 t}\sin(\omega_d t+\varphi)$$

where ζ is the damping ratio, ω_n is undamped natural frequency, a_0 is constant of exponential at $t = d$, ω_d is a damped natural frequency of oscillations of the transient, φ is the phase shift and $\zeta_0 = -\zeta\omega_n$.

The damping ratio ζ and the frequency ω_d are constants if the linear system is time-invariant. With nonlinear time-invariant systems, these indicators will vary with time. It will be necessary to express these indicators as part of a dynamic process, i.e. with differential equations. The procedure will be presented in Section 8.1 by means of the second-order differential equation. In Section 8.2 the performance diagrams of transient responses will be presented. They help us to assess the performance of transient response of a nonlinear control system as a function of some parameter of the system. These diagrams are significant as they enable us to to design a nonlinear control system with the requirement on the performance of transient response.

8.1 Determining Symmetrical Transient Responses Near Periodic Solutions

The characteristics of a transient response in a nonlinear control system of higher order is in most cases evaluated by knowing the transient response curve. The

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

damped oscillatory transient response in the vicinity of a periodic solution is carried out either by numerical procedures or with approximate graphical or analytical procedures. Of paramount interest for engineering practice are approximate procedures which enable a relatively simple determination of the influence of parameters on overall dynamic performance of the nonlinear control system. In these cases the characteristics of a damped oscillatory transient response is evaluated similarly as with linear systems: by overshoot, damping, settling time, the period of oscillations, etc.

Transient response of unforced nonlinear closed-loop control system (see Fig. 1.2) is defined by the solution of the nonlinear differential equation of the nonlinear control system:

$$A(p)x + B(p)F(x, px) = 0 \tag{8.1}$$

The damped oscillatory transient response $x(t)$ in the vicinity of a periodic solution — provided that the system undergoes harmonic linearization — can be determined sufficiently accurately by the analysis of dominant poles. In order to find $x(t)$ in the vicinity of a periodic solution, nonlinear differential equation (8.1) is approximated with the linear second-order differential equation[3]:

$$\frac{d^2x(t)}{dt^2} + 2h\frac{dx(t)}{dt} + c^2x(t) = \varepsilon F(x, px) \tag{8.2}$$

where ε is a small parameter, while the frequency of the periodic solution, in whose vicinity is sought the transient response, is given by:

$$\omega_0 = \sqrt{c^2 - h^2} \tag{8.3}$$

The solution for the nonlinear closed-loop control system[4] (signal at the input of nonlinear element) (8.2) is based on the fact that all higher terms in the series (because of small parameter ε) can be neglected in a first approximation (Petrov et al., 1967; Solodovnikov, 1969; Popov and Pal'tov, 1960; Pal'tov, 1975; Netushil, 1983):

$$x(t) = a(t)\sin\psi(t) + \varepsilon\varphi_1(a, \psi) + \varepsilon^2\varphi_2(a, \psi) + \dots \tag{8.4}$$

Since the amplitude $a(t)$ and the phase $\varphi(t)$ are time-varying quantities, they must be expressed as if they were dynamic processes:

$$\frac{da(t)}{dt} = -ha(t) + \varepsilon\,\Phi_1(a) + \varepsilon^2\Phi_2(a, \psi) + \dots \tag{8.5}$$

[3]The assumption is that the dominant dynamics is oscillatory, i.e. the nonlinear control system can be approximated by a linear system of second order. This assumption must always be kept in mind, since the aforesaid condition may not always be satisfied.

[4]It must be noticed that the solution is sought in the form of a harmonic signal, under the assumption that oscillations exist. If they are absent, it is normally anticipated that the solution form will be different.

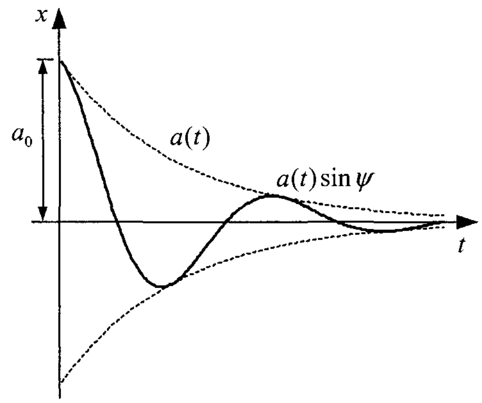


Figure 8.1: Damped oscillatory transient response of a nonlinear control system.

$$\frac{d\psi(t)}{dt} = \omega_o + \varepsilon\, b_1(a) + \varepsilon^2 b_2(a) + \ldots \tag{8.6}$$

A first approximate solution of the transient response (8.4) is:

$$x(t) = a(t)\sin\psi(t) \tag{8.7}$$

A first approximate solution of the transient response (8.5) is:

$$\frac{da}{dt} = -ha + \varepsilon\, \Phi_1(a) \tag{8.8}$$

A first approximate solution of the transient response (8.6) is:

$$\frac{d\psi}{dt} = \omega_o + \varepsilon\, b_1(a) \tag{8.9}$$

The equation (8.8) includes the finite term $-ha$ in the expression for da/dt; this allows a fast change of the amplitude $a(t)$ of nonlinear damped oscillations $x(t)$ (Fig. 8.1). On the basis of (8.7), (8.8), (8.9) and Fig. 8.1, the quickly decaying (or expanding) oscillatory transient responses can be found. For the nonlinear control system signal $x(t)$ (Fig. 1.2), transient responses are close to underdamped transient response of the linear system with the damping indication $\zeta_0 = -h$. Namely, the linear oscillatory system will produce the signal:

$$x(t) = a_0 e^{\zeta_0 t}\sin(\omega_d t + \varphi)$$

where a_0 is a constant amplitude, ζ_0 is constant damping indication, ω_d is constant frequency of oscillations of the transient response and φ is constant phase shift.

In contrast to linear control systems, these indicators are time varying in nonlinear control systems. Combining (8.3) and (8.7), (8.8), (8.9), slowly varying parameters of the transient response are at the input to the nonlinear part of the system:

$$\zeta(t) = -h + \frac{\varepsilon \Phi_1(a)}{a} \tag{8.10}$$

$$\omega(t) = \omega_0 + \varepsilon\, b_1(a) \tag{8.11}$$

where $\zeta(t)$ is 'damping' [5]—a slowly varying quantity and $\omega(t)$ is 'frequency' [6]—a fast varying quantity. From (8.7) and (8.10), the solution for the variable $x(t)$ is sought in the form:

$$x(t) = a(t) \sin \psi(t) \tag{8.12}$$

From (8.8) and (8.10) follows the change of amplitude of oscillations of nonlinear control system:

$$\frac{da}{dt} = a\zeta(t) \tag{8.13}$$

From (8.13) follows that the envelope of oscillations of nonlinear control system transient response $x(t)$ will be:

$$\frac{da}{a}\zeta(t)dt$$

and consequently

$$a(t) = a_0 e^{\int_0^t \zeta(t)dt}$$

So, this envelope consists of elementary exponential segments with continuously changing 'damping' $\zeta(t)$ (Fig. 8.1).

Equation (8.13) shows also that the oscillations of the nonlinear control system transient response will be damped[7] for $\zeta < 0$ and increasing for $\zeta > 0$:

$$\text{damped oscillations}: \frac{da}{dt} < 0, \text{ for } \zeta < 0 \tag{8.14}$$

$$\text{increasing oscillations}: \frac{da}{dt} > 0, \text{ for } \zeta > 0 \tag{8.15}$$

When $\zeta = 0$ the oscillations are not decaying nor increasing, i.e. we have self-oscillations

$$\text{increasing oscillations}: \frac{da}{dt} > 0, \text{ for } \zeta > 0 \tag{8.16}$$

[5]Here damping is quoted because it is the term borrowed from the linear theory.

[6]Frequency is also the term borrowed here from the linear theory.

[7]Note that for linear system transient response damped oscillations exist only for $0 < \zeta < 0.707$ and the response is unstable for $\zeta < 0$. This fact clearly shows that damping ratio $\zeta = \text{const.}$ for linear system is something else than the time-varying 'damping' $\zeta(t)$ for nonlinear systems, although both give us an indication of damping in the system.

From (8.9) and (8.11) follows:

$$\frac{d\psi}{dt} = \omega t \tag{8.17}$$

For the solution of the nonlinear differential equation in the vicinity of a periodic solution, the parameters 'damping' $\zeta(t)$ and 'frequency' $\omega(t)$ can be conveniently expressed by means of the amplitude $a(t)$, i.e. instead of $\zeta(t)$ and $\omega(t)$, $\zeta(a)$ and $\omega(a)$ can be used. Harmonic linearization of nonlinearity will give here specific expressions different from previous cases.

Assuming that the damping indicator is not small, from (8.12), (8.13), (8.17) and (8.1) follows that by differentiating with respect to time $x(t) = a(t) sin\psi(t)$, the product of the two functions is:

$$px(t) = a(t)\left[\omega(t)\cos\psi(t) + \zeta(t)\sin\psi(t)\right] \tag{8.18}$$

where:

$$\sin\psi(t) = \frac{x(t)}{a(t)};\ \cos\psi(t) = \frac{px(t)}{a(t)\omega(t)} - \frac{\zeta(t)x(t)}{a(t)\omega(t)} = \frac{p-\zeta(t)}{a(t)\omega(t)}x(t) \tag{8.19}$$

The first harmonic[8] of the nonlinear function $F(x, px)$, with $x(t) = a(t) sin\psi(t)$, has the form:

$$F(x, px) = \left[P(a,\zeta,\omega) + Q(a,\zeta,\omega)\frac{p-\zeta}{\omega}\right]x \tag{8.20}$$

where:

$$\begin{aligned} P &= \frac{1}{\pi a}\int_0^{2\pi} F\left[a\sin\psi, a(\omega\cos\psi + \zeta\sin\psi)\right]\sin\psi d\psi \\ Q &= \frac{1}{\pi a}\int_0^{2\pi} F\left[a\sin\psi, a(\omega\cos\psi + \zeta\sin\psi)\right]\cos\psi d\psi \end{aligned} \tag{8.21}$$

With small 'damping' $\zeta \approx 0$ (i.e. in the vicinity of periodic solution), the expression (8.26) takes on the form (3.79), i.e. the form for determining the periodic solution of the first harmonic.

With single-valued nonlinearities of the form $F(x)$, the expression for harmonic linearization is identical to the expression for a periodic solution, independent of the value of parameter ζ, providing the substitution of the constant X_m with amplitude $a(t)$ is made.

The characteristic equation of the harmonically linearized closed-loop control system (8.1) for an oscillatory transient response with large damping ζ is:

$$D(p) = A(p) + B(p)(P + Q\frac{p-\zeta}{\omega}) = 0 \tag{8.22}$$

[8] Vanishing or expanding.

where in the general case for the nonlinearity $F(x, px)$:

$$P = P(a, \zeta, \omega);\ Q = Q(a, \zeta, \omega) \tag{8.23}$$

and for nonlinearity $F(x)$:

$$P = P(a);\ Q = Q(a) \tag{8.24}$$

Determining the damped (or expanding) oscillatory transient responses in a harmonically linearized control system is carried out by determining the dominant poles $p = \zeta \pm j\omega$ of the characteristic equation (8.22):

$$\begin{aligned} D(\zeta + j\omega) &= A(\zeta + j\omega) + B(\zeta + j\omega)\left(P(a, \zeta, \omega) + Q(a, \zeta, \omega)\,\frac{\zeta + j\omega - \zeta}{\omega}\right) \\ &= 0 \end{aligned} \tag{8.25}$$

By expanding (8.22) in a series, the characteristic equation becomes:

$$\begin{aligned} D(\zeta + j\omega) = D(\zeta) + \left(\frac{dD}{dp}\right)_{\zeta} j\omega + \frac{1}{2!}\left(\frac{d^2D}{dp^2}\right)_{\zeta} (j\omega)^2 + \dots \\ + \frac{1}{n!}\left(\frac{d^nD}{dp^n}\right)_{\zeta} (j\omega)^n + \dots \end{aligned} \tag{8.26}$$

Subscript ζ of derivations of D indicates that it should be inserted for p after differentiating D with respect to p in the vicinity of periodic solution $\zeta \rightarrow 0$, since all the higher terms of ζ can be neglected.

With (8.26) rearranged:

$$D(\zeta + j\omega) = R(a, \zeta, \omega) + jI(a, \zeta, \omega) = 0 \tag{8.27}$$

By equating the real and imaginary parts of the characteristic equation $D(\zeta + j\omega)$ with respect to p in the vicinity of periodic solution $\zeta \rightarrow 0$, the system of equations is obtained:

$$\begin{aligned} D_1(\zeta + j\omega) &= R(a, \zeta, \omega) = 0 \\ D_2(\zeta + j\omega) &= I(a, \zeta, \omega) = 0 \end{aligned} \tag{8.28}$$

By solving the system of equations (8.28), the solution for 'damping' and 'frequency' as functions of the amplitude of oscillations is obtained:

$$\omega = \omega(a);\ \zeta = \zeta(a) \tag{8.29}$$

From (8.28) for $\zeta = 0$, the amplitude $a(t) = X_m$ is obtained, as well as the frequency of self-oscillations. The approximate form of the oscillatory transient response can be found from (8.29). By satisfying the initial amplitudes $a(t) = a_0$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

at $t = 0$ and by calculating $\zeta = \zeta(a_0)$, the velocity of the envelope of damping of oscillations (Fig. 8.1) at the starting point can be determined. In other words, for $a(t) = a_i$ the values of $\zeta = \zeta(a_i)$ are found, as well as the envelope of the transient response $a(t)$. Similarly, the function $\omega = \omega(a)$ can be calculated.

For the synthesis of the nonlinear control system with the desired transient response, it is necessary to know the influence of the basic system parameters (for example K and T) on the performance indicators $\zeta(a)$ and $\omega(a)$. From (8.27) the influence of the system parameters K and T on the properties of the oscillatory transient response can be determined. The system (8.28) is rewritten in the form:

$$\begin{aligned} D_1(a,\zeta,\omega,K,T) &= 0 \\ D_2(a,\zeta,\omega,K,T) &= 0 \end{aligned} \tag{8.30}$$

From the above system of equations, the dependencies (8.29) are determined for certain parameter values. By knowing the functions $\zeta = \zeta(a)$ and $\omega = \omega(a)$, by means of equation (8.12) as a starting point, the parameters (performances) of the oscillatory transient response can be determined by solving equations:

$$\int_{a_0}^{a} \frac{da}{a\zeta(a)} = t;\ \psi = \int_0^t \omega(a)\,dt + \psi_0 = \int_{a_0}^{a} \frac{\omega(a)}{a\zeta(a)} da + \psi_0 \tag{8.31}$$

The number of oscillations as well as the settling time of the transient response, during which the envelope $a(t)$ decreases from initial value a_0 to some final value a_K (when the response has decayed essentially to zero) comes from (8.31):

$$t_s = \int_{a_0}^{a_K} \frac{da}{a\zeta(a)} \tag{8.32}$$

$$m = \frac{\psi - \psi_0}{2\pi} = \frac{1}{2\pi}\int_0^{t_s} \omega(a)\,dt = \frac{1}{2\pi}\int_{a_0}^{a_K} \frac{\omega(a)}{a\zeta(a)} da \tag{8.33}$$

EXAMPLE 8.1

For the nonlinear electromechanical servo system with the structure as in Fig. 8.2, the $\zeta(a)$ and $\omega(a)$ are to be found.

Solution. *The equations for the system in Fig. 8.2 are:*

$$\begin{aligned} u_1(t) = F_1(x),\ (T_1 p + 1)u_2(t) &= K_1 u_1(t) \\ (T_2 p + 1)u_2(t) = K_2 u_2(t), u_4(t) &= F_2(u_3(t)) \\ m(t) = K_3 u_4(t), px(t) &= -K_4 m(t) \end{aligned}$$

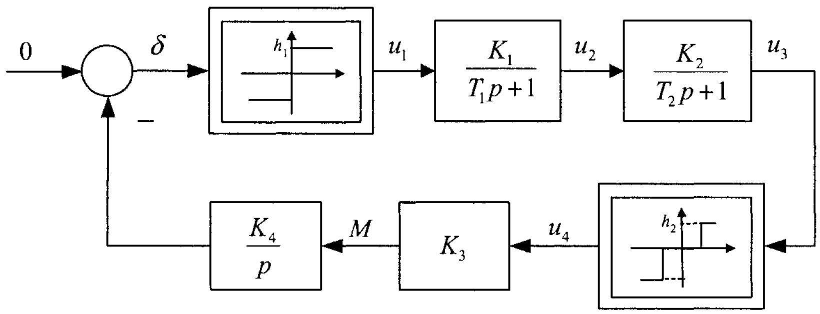


Figure 8.2: Block diagram of a nonlinear servo system in stabilization mode.

Harmonic linearization $F_1(x)$ for $x(t) = a(t)sin\psi(t)$ is:

$$P_1 = \frac{4h_1}{\pi a};\ u_1 = P_1(a)\cdot x$$

and the harmonic linearization $F_2(u_3)$ for $u_3 = a_2(b)sin(\psi+\varphi)$:

$$P_2 = \frac{4h_2}{\pi a_2}\sqrt{1-(\frac{c}{a_2})^2};\ u_4 = P_2(a_2)\cdot u_3$$

From:

$$-P_1(a)\frac{K_1}{(T_1p+1)}\cdot\frac{K_2}{(T_2p+1)}\cdot P_2(a_2)\frac{K_3K_4}{p}x = 0$$

we obtain:

$$\left(1+\frac{K_1K_2K_3K_4P_1(a)P_2(a_2)}{p(T_1p+1)(T_2p+1)}\right)x = 0 \tag{8.34}$$

and the characteristic equation of the harmonically linearized system is:

$$T_1T_2p^3+(T_1+T_2)p^2+p+KP_1(a)\cdot P_2(a_2) = 0 \tag{8.35}$$

where $K = K_1K_2K_3K_4$. The amplitudes of oscillations at the inputs of nonlinearities F_1 and F_2 are coupled by the relation:

$$\frac{x}{u_3} = \frac{4h_2}{\pi a_2}\sqrt{1-\left(\frac{c}{a_2}\right)^2}\frac{K_3K_4}{p} \tag{8.36}$$

Inserting $p = \zeta + j\omega$ in (8.36):

$$\frac{|x|}{|u_3|} = \frac{a}{a_2} = \frac{4h_2}{\pi a_2}\sqrt{1-\left(\frac{c}{a_2}\right)^2}\frac{K_3K_4}{\sqrt{\zeta^2+\omega^2}} \tag{8.37}$$

or:

$$a_2 = \frac{4h_2K_3K_4c}{\sqrt{(4h_2K_3K_4)^2 - \pi^2a^2(\zeta^2+\omega^2)}} \tag{8.38}$$

From (8.38) and (8.35) the characteristic equation of third order follows:

$$p^3 + c_1p^2 + c_2p + c_3 = 0 \tag{8.39}$$

The coefficients of the characteristic equation (8.39) are:

$$c_1 = \frac{T_1+T_2}{T_1T_2};\ c_2 = \frac{1}{T_1T_2}$$

$$c_3 = \frac{KP_1(a)P_2(a)}{T_1T_2} = \frac{K}{T_1T_2}\frac{4h_1}{\pi a}\frac{4h_2}{\pi a_2}\sqrt{1-\left(\frac{c}{a_2}\right)^2}$$

By using (8.38) we will extract a_2:

$$\frac{1}{a_2} = \frac{\sqrt{(4h_2K_3K_4)^2 - \pi^2a^2(\zeta^2+\omega^2)}}{4h_2K_3K_4c}$$

and:

$$\sqrt{1-\left(\frac{c}{a_2}\right)^2} = \sqrt{1-1+\frac{\pi^2a^2(\zeta^2+\omega^2)}{(4h_2K_3K_4)^2}} = \frac{\pi a}{4h_2K_3K_4}\sqrt{\zeta^2+\omega^2}$$

If we substitute the latter equation into the expression for c_3, we obtain:

$$c_3 = \frac{K}{T_1T_2}\frac{4h_1}{\pi a}\frac{4h_2}{\pi}\frac{\sqrt{(4h_2K_3K_4)^2 - \pi^2a^2(\zeta^2+\omega^2)}}{4h_2K_3K_4c}\frac{\pi a}{4h_2K_3K_4}\sqrt{\zeta^2+\omega^2}$$

and after collecting terms:

$$c_3 = \frac{K}{T_1T_2}\frac{4h_1(\zeta^2+\omega^2)}{\pi K_3K_4c}\sqrt{\frac{1}{\zeta^2+\omega^2} - \left(\frac{\pi a}{4h_2K_3K_4}\right)}$$

Substituting $p = \zeta + j\omega$ into (8.39), using $p^3 = \zeta^3 + 3j\zeta^2\omega - 3\zeta\omega^2 - j\omega^3$ and $p^2 = (\zeta^2-\omega^2) + j2\zeta\omega$, yields the characteristic equation of the closed-loop control system

$$\begin{aligned} D(\zeta+j\omega) &= \left[(\zeta^3-3\zeta\omega^2) + j(3\zeta^2\omega - \omega^3)\right] \\ &\quad + c_1\left[(\zeta^2-\omega^2) + j2\zeta\omega\right] + c_2(\zeta+j\omega) + c_3 = 0 \end{aligned}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

from which the equations for the real and imaginary parts follow:

$$
\begin{aligned}
D_1(\zeta + j\omega) &= R(c_i, \zeta, \omega) = \zeta^3 - 3\zeta\omega^2 + c_1(\zeta^2 - \omega^2) + c_2\zeta + c_3 = 0 \\
D_2(\zeta + j\omega) &= I(c_i, \zeta, \omega) = 3\zeta^2\omega - \omega^3 + c_1 2\zeta\omega + c_2\omega = 0
\end{aligned} \tag{8.40}
$$

If the first equation is rearranged and the second divided by ω, the following set of equations is obtained:

$$\zeta^3 + c_1\zeta^2 + c_2\zeta + c_3 = \omega^2(3\zeta + c_1) \tag{8.41}$$

$$3\zeta^2 + c_1 2\zeta + c_2 = \omega^2 \tag{8.42}$$

Substituting 8.42 into 8.41 gives:

$$\zeta^3 + c_1\zeta^2 + c_2\zeta + c_3 = 9\zeta^3 + 9c_1\zeta^2 + 2c_1^2\zeta + 3c_2\zeta + c_1c_2$$

or:

$$8\zeta^3 + 8c_1\zeta^2 + 2c_2\zeta + 2c_1^2\zeta = 2\zeta(4\zeta^2 + 4c_1\zeta + c_2 + c_1^2) = c_3 - c_1c_2$$

and further:

$$\zeta = -\frac{c_1c_2 - c_3}{2\left[(c_1 + 2\zeta)^2 + c_2\right]} \tag{8.43}$$

Now we shall transform (8.41) into the following form:

$$\zeta^3 + c_1\zeta^2 + c_2\zeta + c_3 = \omega^2(2\zeta + c_1) + \omega^2\zeta \tag{8.44}$$

Substituting the term (8.42) for the last right-hand member yields:

$$
\begin{aligned}
\zeta^3 + c_1\zeta^2 + c_2\zeta + c_3 &= \omega^2(2\zeta + c_1) + (3\zeta^2 + c_1 2\zeta + c_2)\zeta \\
\zeta^3 + c_1\zeta^2 + c_2\zeta + c_3 &= \omega^2(2\zeta + c_1) + 3\zeta^3 + c_1 2\zeta^2 + c_2\zeta
\end{aligned}
$$

or:

$$\omega^2(2\zeta + c_1) + 2\zeta^3 + c_1\zeta^2 - c_3 = \omega^2(2\zeta + c_1) + \zeta^2(2\zeta + c_1) - c_3 = 0 \tag{8.45}$$

and further:

$$\omega^2 = \frac{c_3}{c_1 + 2\zeta} - \zeta^2 \tag{8.46}$$

The dependence of indicators for damping and frequency with respect to the amplitude of oscillations is illustrated in Fig. 8.3. ∎

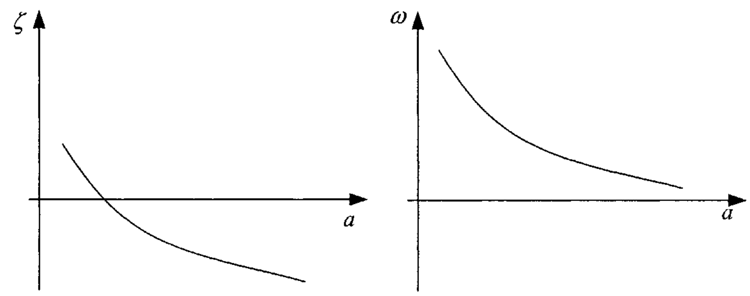


Figure 8.3: The dependence of 'damping' and 'frequency' on the amplitude of oscillations.

EXAMPLE 8.2

It is necessary to solve the Van der Pol equation by means of harmonic linearization.

Solution. *The nonlinear Van der Pol differential equation is:*

$$\frac{d^2x}{dt^2} - \mu\left(1-x^2\right)\frac{dx}{dt} + x(t) = 0$$

or:

$$\left(p^2 - \mu p + 1\right)x + \mu x^2 px = 0 \tag{8.47}$$

where μ is generally not a small parameter. The nonlinearity in (8.47) has the form:

$$F(x, px) = \mu x^2 px = 0$$

If we substitute $\mu x^2 = \mu a^2 \sin^2\psi$ and $px = a(\omega\cos\psi + \zeta\sin\psi)$ we obtain:

$$F(x, px) = \mu a^3 \sin^2\psi(\omega\cos\psi + \zeta\sin\psi) \tag{8.48}$$

The coefficients of harmonic linearization are determined by applying the expression (8.21):

$$\begin{aligned} P &= \frac{1}{\pi a}\int_0^{2\pi} \mu a^2 \sin^2\psi \cdot a(\omega\cos\psi + \zeta\sin\psi)\sin\psi d\psi = \frac{3\mu}{4}a^2\zeta \\ Q &= \frac{1}{\pi a}\int_0^{2\pi} \mu a^2 \sin^2\psi \cdot a(\omega\cos\psi + \zeta\sin\psi)\cos\psi d\psi = \frac{\mu}{4}a^2\omega \end{aligned} \tag{8.49}$$

Combining (8.49) and (8.20) gives:

$$F(x, px) = \mu x^2 px = \left(\frac{3\mu}{4}a^2\zeta - \frac{\zeta}{\omega}\frac{\mu}{4}a^2\omega\right)x + \frac{\mu}{4}a^2 px \tag{8.50}$$

From (8.50) and (8.47) the harmonically linearized differential equation results:

$$\left[p^2-\mu\left(1-\frac{a^2}{4}\right)p+1+\frac{\mu}{2}a^2\zeta\right]x=0 \tag{8.51}$$

with the roots of the characteristic equation:

$$\lambda_{1,2}=\zeta\pm j\omega=\frac{\mu}{2}\left(1-\frac{a^2}{4}\right)\pm j\sqrt{1+\frac{\mu}{2}a^2\zeta-\frac{\mu^2}{4}\left(1-\frac{a^2}{4}\right)^2} \tag{8.52}$$

where:

$$\zeta=\frac{\mu}{2}\left(1-\frac{a^2}{4}\right) \tag{8.53}$$

$$\omega^2=1+\frac{\mu}{2}a^2\zeta-\frac{\mu^2}{4}\left(1-\frac{a^2}{4}\right)^2 \tag{8.54}$$

For $\zeta=0$, the periodic solution (self-oscillations) is obtained from (8.53) and (8.54):

$$a=X_m=2$$

and from (8.54), since $\zeta=0$ when $a=2$:

$$\omega^2=1,\ \text{i.e. } \omega=1$$

Further, (8.53) and (8.32) yield the envelope of the transient response:

$$\begin{aligned} t&=\int_{a_0}^{a}\frac{da}{a\zeta(a)}=\int_{a_0}^{a}\frac{da}{a\frac{\mu}{2}\left(1-\frac{a^2}{4}\right)}=\frac{8}{\mu}\int_{a_0}^{a}\frac{da}{a\left(4-a^2\right)}\\ &=\frac{1}{\mu}\left(ln\frac{a^2}{|4-a^2|}-\ln\frac{a_0^2}{|4-a_0^2|}\right);\ \text{i.e.}\\ a(t)&=\frac{2a_0}{\sqrt{a_0^2+\left(4-a_0^2\right)e^{-\mu t}}}\end{aligned} \tag{8.55}$$

The derived solution by the method of harmonic linearization corresponds to the solution of the Van der Pol equation, for $\mu=\varepsilon$ as a small parameter. ■

8.2 Performance Diagrams of Nonlinear System Transient Responses

The indicators of the performance of an oscillatory transient response can be conveniently found by means of the *performance indices diagram* of E. P. Popov,

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

(Popov and Pal'tov, 1960). These diagrams for nonlinear symmetrical transient responses show the relations between the indicators of performance 'frequency' ω and 'damping' ζ upon the amplitude of the oscillations $a(t)$, plus a varying system parameter (e.g. gain K). They can be found from complex equation (8.27), which is transformed into the form:

$$D(a,\zeta,\omega,K) = R(a,\zeta,\omega,K) + jI(j\omega,a,\zeta,K) = 0 \tag{8.56}$$

or:

$$\begin{aligned} D_1(a,\zeta,\omega,K) &= I(a,\zeta,\omega,K)0 \\ D_2(a,\zeta,\omega,K) &= I(a,\zeta,\omega,K)0 \end{aligned} \tag{8.57}$$

By solving the system of equations (8.57) with respect to the indicators of performance ω and ζ, the following expressions are obtained:

$$\zeta = \zeta(K,a);\ \omega = \omega(K,a) \tag{8.58}$$

Setting $\zeta = const.$ and $\omega = const.$, by means of (8.58) the functions $\zeta_i(a,K) = \zeta_i = const.$ and $\omega_i(a,K) = \omega_1 = const.$ are plotted in the parameter plane K (Fig. 8.4).

By appropriate choice of initial conditions (Fig. 8.1) for $t = 0$, $x(0) = a_0$, $\psi = \pi/2$, the essential properties of the transient response can be found for an

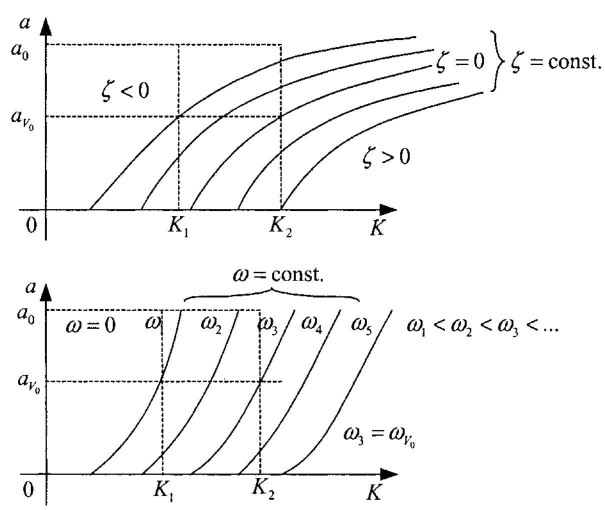


Figure 8.4: Diagrams of performance indicators for transient response of nonlinear control system.

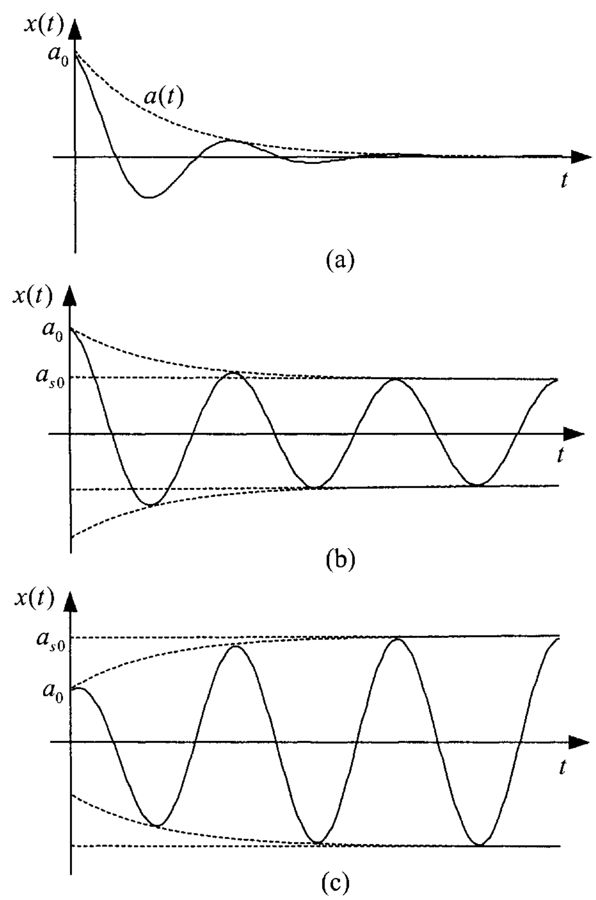


Figure 8.5: Transient responses of a nonlinear control system: (a) damped oscillatory, (b) damped oscillatory ending in self-oscillations, (c) increasing oscillatory ending in self-oscillations

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

arbitrary value of varying parameter K. For example, with $K = K_1$ (Fig. 8.4) and any value of $x(0) = a_0$, the transient response will be damped, i.e. the amplitude $a(t)$ will vary between a_0 (point A) and 0, the frequency of oscillations ω will be $\omega_1 < \omega < \omega_2$, and the damping indicator will be in the region $\zeta < 0$. If the value of $K = K_2$ and any value of $x(0) = a_0$ are chosen, $\zeta = 0$, a periodic transient response is established in the system—with self-oscillations of amplitude $a_{s0} = X_m$ and frequency $\omega_{s0} = \omega_3$.

If $a_0 > a_{s0}$ (point B), the oscillation amplitude will decrease until $a = a_{s0}$ is reached (point D), see Fig. 8.5b. If $a < a_{s0}$ the oscillation amplitude will increase until $a = a_{s0}$ is reached (point D), see Fig. 8.5c. In both cases self-oscillations are established. The region of self-oscillations is for $K > K_m$. In the region $\zeta > 0$, the oscillatory (for $K > K_m$) transient response will increase until reaching the $\zeta = 0$ curve. In the region left of the curve $\omega = 0$, the transient response will be aperiodic and decreasing. By determining the performance indices diagram of a nonlinear oscillatory transient response, a broader picture of system stability with the desired performance indices is established.

EXAMPLE 8.3

Determine the performance indices diagram of the relay servo system from Example 5.4. Control system parameters are $T_1 = 0.5[s]$; $T_2 = 0.3[s]$; $c = 6[V]$; $b = 0.1$ and for the gains of the linear part $K = 0.2$ and $K = 0.5$.

Solution. *The transfer function of the closed-loop system is:*

$$G_{cl}(s) = \frac{\Theta(s)}{\Theta_n(s)} = \frac{G_N(a) \cdot G_L(s)}{1 + G_N(a)G_L(s)} = \frac{KG_N(a)}{(T_1 s + 1)(T_2 s + 1)s + KG_N(a)}$$

where:

$$G_N(a) = \frac{4c}{\pi a}\sqrt{1 - \frac{b^2}{a^2}}$$

The characteristic equation of the system is:

$$D(a,s) = (T_1 s + 1)(T_2 s + 1)s + KG_N(a) = 0 \tag{8.59}$$

or:

$$0.03s^3 + 0.4s^2 + s + KG_N(a) = 0 \tag{8.60}$$

By putting $s = \zeta + j\omega$ in (8.60) and rearranging, the system of equations (8.57) results in:

$$\begin{aligned} R(a,\zeta,\omega,K) &= D_1(a,\zeta,\omega,K) \\ &= 0.03\zeta^3 - 0.09\zeta\omega^2 + 0.4\zeta^2 - 0.4\omega^2 + \zeta + KG_N(a) = 0 \\ I(a,\zeta,\omega,K) &= D_2(a,\zeta,\omega,K) \\ &= 0.09\zeta^2\omega - 0.03\omega^3 + 0.8\zeta\omega + \omega = 0 \end{aligned} \tag{8.61}$$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

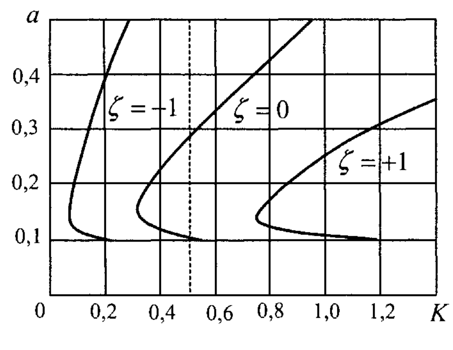


Figure 8.6: Diagram of performance indicators of a servo system from Example 8.3.

The 'frequency' $\omega = \omega(\zeta)$ is found from the second equation (8.61):

$$\omega^2 = 3\zeta^2 + 26.7\zeta + 33.3 \tag{8.62}$$

From (8.62) and the first equation of system (8.61) it follows that $\zeta = \zeta(K)$:

$$0.24\zeta^3 + 3.2\zeta^2 + 12.7\zeta + 13.3 - KG_N(a) = 0, \tag{8.63}$$

where:

$$G_N(a) = \frac{7.64}{a}\sqrt{1 - \frac{0.01}{a^2}} \tag{8.64}$$

From (8.64) it follows that the transient response is possible for $a > 0.1$. The performance indices diagram can be appropriately drawn by determining parameter K from (8.64):

$$K = \frac{1}{G_N(a)}\left(0.24\zeta^3 + 3.2\zeta^2 + 12.7\zeta + 13.3\right) \tag{8.65}$$

By inserting $\zeta =$ const. in (8.65) and by varying a within the boundaries 0.1 and 0.5, the performance indices diagram ensues (Fig. 8.6). The curves of constant values of frequency in the plane (a, K) coincide with the curves for $\zeta =$ const. According to (8.62), the values of parameter ω for $\zeta = 0$, $\zeta = -1$ and $\zeta = +1$ are $\omega = 3.1[s^{-1}]$, $\omega = 5.77[s^{-1}]$ and $\omega = 7.94[s^{-1}]$, respectively.

From the diagram in Fig. 8.6 follows that for $K = 0.2$ the transient response is damped until $a \leq 0.1$ with arbitrary initial conditions, i.e. when the difference signal becomes smaller that the dead zone. For $K = 0.5$ (dashed line), with any initial conditions, if $a > 0.105$ self-oscillations with the amplitude $a_{s0} = X_m = 0.27$ and the frequency $\omega_{s0} = 5.77s^{-1}$ are established. If the initial deflection is within *the limits* $0.1 < a < 0.105$*, the envelope of the transient response will decrease until the asymptote* $a \leq 0.1$ *is reached. From the standpoint of system stability, two limit cycles are possible: a stable limit cycle with* $X_m = 0.27$ *and an unstable limit cycle with the amplitude* $a_{s0} = 0.105$. ■

EXAMPLE 8.4

Find the performance diagram of the nonlinear control servo system transient response. The block diagram of the servo system is given in Fig. 8.7.

Solution. *The characteristic equation of the system in Fig. 8.7 is:*

$$D(a,s) = T_1 T_2 s^3 + (T_1 + T_2) s^2 + \left(1 + T_1 K_2 K_3 \frac{4h}{\pi a}\right) s + (K_1 + K_3) K_2 \frac{4h}{\pi a} = 0 \tag{8.66}$$

By inserting $s = \zeta + j\omega$ *and rearranging, the following system of equations is obtained:*

$$\begin{aligned} R(a,\zeta,\omega,K) &= D_1(a,\zeta,\omega,K) \\ &= (K_1 + K_3) K_2 \frac{4h}{\pi a} + \left(1 + T_1 K_2 K_3 \frac{4h}{\pi a}\right) \zeta + (T_1 + T_2)\zeta^2 \\ &\quad + T_1 T_2 \zeta^3 - (T_1 + T_2 - 3T_1 T_2 \zeta)\omega^2 = 0 \\ I(a,\zeta,\omega,K) &= D_2(a,\zeta,\omega,K) \\ &= f\left(1 + T_1 K_2 K_3 \frac{4h}{\pi a} + 2(T_1 + T_2)\zeta + 3T_1 T_2 \zeta^2\right)\omega + T_1 T_2 \omega^3 = 0 \end{aligned} \tag{8.67}$$

From the last equation in (8.67) follows:

$$\omega^2 = \frac{1}{T_1 T_2} + \frac{4hK_2K_3}{\pi T_2 a} + 2\frac{T_1 + T_2}{T_1 T_2}\zeta + 3\zeta^2 \tag{8.68}$$

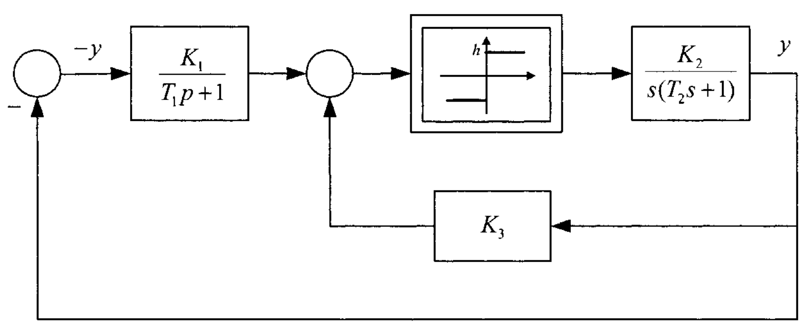


Figure 8.7: Block diagram of a nonlinear servo system.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

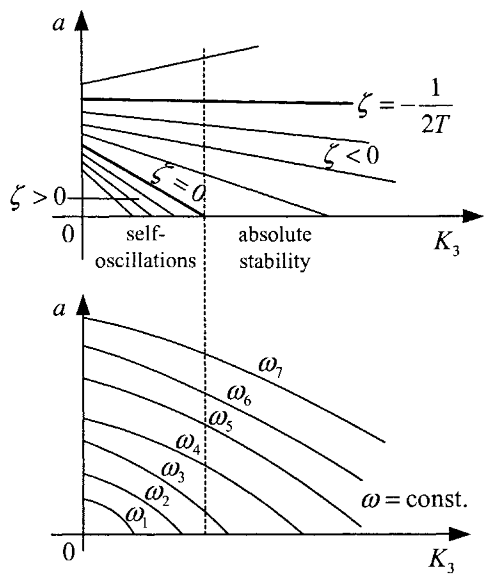


Figure 8.8: Diagram of performance indicators for transient response of a nonlinear servo system—Example 8.4.

The first equation in (8.67) and (8.68) make it possible to find $a(K_3)$:

$$a = \frac{4hK_2}{\pi f(\zeta)}\left[K_1 - T_1\left(\frac{1}{T_2} + 2\zeta\right)K_3\right] \tag{8.69}$$

where:

$$f(\zeta) = \frac{T_1 + T_2}{T_1 T_2} + 2\left[1 + \frac{(T_1 + T_2)^2}{T_1 T_2}\right]\zeta + 8(T_1 + T_2)\zeta^2 + 8T_1 T_2 \zeta^3 \tag{8.70}$$

From (8.69) and (8.70), the lines for $\zeta = const.$ in the plane (a, K_3) are drawn for the amplitudes from equation (8.69). By means of (8.68), (8.69) and (8.70), the curves for $\omega = const.$ are drawn in Fig. 8.8). From Fig. 8.8) it is seen that above the line $\zeta = 0$ are lines with $\zeta < 0$, and below are lines with $\zeta > 0$. This means that for $K_3 < K_m$ the transient response will be oscillatory with damped oscillations. ∎

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Chapter 9

Describing Function Method in Fuzzy Control Systems

Fuzzy control is an extremely successful application of fuzzy sets and systems theory to practical problems. The reason for that can be sought in theoretical and practical success stories of this approach. A good engineering approach to control of various plants is an effective use of all the available information. Often this information comes not only from the plant sensors, but also from the mathematical model of the plant and human expert knowledge about the plant operation. Experts provide linguistic descriptions about the system and control instructions. Conventional controllers cannot incorporate the linguistic fuzzy information into their design. A majority of them need an appropriate mathematical model of the system. Contrary to them, fuzzy control is a model-free approach which heavily relies on the expert linguistic fuzzy information. Fuzzy control provides nonlinear controllers which are general enough to perform any nonlinear control actions. In this chapter the basics of fuzzy logic and fuzzy set theory upon which the fuzzy control is formed are briefly shown. Our goal was not to be exhaustive but informative enough for the reader to be able to understand the basics of this prospective field. Due to the fact that this book is dealing mainly with the frequency methods for nonlinear systems, we presented in this chapter application of the describing function method in fuzzy control systems. It is shown that it is possible to obtain the describing function of the particular type of fuzzy element and consequently use stability analysis. Moreover, it is shown that a fuzzy element can diminish or even eliminate the resonance jump which can exist in a nonlinear system and aggravate effects during operation.

9.1 Basics of Fuzzy Logic

9.1.1 Introduction

Classical set theory is the child of 19th century mathematics. It was introduced by the German mathematician G. Cantor. Today we know these sets as *simple* or *sharp sets* or as *crisp sets*. For each set element it is possible to write down a single-valued membership to the terms: yes/no, high/low, dark/bright, right/wrong or to the Boolean logic: 1/0. This means that the membership to the set is strictly defined, the particular object or term is either the member of the set or it is not.

The next important step which may be characterized as a promoting one was made by a mathematician Lotfi A. Zadeh, Professor at Berkeley (1965). He introduced the idea of *fuzzy*, *soft* membership to the set used for quality description of the object or term, e.g. very large, large, medium large, small, very small. The suggested procedures are both founded and defined mathematically. The theory of fuzzy sets enables the introduction of rough description using linguistic expression. Each object (term) which is an element of the set was attributed a membership function by Zadeh, called μ, which says to what degree the object (term) is a member of the set. The standard membership function shape assumes the value of $\mu = 0$ (the object is not the member of the set) and $\mu = 1$ (the object belongs to the set completely). For sets defined in that way, Zadeh (Zadeh, 1965) introduced the expression *fuzzy*. It brought about the following expressions, i.e. fuzzy set, fuzzy logic, fuzzy arithmetic, fuzzy control, fuzzy expert systems, fuzzy computing, fuzzy model, etc. In this way, 0/1 set membership increased the use of qualitative (formally linguistic) human perception.

The following expressions have been frequently used:

1. Accuracy value: completely correct, very correct, almost wrong, wrong, ...
2. Predicates: larger than, obsolete, soon, almost, ...
3. Probabilities: near zero, unlikely, ...
4. Possibilities: almost possible, often possible, ...

The mathematical term "fuzzy set" is used to describe "soft" or "fuzzy". Here, membership function μ becomes very important. Membership of an element $x \in X$ to the set A may be described with classic (crisp) sets with characteristic function $\mu(x)$, where $x \in X$ as an element belongs (yes) or does not belong (no) to the set A. With fuzzy sets membership of the element $x \in X$ to the set A is not absolute. It is described by means of the membership function $\mu(x)$ which can now assume all values between 0 and 1. In this way it is possible to say that classic (crisp) sets represent a special example of fuzzy sets, or their subset. Simply put, element x can (in %) belong to bigger number of sets $X_1, X_2, \dots, X_n$. This being the case,

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

certain sets may exclude each other and the percentage sum of particular parts may exceed 100%.

This fuzzy set theory became very popular 25 years later. The fuzzy property of determining real objects and systems with an absence of absolutely defined borders is widely applicable, particularly with complex objects and systems where classical methods cannot be used or when their application is very demanding.

Fuzzy control is particularly important for automatic control systems because it plays an important role in the area of fuzzy systems.

The original definition of fuzzy sets according to L.A. Zadeh (1965) is:

> Fuzzy set A defined in the set X is characterized by the membership function (characteristic function) $\mu(x)$ which attributes to each element x of the set X a real number in the interval $[0, 1]$, with the value $\mu(x)$, where $\mu(x)$ represents the degree of the membership of the element x to the set A.

The basic problem of fuzzy set theory practical realization is how to determine or choose an adequate membership function. In his article, N. Nakajama (1988) says:

> Unfortunately, as a matter of fact, there are neither theoretical nor empirical methods for determining function membership.

As opposed to probability which is objective (can be proved experimentally), membership degree is subjective, although the intervals, when they both appear, overlap.

Given principles enable the use of fuzzy set theory in mathematical procedures of not completely known systems. These are systems for which it is difficult to determine a good enough simple mathematical model (make comparison with the principle of compatibility). Usually such systems are processed by means of probability theory. Fuzzy set theory with the probability theory opens new possibilities here.

Whether we are strongly critical or enthusiastic, it is obvious that the fuzzy concept offers something new. It offers wider possibilities of empirical and verbal description regarding control and regulation strategy within automated industrialized systems, particularly complex ones.

9.1.2 Fuzzy Sets Fundamentals

Fuzzy sets represent the fundamentals of fuzzy theory. In control technique they are used mostly for mathematical process description with language rules of the following forms:

IF *premise*, **THEN** *conclusion*

in which language variables are used, e.g.:

very short, **short**, **tall**, **very tall**, **very very tall**
or similar *language variables* known as *linguistic variables*.

Language or linguistic variables may be defined as a variable whose statement values are in natural or artificial language. If "tall", "very tall", "very very tall" represent the value of "height", then "height" is a language variable. To numerical values or data the *membership functions*—language variables—are attached. Mathematical notation of language variables is *fuzzy* or *soft* set. Operators *IF, THEN* which connect language variables are the rules corresponding to the simplest forms of human decisionmaking processes, because *premise* and *conclusion* are fuzzy statements, significant for human thinking. That's the way a human thinks:

IF the temperature in the room is low,
THEN put on the central heating.

In human thinking further logical connecting of such statements is present, e.g.:

IF the temperature in the room is low,
AND outside temperature is decreasing,
THEN put on the central heating to the maximum.

Descriptive forms, which enable mathematical formulation of such expressions, represent *fuzzy theory* or *fuzzy (soft) set theory*.

9.1.3 Crisp and Fuzzy Sets and Their Membership Functions

In the introduction, set A is said to be a set of specific various objects from the material and rational world which make a whole. This object membership set is a part of a *basic set* X. Single objects are *elements* x of set A. If for example set A consists of six discrete values $x_1, x_2, x_3, x_4, x_5, x_6$:

$$A = \{x_1, x_2, x_3, x_4, x_5, x_6\} \tag{9.1}$$

of real numbers between 10.0 and 20.0 it is a part of real numbers basic set $\Re$:

$$A = \{x, x \in \Re, 10.0 \leq x \leq 20.0\} \tag{9.2}$$

For *crisp* sets binary form *membership function* μ is used:

$$\mu_A(x) = \begin{cases} 1, \text{ if } x \in A \\ 0, \text{ if } x \notin A \end{cases} \tag{9.3}$$

In crisp sets each element is assigned single-valued set membership ($\mu_A = 1$) or set non-membership ($\mu_A = 0$). As an example, Fig. 9.1 shows membership function $\mu_A(x)$ for temperature value x of crisp set A which represents language value "frost".

If for the same example set A is defined so that its temperature values x denote language value "cold", the single-valued membership or non-membership of all

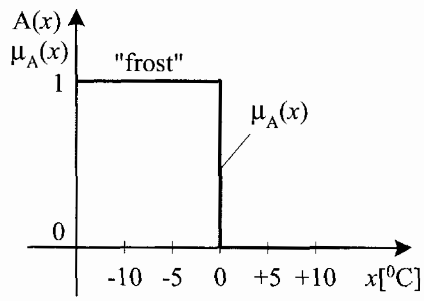


Figure 9.1: Crisp set A membership function μ_A.

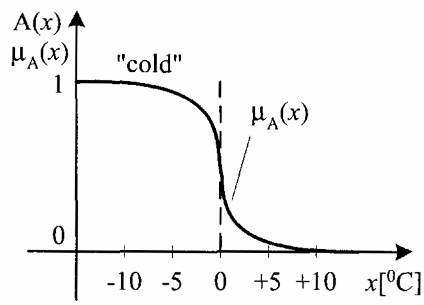


Figure 9.2: Fuzzy set A membership function.

elements x in set A is not crisp anymore, because the expression "cold" is a fuzzy-soft term and the temperature x may be higher or lower than $0[°C]$. Fuzzy area is described by means of *continuous membership function* $\mu_A(x)$. Continuous membership function for the above example is presented in Fig. 9.2.

By using the continuous membership function it is possible to show that the element x can belong partly to a certain set, or that it can be a part of several sets, respectively. At the same time, by expressing fuzzy membership, sets maintain their border value property. For example, the temperature of a room has the value of $x = 22[°C]$. If we wish, by language values it is possible to say that $x = 22[°C]$ belongs to the A set area which should mark values "cold", but at the same time it belongs to the B set area which marks "warm". Hence, language value description

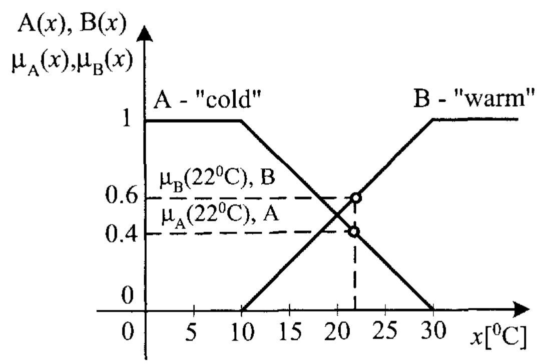


Figure 9.3: Temperature membership function of sets A and B.

is implemented in two fuzzy sets:

$$COLD : \mu_A A(x) \in [0,1], \text{ and}$$
$$WARM : \mu_B B(x) \in [0,1].$$

If membership functions $\mu_A(x)$ and $\mu_B(x)$ are chosen to be linearly dependent, which happens very often in practice, temperature membership $x = 22[°C]$ in both sets can be presented by the Fig. 9.3.

The temperature $x = 22[°C]$ has the following membership:

$$\text{to the set } A : \mu_A A(x) = 0.4$$
$$\text{to the set } B : \mu_B B(x) = 0.6$$

Figure 9.3 shows that the characteristic binary value $\mu(x) = 0/1$ for crisp sets becomes a continuous fuzzy set membership function $\mu(x)$. The continuous membership function $\mu(x)$ also gives the degree of set element membership to the fuzzy set. Therefore, in literature beside the term "membership function", the term *membership degree* is often found.

Fuzzy set A is completely described by membership function μ_A. For each element x belonging to the basic set X, the membership function $\mu_A(x)$ gives the membership degree of the element x to the set A. In that way, fuzzy set A is described by means of sorted pairs:

$$A = \{x, \mu_A(x)\}, x \in X \tag{9.4}$$

Membership function also gives the membership degree of elements x to the set A, which means that it copies set X to the membership area A. Membership function area is the subset consisting of real numbers with upper and lower borders which equals zero. If for the membership function upper border value 1 is chosen, *normalized membership function* and *normalized fuzzy set* are obtained. In control practice, normalized membership functions and standard fuzzy sets are most often used. Hence, from now on normalized set forms will be dealt with in the first place.

9.1.4 Fuzzy Set Parameter Presentation

In order to show fuzzy set parameter presentation, graphic shape of normalized membership function is used. Trapezoid, triangle and rectangle are used as basic shapes in control technique. Figure 9.4a shows a trapezoid with parameters $[x_1, x_2, x_3, x_4]$. In the case $[x_1 = x_2 = x_3 = x_4]$ we get Fig. 9.4b which is known as *singletons sets*. It is used to show crisp sets (like a type of Dirac's impulse). For $[x_1 = x_2, x_3 = x_4]$ a rectangle is obtained (Fig. 9.4c) which is rarely used in control technique. For $[x_1, x_2 = x_3, x_4]$, a triangle is obtained (Fig. 9.4d). This shape as well as trapezoid is the basic shape in control technique.

For $[x_1, x_2, x_3 > x_4]$ and $[x_1 < x_2, x_3, x_4]$ we get graphic presentations (Figs. 9.4e and 9.4f) which are often used with output values. Figure 9.5 shows common membership function in the shape of trapezoid and triangle. This being the case, trapezoid is usually given by parameters $[x_1, x_2, x_3, x_4, a, b, y_{min}, y_{max}]$, and triangle by $[x_0, a, b, y_{min}, y_{max}]$.

Triangle presentation (singleton) that is obtained for $a = 0$ or $b = 0$ is often used.

For presenting fuzzy sets in other areas different, often nonlinear, forms of function dependence are used, such as:

$$\mu(x) = e^{-x}$$

$$\mu(x) = \frac{1}{1 + x^2}$$

Instead of membership function $\mu(x)$ graphic representation, numerical entry, which can have two forms, is also used. The first form, which is often used, is expressed as follows:

$$A = \frac{\mu_1(x)}{x_1} + \frac{\mu_2(x)}{x_2} + \cdots + \frac{\mu_i(x)}{x_i} + \cdots \tag{9.5}$$

and explains that the fuzzy set A consists of elements x_i having membership values $\mu_i(x)$.

The second form gives a more detailed entry, but it is rarely found in literature. That form is shown by the following expression:

$$A = \{(x_1, \mu_1(x)), (x_2, \mu_2(x)), \ldots, (x_i, \mu_i(x)), \ldots\} \tag{9.6}$$

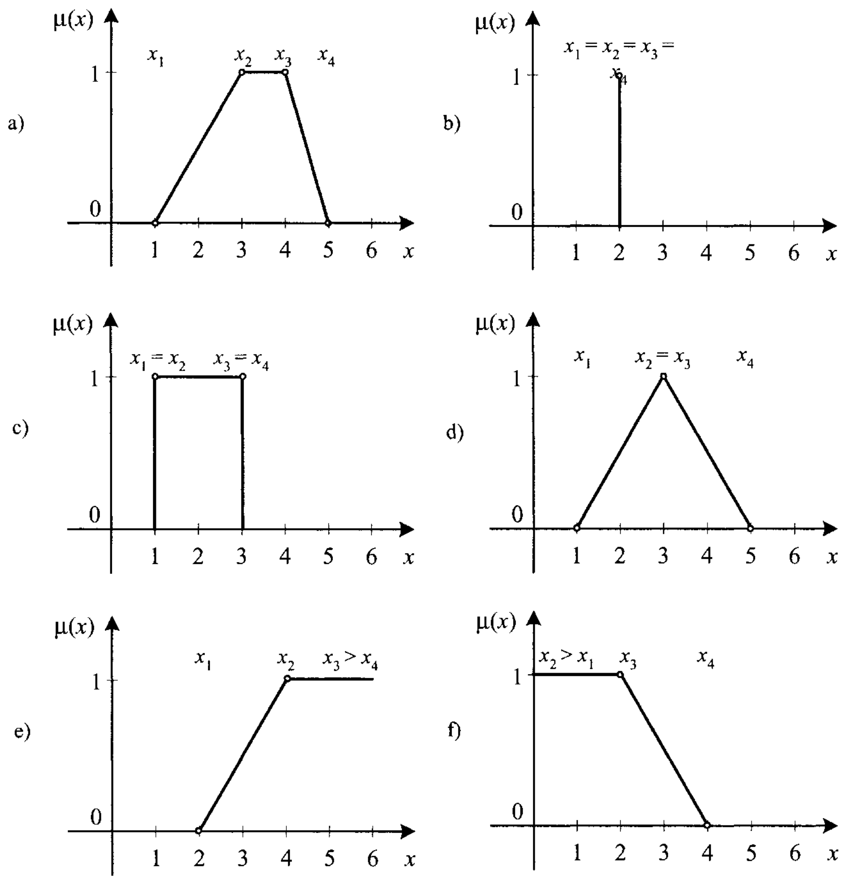


Figure 9.4: Membership functions in the shape of trapezoid and its derivations.

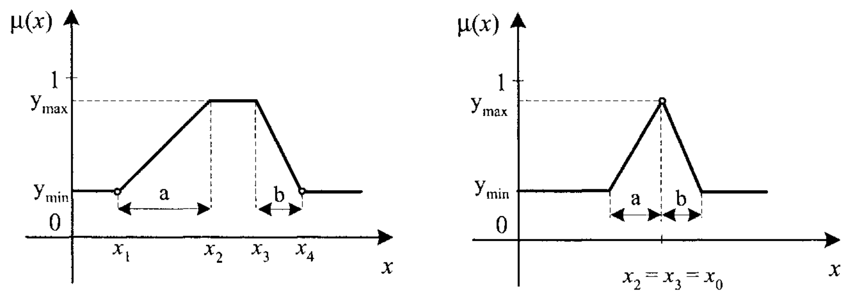


Figure 9.5: Common membership function in the shape of trapezoid and triangle.

Fuzzy set A "cold" (Fig. 9.3) can be written mathematically:

$$A = \{x, \mu_A(x)\} = \{(0,100), (5,1), (10,1), (15,0.75), (20,0.5), (25,0.25), (30,0)\}$$

9.1.5 Basic Operation on Fuzzy Sets in Control Systems

For crisp sets whose elements x come from the same basic set X, three operations have been defined:

1. *Disjunction* : OR
2. *Conjuction, intersection* : AND
3. *Complement* : NOT

Their presentation by characteristic functions leads to well-known *binary logic.* Before these basic operations (operators) are extended to be used with fuzzy sets, it is necessary to study their meaning with crisp sets. As an example, all three operations on sets A and B are studied (Fig. 9.6).

1. Disjunction (OR) $C = A \cup B$ is given to those elements x which belong to set A, set B or to sets A and B. Disjunction characteristic function C is given by maximum operator:

$$C = A \cup B = \max\{A, B\} \tag{9.7}$$

2. Conjunction (AND) $C = A \cap B$ is given to those elements x which belong to both A and B sets. Disjunction characteristic function C is given by minimum operator:

$$C = A \cap B = \min\{A, B\} \tag{9.8}$$

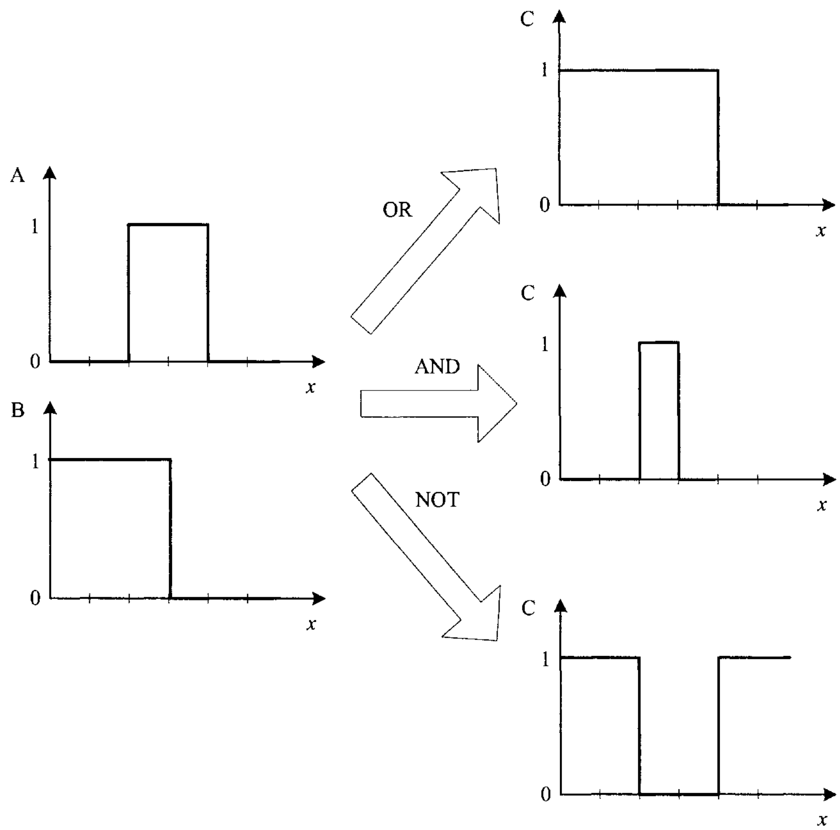


Figure 9.6: Basic procedures on crisp sets.

3. Negation (NOT) $C = \neg A = \overline{A}$ is given by those elements x which do not belong to set A. Therefore, complement characteristic function is given by negation operator:
$$C = \neg A = \{1 - A\} \tag{9.9}$$

Since crisp sets are a special example of fuzzy sets, the described basic procedures may be transferred directly onto fuzzy sets. In this way, the following three operators can be obtained:

a) for disjunction:
$$\mu_c C(x) = \max\{\mu_A A(x), \mu_B B(x)\} \tag{9.10}$$

b) for conjunction:
$$\mu_c C(x) = \min\{\mu_A A(x), \mu_B B(x)\} \tag{9.11}$$

c) for complement:
$$\mu_c C(x) = \{1 - \mu_A A(x)\} \tag{9.12}$$

All three fuzzy operators (a, b and c) for fuzzy sets $\mu_A A(x)$—trapezoid—and $\mu_B B(x)$—cut trapezoid—are shown in Fig. 9.7.

Figure 9.8 shows the meaning of fuzzy operators (a, b and c), triangle shape for sets $A(x)$ and $B(x)$.

Beside the three described operations most other Boolean operators may be used (communicativeness, associativity and distributiveness for disjunction and conjunction) and De Morgan's laws. So, for example the following operations (operators) are valid in the two sets:

1. Equivalence:
$$A(x) = B(x) \Leftrightarrow \mu_A = \mu_B \tag{9.13}$$
Two fuzzy sets are equal if they have equal membership functions.

2. Content:
$$A(x) \subset B(x) \Leftrightarrow \mu_A \leq \mu_B \tag{9.14}$$
Fuzzy set B contains fuzzy set A, if membership function value μ_B is bigger or equal to membership function value μ_A for all possible variable values.

In order to generalize operator AND the following procedures are used:

a) algebraic product:
$$\mu_C = \mu_A \mu_B \tag{9.15}$$

b) algebraic sum:
$$\mu_C = \mu_A + \mu_B - \mu_A \mu_B \tag{9.16}$$

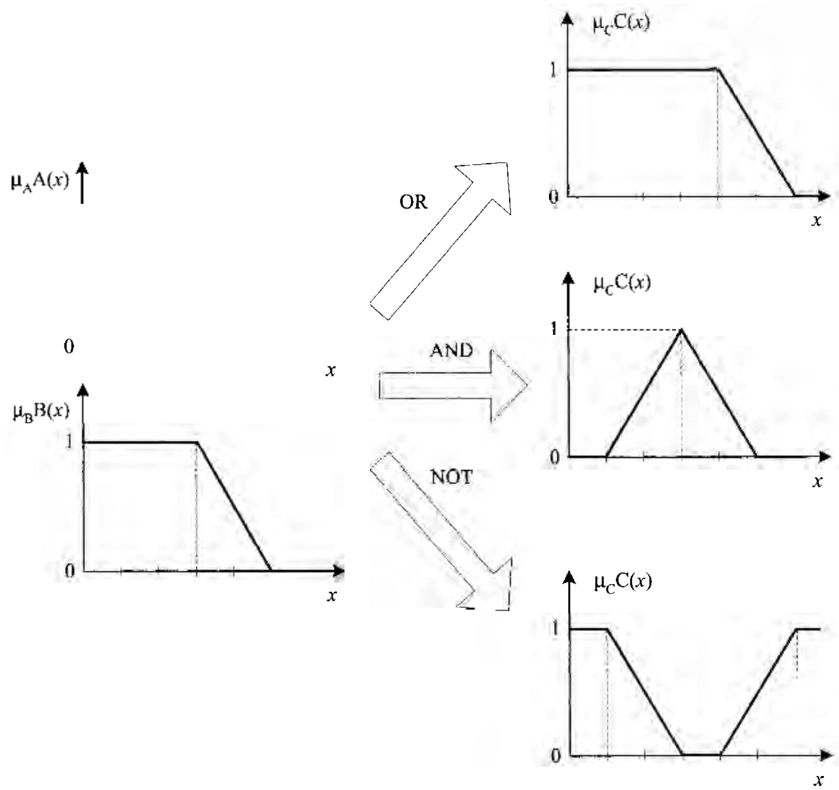


Figure 9.7: Basic procedures on fuzzy sets.

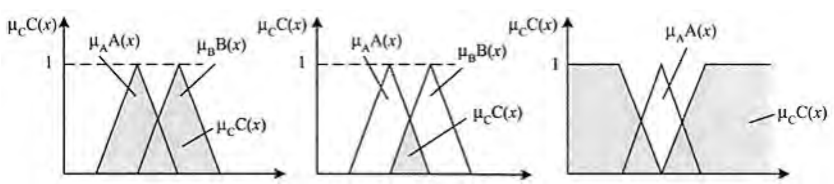


Figure 9.8: Basic procedures on triangle fuzzy sets.

c) maximum:

$$\mu_C = \max\{\mu_A, \mu_B\} \tag{9.17}$$

d) minimum:

$$\mu_C = \min\{\mu_A, \mu_B\} \tag{9.18}$$

e) bounded sum:

$$\mu_C = \min\{1, \mu_A + \mu_B\} \tag{9.19}$$

f) bounded difference:

$$\mu_C = \max\{0, \mu_A + \mu_B - 1\} \tag{9.20}$$

For fuzzy sets the following expressions are not valid:

$$A \cup \overline{A} = 1 \tag{9.21}$$

$$A \cap \overline{A} = 0 \tag{9.22}$$

which are basic expressions for crisp sets.

9.1.6 Language Variable Operators

Mathematical description of language variables is very important when fuzzy sets are dealt with. That description has to be sensitive to language variables, consistent, understandable and easy to survey. This must be respected when a language variable is used repeatedly for the same set. These fundamentals' powers (whole or fraction) describe *langauge predicative values'* operation. For example, language predicative values "little", "enough", "more or less", etc. are included into fuzzy sets through their powers. For that purpose, *language variable operators* are used, the most important being:

1. **Normative operator (NOR)** is obtained by dividing all fuzzy set values with maximum value of the set. As a result, a normative fuzzy set is obtained.

2. **Concentration operator (CON)** is designed for reducing fuzziness of the fuzzy set. It is expressed by:

$$CON(A) = A^2; \quad \text{for set } A \in X \tag{9.23}$$

Concentration operator may be used, e.g. for obtaining membership function "very big" from membership function "big".

$$\mu_{very\ big}(x) = [\mu_{big}(x)]^2$$

3. **Dilatation operator (DIL)** is designed for increasing fuzziness of the fuzzy set. It is expressed by:

$$DIL(A) = A^{1/2}; \text{ for set } A \in X \tag{9.24}$$

Dilatation operator is used when predicative variable "more or less" is realized.

4. **Operator for increasing contrast (POV)** is designed for increasing contrast in the fuzzy set. It is expressed by:

$$POV(A) = \begin{cases} CON(A); & x < 0.5 \\ DIL(A); & x \geq 0.5 \end{cases} \tag{9.25}$$

for set A with elements $x \in X$.

5. **Blurring operator (BLR)** is designed for increasing blur in the fuzzy set. It is expressed by:

$$BLR(A) = \begin{cases} DIL(A); & x < 0.5 \\ CON(A); & x \geq 0.5 \end{cases} \tag{9.26}$$

For better understanding later in the text last four operators for the fuzzy set will be shown:

$$A(x) = [0.20; 0.40; 0.60; 0.80; 1]$$

and we get:

$$\begin{aligned} A_{CON}(x) &= [0.04; 0.16; 0.36; 0.64; 1] \\ A_{DIL}(x) &= [0.45; 0.63; 0.77; 0.89; 1] \\ A_{POV}(x) &= [0.04; 0.16; 0.77; 0.89; 1] \\ A_{BLR}(x) &= [0.45; 0.63; 0.36; 0.64; 1] \end{aligned}$$

Beside the above operators, language variable operators which change membership to the function or to the set are also used.

9.1.7 General Language Variable Operators

Beside operators for carrying out basic operations and some more important operators of specific procedures on language variables, it is necessary to describe the most important general operators as well, since they are the fundamentals of the above procedures.

T-operator. T-operator defines the general class of operators by means of which conjunctions are shaped. It belongs to the so-called "triangular norms" group which are used to design conjunctions of two fuzzy sets as well as for modeling operator "AND". T-operator is a two-parameter function $t[\mu_A A(x), \mu_B B(x)]$, its definition area being $[0,1] \times [0,1]$.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

S-operator. S-operator defines a general class of operators used for joining of variables. It belongs to the so-called "triangular-co-norms" group. Most often it is labeled s-operator although it is possible to find co-t-operator. S-operator corresponds to t-operator, so it is used for modeling operator shape "OR".

Algebraic sum operator, maximum operator and limited sum operator as well as some other operators belong to s-operators.

T-norms and s-norms operators are linked together by means of De Morgan's theorem:

$$t[\mu_A A(x), \mu_B B(x)] = 1 - s\{s[1 - \mu_A A(x)], s[1 - \mu_B B(x)]\} \tag{9.27}$$

There are so many good operators that it is difficult to choose the best one. In the literature it is possible to find several good criteria for choosing operators, but beside mathematical requirements which are set for certain operators it is necessary to respect others as well.

For control technique the following operators are extremely important:

- Minimum operator,
- Maximum operator,
- Fuzzy operator "AND" (conjunction), and
- Fuzzy operator "OR" (disjunction).

9.1.8 Fuzzy Relations

So far all operations on fuzzy sets presented here referred to the elements of the same set. In practice, it is often necessary to deal with elements from two or more different sets or elements of two or more subsets.

Let X and Y be two normative sets. From their elements $x \in X$ and $y \in Y$ we have to get, according to a certain rule, elements $(x,y) \in R(x,y)$. The operation is called *binary relation* (because there are two elements in it) or just *relation*. If the set R contains elements from several sets, it is generally defined as a subset of a Cartesian product of initial sets. Therefore, binary fuzzy relation is defined as follows:

$$R = \{(x,y), \mu_R(x,y)\};\ (x,y) \in P = A \times B \tag{9.28}$$

where:

$\times$ – Cartesian product operator

A, B – subsets of X and Y sets

$$P = A \times B = \{(x,y)|x \in A, y \in B\} \tag{9.29}$$

Binary relation membership function μ_R is the function of both variables. The general form of the relation is written as a matrix of elements, i.e. membership functions. As an example of a "fuzzy relation", a membership function which is the number interval between 0.0 and 1.0 can be taken.

Cartesian product subset realization is mentioned only in general. Hence, (fuzzy) membership function Cartesian products, as a rule, are defined by a minimum operator. For the binary example it is:

$$\mu_R(x,y) = \min_{A \times B}[\mu_A(x), \mu_B(y)] \tag{9.30}$$

or shorter:

$$\mu_R(x,y) = \mu_A(x) \wedge \mu_B(y) \tag{9.31}$$

where:

$\wedge$ abbreviation for minimum operator

$\vee$ abbreviation for maximum operator

The general form of the last term has the shape of fuzzy relation matrix whose dimension is $n \times m$, where m is the number of membership function discrete values μ_A (for fuzzy set A) and n is the number of membership function discrete values μ_B (for fuzzy set B).

In order to explain more fully, an example of determining the described binary relation is given. The example of the rule "IF - THEN" is studied, i.e. IF A, THEN B, and its membership function is looked for (then, fuzzy sets A and B belong to different basic sets X and Y).

Hence, it is:

X – basic fuzzy set of the controlled value x;

A – fuzzy (sub)set of the controlled values with "big" magnitudes;

Y – basic fuzzy set of the reference value magnitude y;

B – fuzzy (sub)set of reference values with "very small" magnitude y.

The set, which should be shaped according to the rules IF A, THEN B, fits the pairs (x,y), which belong to the new basic set $P = A \times B$ put up according to a Cartesian product rule. This means that the new basic set is the origin of relationship description and dependence between sets whose elements belong to different basic sets. If it is known that basic sets have, as a rule, discrete elements, binary dependence from rectangular presentation may be taken as origin. Membership functions $\mu_A(x)$ and $\mu_B(x)$ are represented graphically as coordinates in Fig. 9.9.

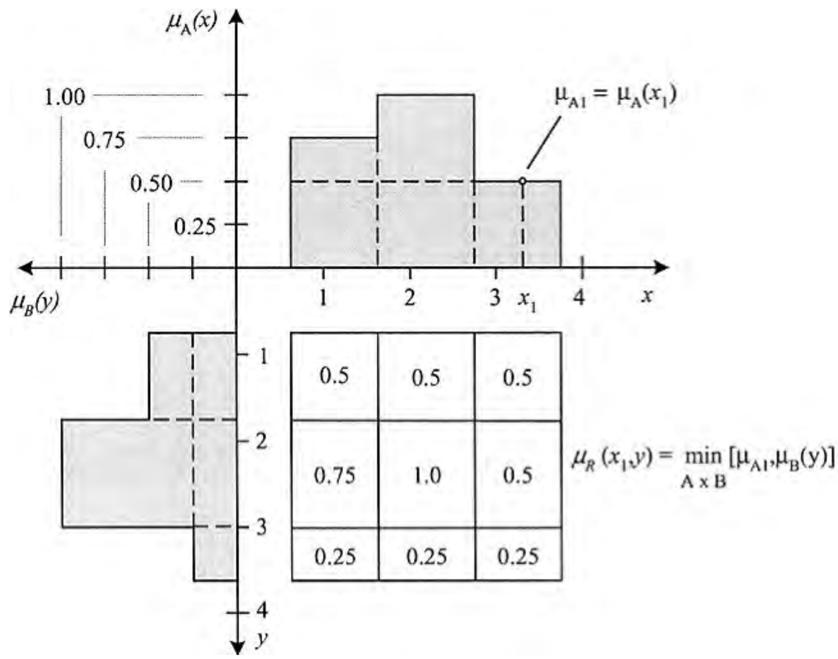


Figure 9.9: An example of membership function defining for the given fuzzy relation.

Membership function of the rule IF A, THEN B is obtained from equation (9.30). For element x_1 in fuzzy set A with attributed value $\mu_A(x_1)$ the last operation announces membership function $\mu_B(y)$ restriction to the value $\mu_{A1} = \mu_A(x_1)$. In this way the last expression changes and becomes:

$$\mu_R(x_1, y) = \min_{A \times B}[\mu_{A1}, \mu_B(y)]$$

Minimum shaping shows that the conclusion part (conclusion $\equiv$ THEN) can have premise degree (premise $\equiv$ IF). This is seen in the table (Fig. 9.9) from the example (value μ_R).

Membership function rule is determined in this way.

9.1.9 Fuzzy Relational Equations

Fuzzy relational equations connect fuzzy relations and fuzzy sets. The operators which they use belong to the group of *composition operators*. They represent an important operator group which enables membership function choice (*regulation law*) from the group of membership functions with similar properties. This procedure is called *inference* and is one of the most important procedures of fuzzy control. Its fundamentals are IF-THEN rules. In literature a large number of composition operators is found although there is just a slight difference in operation between them. The most common of these operators are the following two:

a) MAX-MIN operator (MAX-MIN inference), known as *Mamdani's minimum operator*. It consists the two known operators MAXIMUM (for fuzzy OR) and MINIMUM (for fuzzy AND).

b) MAX-PRODUCT operator (MAX-PROD inference), known also as *Larsen's product operator*. It consists the known operators MAXIMUM (for fuzzy OR) and ALGEBRAIC-PRODUCT (for fuzzy AND).

Both operators can be written like fuzzy relational equations:

$$\mu_A A(x) \circ R(x,y) = \mu_B B(y) \tag{9.32}$$

where:

$\mu_A A(x)$ multidimensional fuzzy set A with argument x and membership function discrete values:

$$\mu_A = [\mu_{A1}, \mu_{A2}, \ldots, \mu_{Ai}, \ldots, \mu_{Am}] \tag{9.33}$$

$\mu_B B(y)$ multidimensional fuzzy set B with argument y and membership function discrete values

$$\mu_B = [\mu_{B1}, \mu_{B2}, \ldots, \mu_{Bi}, \ldots, \mu_{Bm}] \tag{9.34}$$

$R(x,y)$ fuzzy relation between set A and B with arguments x and y, given by relational matrix with elements $[r_{ij}]$.

$\circ$ composition operator.

In equation (9.32) the $\mu_B B(y)$ is solved by inference of $\mu_A A(x)$ and $R(x,y)$. Two solutions are possible:

a) for MAX-MIN operator, the equation becomes:

$$\mu_B B(y) = \max[\min(\mu_A A(x), R(x,y))] \tag{9.35}$$

i.e. for individual components:

$$\mu_{Bj} b_j = \max[\min(\mu_{Ai} a_i, r_{ij})]; \;\; 1 \leq i \leq m; \; 1 \leq j \leq n \tag{9.36}$$

b) for MAX-PROD operator the equation becomes:

$$\mu_B B(y) = \max[\mu_A A(x) \times R(x,y)] \tag{9.37}$$

i.e. for individual components:

$$\mu_{Bj} b_j = \max[\mu_{Ai} a_i r_{ij}]; \;\; 1 \leq i \leq m; \; 1 \leq j \leq n \tag{9.38}$$

In both examples it is seen that matrix product operations, i.e. addition and multiplication are replaced by maximum and minimum operations. In the equation (9.37) minimum is replaced by simple multiplication.

9.1.10 Use of Language Variables and Language Expressions

When using fuzzy sets language expressions are often used. Language expressions make/form larger fuzzy set $M(hx)$ from fuzzy set $M(x)$ where:

$$hM(x) \Rightarrow M(hx) \tag{9.39}$$

In this case different operators are used: concentration operator, dilatation operator, negation operator, etc. For better understanding, an example follows:

EXAMPLE 9.1
The term "tall" has been given with fuzzy set:

$$M(x) = (0.0/10), (0.1/20), (0.4/30), (0.8/40), (1.0/50)$$

It is necessary to write expression:

a) "very tall" (operator h = "very")

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

b) "very very tall" (operator h= "very very")

c) "not very tall" (operator h = "not very")

Solution.

a) For "very tall" concentration operator given in (9.23) is used:

$$M_a(hx) = CON[M(x)] = [M(x)]^2$$

and we get:

$$M_a(hx) = (0.0/10), (0.01/20), (0.16/30), (0.64/40), (1.0/50)$$

b) For "very very tall" concentration operator is used twice:

$$M_b(hx) = CON[M_a(hx)] = CON\{CON[M(x)]\}$$

and we get:

$$M_b(hx) = (0.0/10), (0.0001/20), (0.0256/30), (0.4096/40), (1.0/50)$$

c) For "not very tall" negation operator above operator "very tall" is used. Now it is:

$$M_c(hx) = 1 - M_a(hx) = 1 - CON[M(x)]$$

and we get:

$$M_c(hx) = (1.0/10), (0.99/20), (0.84/30), (0.36/40), (0.0/50)$$

Other operators for making wider fuzzy sets are used in a similar way. ■

9.1.11 Fuzzification

If the data already have language values, then writing them by means of basic fuzzy sets is simple. When we have measurement results (real numbers) individual data should be associated with corresponding basic fuzzy sets. In control applications such situations are often found before inference (decisionmaking process).

Real number x_0 in the set of real numbers X is a degenerated fuzzy set with membership function:

$$\mu(x) = \begin{cases} 1, & \text{for } x = x_0 \\ 0, & \text{for } x \neq x_0 \end{cases} \tag{9.40}$$

where $x_0 \in X$.

Basic fuzzy sets are defined in the space of X. The condition of overlapping must be fulfilled. At least one fuzzy set must overlap each x in X. In other words, basic fuzzy set disjunction covers all X.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

9.1.12 Language Description of the System by Means of IF-THEN Rules

The target of system language description by the IF-THEN rules is the realization of the inference procedure which, in a way, changes fuzzy and crisp information into shapes that the human describes as "possible".

The system is described by IF-THEN rule set:

$$C = \{R_i;\ i = 1,2,3,\ldots,N\} \tag{9.41}$$

where R_i is an individual rule and N is the number of the rules.

Individual rule R_i is a fuzzy statement with *basic form*, which has one condition (input) and one consequence (output):

$$R_i:\ \ IF(X \text{ is } A_i), THEN(Y \text{ is } B) \tag{9.42}$$

where X, Y are fuzzy sets with elements A and B; A is the system input and B is the system output.

In classic (binary) logic only one implication is used which logic algebra postulates make possible:

$$I(x,y) = x \Rightarrow y = \bar{x} \cup y \tag{9.43}$$

where $\Rightarrow$ is the implication operator; $\bar{x}$ is the negation operator and $\cup$ is the disjunction operator.

In this relation operands x and y can have only two possible values: true (1) or false (0). The same applies to the result.

With fuzzy sets for calculating $Y = B'$ at $X = A'$ from the shape rule (9.42): IF $(X = A)$, THEN $(Y = B)$ procedure *compositional rule of inference* is used. Fundamental procedure is the use of operator *T-normal* and different sorts of implications I, i.e. inference compositional rules, which have the following shape:

$$\mu_B(y) = \sup T\left\{\mu A'(x), I[\mu_A(x), \mu_B(y)]\right\} \tag{9.44}$$

As implication operator I, a large number of different implications can appear. The following operators are often used:

- Algebraic product (expression (9.15)),
- Bounded difference (expression (9.20)),
- Drastic product, or
- Minimum (expression (9.18)).

It is possible to connect implication operators to operator max() and its dual operators. This could be the basis for increasing the list of possible implication operators.

Calculating $Y = B'$ by expression (9.42) compositional rules of inference are used whose shape is (9.44) and bring *T-norm* as a base with them. With T-norm different operators are also used, mostly minimum, algebraic product, bounded difference and drastic product, i.e. operators said to be used as implication operators.

For describing the example, only one rule R_i has been used and with one shape supposition (9.42). For practical use it is not enough, so when modeling the system it is necessary to use more basic or combined rules. In the language model each rule is a fuzzy relation. This can also be said for the overall fuzzy relation which represents all rules of the language model:

$$\begin{aligned}
&R_1: \ IF(X \text{ is } A_1)\ THEN\ (Y \text{ is } B_1)\\
&R_2: \ IF(X \text{ is } A_2)\ THEN\ (Y \text{ is } B_2)\\
&\vdots\\
&R_i: \ IF(X \text{ is } A_i)\ THEN\ (Y \text{ is } B_i)\\
&\vdots\\
&R_n: \ IF(X \text{ is } A_n)\ THEN\ (Y \text{ is } B_n)
\end{aligned} \tag{9.45}$$

In this way system language description can be written as a single rule disjunction (relation):

$$R_S = \bigcup_i R_i;\ 1 \leq i \leq q \tag{9.46}$$

where:

R_S system model fuzzy relation (fuzzy relation matrix of the system)

R_i fuzzy relation of the single rule (fuzzy relation matrix of single/individual rule)

q system model language rule number

$\bigcup_i$ disjunction operator defined by maximum value choice

$$\bigcup_i (y_1, y_2, \ldots, y_i, \ldots) = \max(y_1, y_2, \ldots, y_i, \ldots)$$

The whole procedure is very complex and requires a lot of work. An adequate implicator is chosen for the particular rule. The rules which have been used make the corresponding language model.

In the way which has been described so far, the language model of the system may be calculated by means of a fuzzy relation matrix. This model cannot be used for dynamic systems. In order to describe dynamic systems it is necessary to use "IF - THEN" and their expanding operators, i.e. AND and OR in their conditional part. Then the rule (9.42) becomes (as an example of operator AND use):

$$R_i: \ IF(X \text{ is } A)\ AND\ (Z \text{ is } C)\ THEN\ (Y \text{ is } B) \tag{9.47}$$

where X and Z are input quantities; A and C are observed values of the input quantities (e.g. desired values); Y is the output quantity; and B is the momentary value of the output quantity.

Operator THEN as well as operator AND can be used several times (in control applications it is often $C_i = dA_i/dt$).

For the values A, C, and B predicative expression combinations are often used: N - negative, P - positive, ZE - zero, S - small, M - medium, B - big. From the above predicates it is possible to make seven combinations which are used very often:

NB Negative Big
NM Negative Medium
NS Negative Small
ZE ZEro
PB Positive Big
PM Positive Medium
PS Positive Small

It is possible to present (by using the above labels) the language model, i.e. individual shape rule (9.46) by a table. Here, language variables are presented as fuzzy sets (expressions, terms). Such presentation is often found in the literature because it makes detailed model design possible. The above presentation is explained by means of the rule example:

$$IF\ (v = PS)\ AND\ (v' = NS)\ THEN\ (y = NS)$$

Figure 9.10 shows seven-degree combination by means of the table, and Fig. 9.11 five-degree combination (PB, PM ZE, NS, NB), where v is the control error, v' is the control error change ($v' = dv/dt$), and y is the actuator position (given in the table).

Table or table-graphic presentation is simple for presenting a connection of three variables (in given example connection of two variables in the cause and one variable in the effect part of the rule).

$v \downarrow$ / $v' \rightarrow$

	NB	NM	NS	ZE	PS	PM	PB
NB							
NM							
NS							
ZE							
PS							
PM							
PB			NS				

Figure 9.10: Table presentation of individual rule in seven-column combination.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

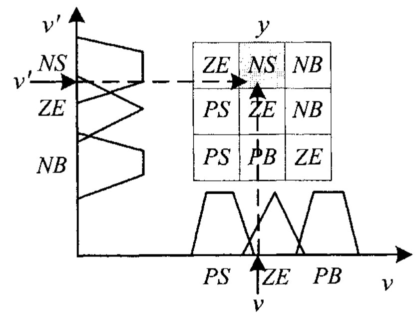


Figure 9.11: Graphic-table presentation of individual rule in five-column combination.

It is possible to present the dynamic system model in different ways. An example of that is fuzzy decisionmaking.

9.1.13 Language Description of the System with Fuzzy Decision Making

Fuzzy decisionmaking procedure is used with systems defined by two or more augmented IF-THEN rules.

The procedure belongs to the fuzzy logic area defined by the strict mathematical accuracy. In control application two forms of fuzzy decisionmaking are often found. They are simple to calculate.

First form of decisionmaking. IF-THEN rule is observed. It is given in the expression (9.47) meaning rules for system with two inputs and one output, given by one rule:

$$R_i: \; IF\ (X\ is\ A_i)\ AND\ (Z\ is\ C_i)\ THEN\ (Y\ is\ B_i) \tag{9.48}$$

If we denote:

$$w_{ai} := (x\ is\ A_i),\ and\ w_{ci} := (z\ is\ C_i) \tag{9.49}$$

then w_{ai} and w_{ci} are *partial causes* (premises) which can be processed in different ways. In control applications it is often calculated by means of expression for i-th rule:

$$w_{ai} = (x\ is\ A_i) = \max[\min(x, A_i)] \tag{9.50}$$

$$w_{ci} = (z\ is\ C_i) = \max[\min(z, C_i)] \tag{9.51}$$

Values $w_{ai}, w_{ci}, \ldots$ are known as *cause (premise) weight* or *cause measure*:

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

In the rule (9.48) we have conjunction AND, which connects cause partial weight into *complete weight* — w_i. For the example of the system with two inputs, the following expression is used:

$$w_i := \min(w_{ai}, w_{ci}) \tag{9.52}$$

In the rule (9.48), operator OR may appear instead of operator AND. In that case w_i is most often calculated by means of the expression:

$$w_i := \max(w_{ai}, w_{ci}) \tag{9.53}$$

In the rule (9.48) operators AND and OR may also appear on the right side. Then the rule becomes:

$$R_i: \ IF\ (X\ is\ A_i)\ AND\ (Z\ is\ C_i)\ THEN\ (Y_1\ is\ B_i)\ AND\ (Y_2\ is\ D_i) \tag{9.54}$$

In such cases the conjunction AND on the right side is present as syntactic connection at numerating independent outputs $Y_1, Y_2, \ldots$. On the contrary, when it is necessary to create one output, the same rules as those on the left side can be used.

Let's see the same system (9.48) with two IF-THEN rule forms:

$$\begin{aligned} R_1: \ & IF\ (X_1\ is\ A_{11})\ AND\ (Z_1\ is\ C_{11})\ THEN\ (Y\ is\ B_1) \\ R_2: \ & IF\ (X_1\ is\ A_{21})\ AND\ (Z_1\ is\ C_{21})\ THEN\ (Y\ is\ B_2) \end{aligned} \tag{9.55}$$

Weights are mostly calculated in the following two ways (other ways are possible):

1st way (using min operator):

$$\begin{aligned} w_i &= \min[A_{11}(x_0), C_{11}(z_0)] \\ w_i &= \min[A_{21}(x_0), C_{21}(z_0)] \end{aligned} \tag{9.56}$$

2nd way (using product operator):

$$\begin{aligned} w_i &= A_{11}(x_0) C_{11}(z_0) \\ w_i &= A_{21}(x_0) C_{21}(z_0) \end{aligned} \tag{9.57}$$

Then the following expressions are given, i.e. $w_1 B_1$ and $w_2 B_2$, and a fuzzy set is formed:

$$B^* = w_1 B_1 \cup w_2 B_2 \tag{9.58}$$

Second form of decisionmaking. If we study the two-input and one-output systems again, given by two rules (system form (9.55)), but shaped by means of a fuzzy positive (P) and negative (N) set:

$$\begin{aligned} R_1: \ & IF\ (X_1\ is\ N)\ AND\ (X_2\ is\ P)\ THEN\ (Y_1\ is\ N) \\ R_2: \ & IF\ (X_1\ is\ P)\ AND\ (X_2\ is\ N)\ THEN\ (Y_2\ is\ P) \end{aligned} \tag{9.59}$$

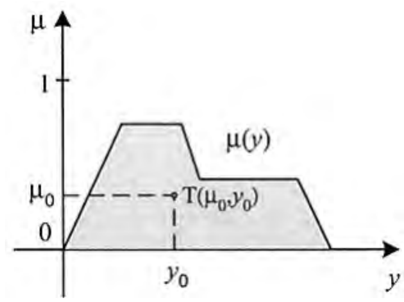


Figure 9.12: Calculating output value by method of gravity center.

In this case membership functions P and N are monotonous functions. Weights w_1 and w_2 are obtained by the expression:

$$\begin{aligned} w_1 &= N(Y_1) \\ w_2 &= P(Y_2) \end{aligned} \tag{9.60}$$

The first decisionmaking variant is most often found in control systems. Beside the above forms, the third form is also found. It is used for systems given by rules (9.55), which have the output defined by the expression $y_i = f_i(x_1, x_2, \ldots, x_i, \ldots, x_n)$.

9.1.14 Defuzzification or Fuzzy Set Adjustment (Calculating Crisp Output Values)

The result of the system calculated with fuzzy sets is also a fuzzy set. The influence of such fuzzy value on the given (technical) system is possible only by means of crisp signals. If, at the beginning, it was necessary to soften or fuzzify the system, now the real numbers are required to control it. This leads to the problem of inference result *adjustment* or *defuzzification*. In control applications it means that fuzzy sets should not be brought to an actuator (motor, valve, potentiometer, etc.) for its control, because the system cannot understand them. Real values (real numbers) of the voltage, pressure, flow or other control value should be applied.

The fuzzification process (inference and composition—subsection 9.1.9) for given momentary input variable values gives as an output information (based on defined rule) the membership function $\mu(y)$, which is a fuzzy set. An example of such a result may be seen in Fig. 9.12.

Looking at membership function $\mu(y)$, the crisp value for y should be calculated. It is the output of the system (set value — regulator output). There are several methods for defuzzification in control applications, but three of them are most often used:

1. *Max-height* method,
2. *Mean of maximum* method, and
3. *Center of gravity* method or *center of area* method.

The first method gives output value y_0 by means of membership function $\mu(y)$ belonging to output set Y according to the following expression:

$$\mu(y_0) = \max_y \mu(y_0) \tag{9.61}$$

The method becomes cumbersome if membership function $\mu(y)$ has maximum at more discrete values.

With the second method the problem is solved by calculating y_0 as the mean arithmetic value of maximum membership function value $\mu(y)$. However, this method is also inconvenient, i.e. if the output set membership function has a particular area (length) where it can have maximum value.

The third method takes into account the area below a membership function $\mu(y)$ like surface area (Fig. 9.12). Crisp output value y_0 is determined according to the gravity center coordinates $T(\mu_0, y_0)$ of the area and is calculated as follows:

$$y_0 = \frac{\int_y y\mu(y)dy}{\int_y \mu(y)dy} \tag{9.62}$$

If the membership function consists of forms made up of straight lines, then the integrals in the last expression can be replaced by additions and calculation can be carried out analytically.

In case of m-polygon segments described by two points, $P_k(\mu_k, y_k)$ and $P_{k+1}(\mu_{k+1}, y_{k+1})$, the gravity center abscissa (9.62) is obtained from the expression:

$$y = \frac{\sum_{k=1}^{m}(y_{k+1} - y_k)[(2y_{k+1} + y_k)\mu_{k+1} + (2y_k + y_{k+1})\mu_k]}{3\sum_{k=1}^{m}(y_{k+1} - y_k)(\mu_k + \mu_{k+1})} \tag{9.63}$$

Even more difficult than calculating the gravity center abscissa is preliminary calculation of intersection points of overlapping membership functions of individual fuzzy rules.

The center of gravity method is most often used in fuzzy control since compared with other methods it has two advantages. Its first advantage is that it not only processes membership function maximum values, but it includes its complete flow. The second advantage of this method is a relatively simple way of calculating crisp values for y_0.

In the literature, an even simpler way of calculating crisp values for y_0 can be found. In that procedure, membership functions which refer to the parts of the rule THEN are represented by only one numerical value different from zero

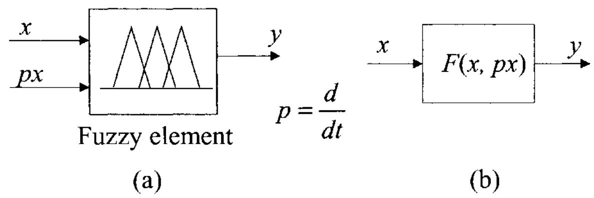


Figure 9.13: Description of the fuzzy logic element with SISO nonlinear element.

(singleton). If S_i denotes n-limited singleton values for n-rules, the expression (9.63) becomes:

$$y_0 = \frac{\sum_{i=1}^{n}(y_i S_i)}{\sum_{i=1}^{n}(S_i)} \tag{9.64}$$

such expression makes the whole defuzzification process less complicated.

9.2 Describing Function of SISO Fuzzy Element

From the basic properties of fuzzy elements, described in a previous section, we can conclude that the fuzzy element is a static nonlinear element which cannot be exactly mathematically formalized. This fact makes the determination of dynamics of systems with fuzzy elements more difficult. Analogously to procedures described in previous chapters, the describing function method can be applied to a class of nonlinear systems with a fuzzy control algorithm. For this purpose it is necessary to determine the describing function of a fuzzy element. Knowing the describing function of a fuzzy element, and describing functions of other nonlinear elements in the system, enables the analysis of a nonlinear system in the frequency domain.

To determine the describing function of a fuzzy element, it is necessary to substitute the fuzzy element with equivalent nonlinear element with a single input and single output (SISO element), see Fig. 9.13 (Kuljača et al., 1999). The describing function can be obtained by experimental method[1], using the simulation model shown in Fig. 9.14, where:

X_m, ω – amplitude and frequency of periodic input signal
k_p, k_d – proportional and derivative coefficients, respectively
$y_N(t)$ – signal at the output of fuzzy element,
Y_m – amplitude of a first harmonic of $y_N(t)$

[1]This procedure is, contrary to mathematical procedures, relatively simple and efficient.

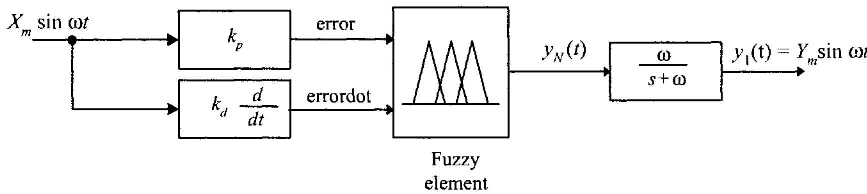


Figure 9.14: The circuit scheme for experimental determination of the describing function.

EXAMPLE 9.2

Determine the describing function $G_N(X_m)$ of a fuzzy element (Fig. 9.14), specified by rules in Table 9.1, membership functions in Fig. 9.15, and coefficients $k_p = 1$; $k_d = 0.15[s^{-1}]$.

The experimental results from simulation model in Fig. 9.14 are shown in Table 9.2 and Figs. 9.16 and 9.17.

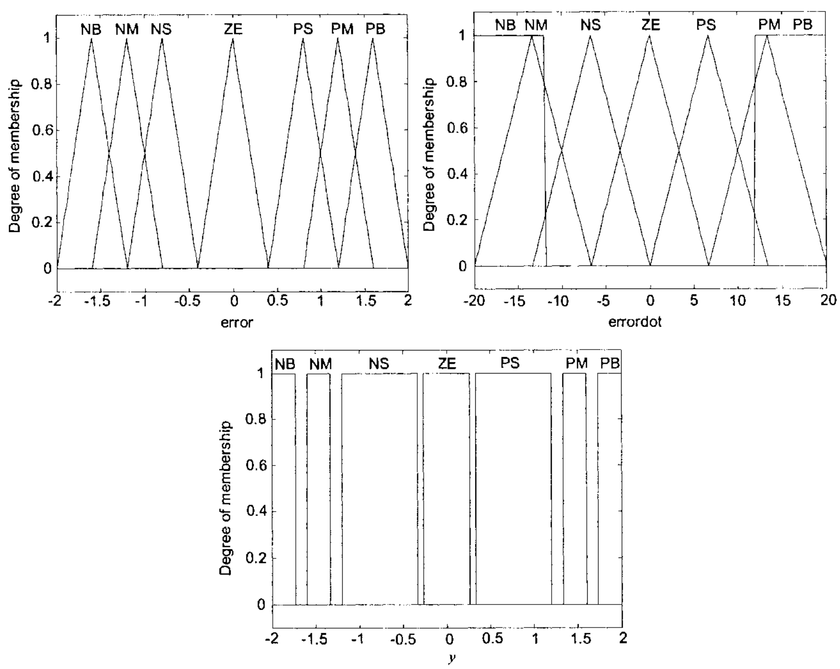


Figure 9.15: Membership functions of the fuzzy element inputs 'error' and 'errordot' and fuzzy element output $y = y_N$.

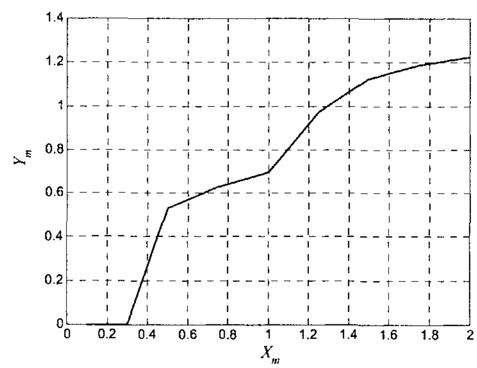


Figure 9.16: The static characteristic of $Y_m = f(X_m)$.

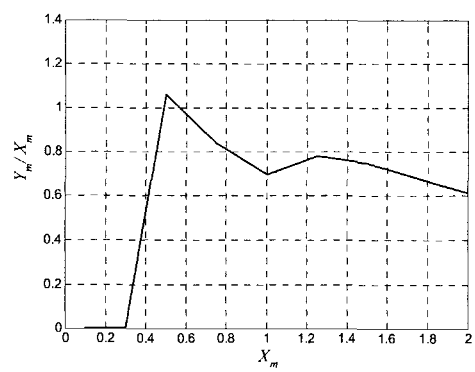


Figure 9.17: Describing function $G_N = Y_m/X_m = f(X_m)$.

Table 9.1: Rule base of the fuzzy element.

$errordot \backslash ^{error}$	NB	NM	NS	ZE	PS	PM	PB
NB	NB	NB	NB	NB	NM	NS	ZE
NM	NB	NB	NB	NM	NS	ZE	PS
NS	NB	NB	NM	NS	ZE	PS	PM
ZE	NB	NM	NS	ZE	PS	PM	PB
PS	NM	NS	ZE	PS	PM	PB	PB
PM	NS	ZE	PS	PM	PB	PB	PB
PB	ZE	PS	PM	PB	PB	PB	PB

Table 9.2: Simulation results.

X_m	Y_1	$G_N(X_m)$
1.999	1.224	0.61234
1.75	1.1876	0.6786
1.5	1.1237	0.7491
1.25	0.9725	0.7780
1	0.697	0.697
0.75	0.629	0.8387
0.5	0.5289	1.0578
0.3	1.123e-3	0.0037
0.1	2.47e-4	0.0025

The presented procedure is relatively simple and efficient for application in engineering practice. From the graphical representation of $G_N(X_m)$ it is possible to determine appropriate mathematical approximations necessary for mathematical description of a class of nonlinear systems with a fuzzy control algorithm.

9.3 Stability Analysis of a Fuzzy Control System

Stability is the most important feature of any feedback system. Without stability disturbance compensation, steady-state performance and any other performance index are not possible. In order to reduce the risk of implementing a fuzzy controller it is necessary to analyze stability. Fuzzy control systems are essentially knowledge-based systems that use fuzzy set theory for knowledge representation and inference. The definition of the fuzzy control rules is relatively simple. However, the assumption that the heuristic rules are robust is not correct and some applications prove that. This fact has motivated the research on the analysis techniques since mid-seventies. As a fuzzy control system can be considered as a nonlinear system, achievements from the nonlinear stability theory should be explored. In the literature (Maeda and Murakami, 1991; Opitz, 1993; Kiendl and Rüger, 1993; Kiendl, 1993; García-Cerezo, Ollero and Aracil, 1992; Tanaka and Sugeno, 1992) results from this research can be found. However, in this book

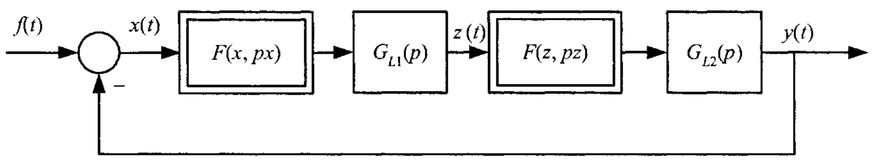


Figure 9.18: The structure of a nonlinear system with fuzzy regulator.

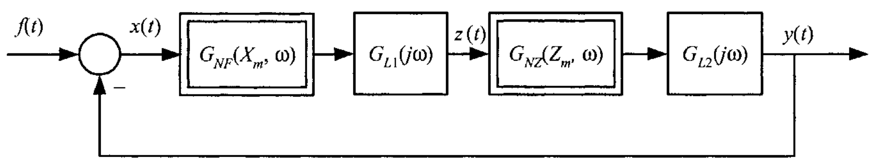


Figure 9.19: Block diagram of harmonically linearized system given in Fig. 9.18.

we explored the describing function method for stability analysis. The method is attractive for the engineering practice because it is simple and gives better insight of effects which the fuzzy element can have on the stability of the closed-loop control system.

The procedure to determine the describing function of a fuzzy element that can be substituted with equivalent SISO element was presented in Section 9.2. By knowing the describing function of a fuzzy element, the mathematical formalization of nonlinear systems with fuzzy regulator is made possible. We will present the procedure to determine stability of a class of nonlinear systems with structure as in Fig. 9.18.

The notation in Fig. 9.18 is:

$F(x, px)$ – nonlinear fuzzy regulator

$F(z, pz)$ – nonlinearity in the system

$G_{L1}(s), G_{L2}(s)$ – transfer functions of linear parts of the system that conform to the filter hypothesis

By conducting the harmonic linearization of the system given in Fig. 9.18, the block diagram in Fig. 9.19 can be obtained.

The notation in Fig. 9.19 is:

$G_{NF}(X_m)$ – describing function of the fuzzy regulator

$G_{NZ}(Z_m)$ – describing function of the nonlinearity $F(z, pz)$

$G_{L1}(j\omega), G_{L2}(j\omega)$ – frequency characteristics of linear parts of the system

The closed-loop characteristic equation of the system in Fig. 9.19 is:

$$1 + F(X_m)G_{L1}(j\omega)F_Z(Z_m,\omega)G_{L2}(j\omega) = 0 \tag{9.65}$$

The stability can be analyzed by use of some previously presented stability analysis criteria.

EXAMPLE 9.3
Determine the stability of the power system secondary-load frequency control (Šijak et al., 2002), with model shown in Fig. 9.20, where:

$G_G(s) = 1/(1+sT_G) = 1/1+0.08s$ – *transfer function of the turbine governor*

T_{CH} – *stream turbine time constant*

ΔP_m – *change of the mechanical power of the turbine*

ΔP_L – *change of power system load*

ΔP_r – *power system active power reference change*

ΔP_r – *power system frequency change*

$G_s(s) = K_s/(1+sT_s) = 120/1+20s$ – *the power system transfer function*

$R = 2.4$ – *the static speed drop of the uncontrolled system*

$F(z)$ – *the power system generation rate constraint, static characteristic of saturation nonlinearity (Fig. 9.21)*

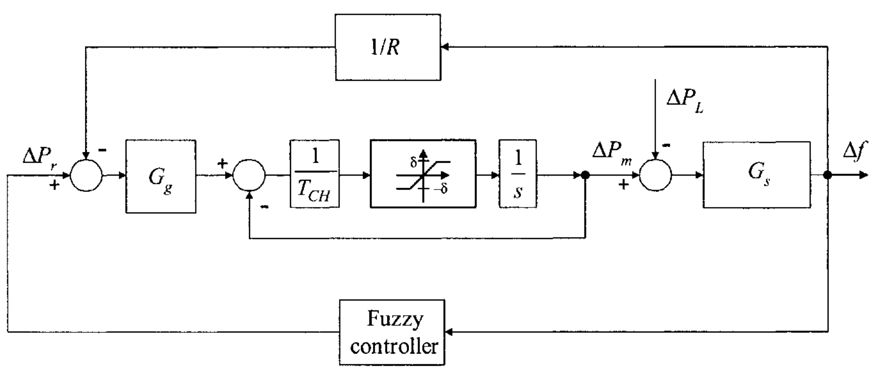


Figure 9.20: Power system secondary-load frequency control model (Example 9.3).

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

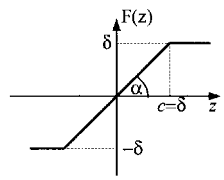


Figure 9.21: Saturation nonlinearity $F(z)$.

The describing function of nonlinearity in Fig. 9.21 is:

$$G_N(Z_m) = \frac{4}{\pi}\left(\frac{\alpha}{2} - \frac{\sin 2\alpha}{4} + \frac{cos\alpha}{\frac{Z_m}{\delta}}\right) \tag{9.66}$$

$$\alpha = \arcsin\frac{\delta}{Z_m} \tag{9.67}$$

The fuzzy regulator is determined with Table 9.3 and Figs. 9.22a, b and c.

The Ziegler-Nichols tuning method for conventional PD regulator parameter tuning (Ziegler and Nichols, 1942) was used to determine k_p and k_d pa-

Table 9.3: Rule base of fuzzy regulator.

errordot \ *error*	NB	NM	NS	ZE	PS	PM	PB
NB	NB	NB	NB	NB	NM	NS	ZE
NM	NB	NB	NB	NM	NS	ZE	PS
NS	NB	NB	NM	NS	ZE	PS	PM
ZE	NB	NM	NS	ZE	PS	PM	PB
PS	NM	NS	ZE	PS	PM	PB	PB
PM	NS	ZE	PS	PM	PB	PB	PB
PB	ZE	PS	PM	PB	PB	PB	PB

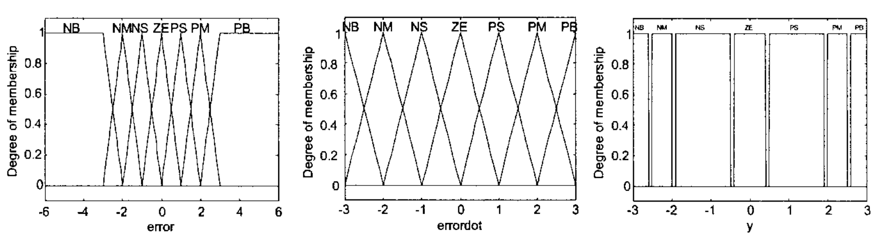


Figure 9.22: Membership functions of the fuzzy regulator. (a) proportional input membership function, (b) derivative output membership function, (c) output membership function.

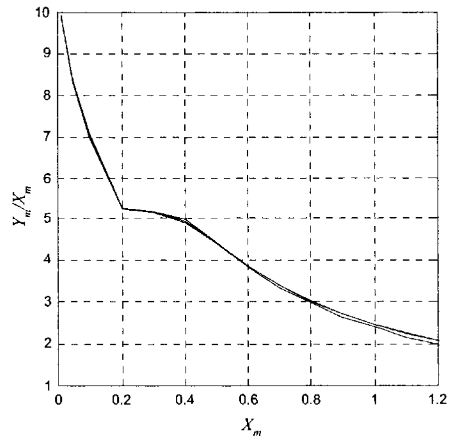


Figure 9.23: Describing function of fuzzy regulator.

rameters of SISO equivalent for given fuzzy regulator. The obtained values are $k_p = 6.9134$; $k_d = 0.4564$. *The graphical representation of describing function of fuzzy element—obtained by the procedure in previous section—is given in Fig. 9.23*

Solution. *Substitution of given values into (9.65) yields the characteristic equation of the system:*

$$\begin{aligned} T_{CH}T_sT_Gs^3 &+ [T_GT_sG_{NZ} + T_{CH}(T_G+T_s)]s^2 \\ &+ [T_{CH} + (T_G+T_s)G_{NZ}]s \\ &+ G_{NZ} + (K_0 + b_1G_{NF})K_sG_{NZ} = 0 \end{aligned} \tag{9.68}$$

where $K_0 = 1/R$.

Applying the Hurwitz stability criterion on equation (9.68) yields the following stability condition:

$$\begin{aligned} &[T_GT_sG_{NZ} + T_{CH}(T_G+T_s)][T_{CH} + (T_G+T_s)G_{NZ}] \\ &\quad -T_{CH}T_sT_G[G_{NZ} + (K_0 + b_1G_{NF})K_sG_{NZ}] > 0 \end{aligned} \tag{9.69}$$

From (9.69) the parametric equation of stability boundary is derived:

$$\begin{aligned} G_{NF} = &\frac{(T_G+T_s)G_{NZ}}{b_1T_{CH}K_s} + \frac{T_{CH}(T_G+T_s)^2 - T_GT_sT_{CH}K_0K_s}{b_1T_{CH}T_sT_GK_s} \\ &+ \frac{T_{CH}(T_G+T_s)}{b_1T_sT_GK_sG_{NZ}} \end{aligned} \tag{9.70}$$

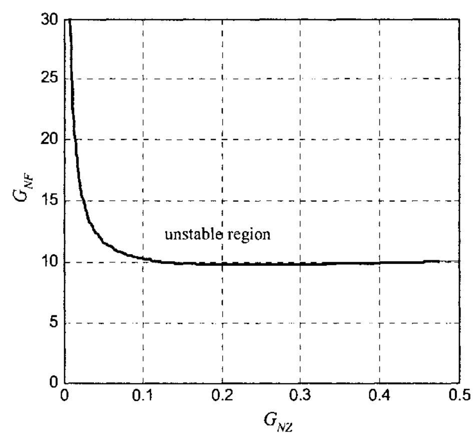


Figure 9.24: Stability boundary $G_{NF} = f(G_{NZ})$.

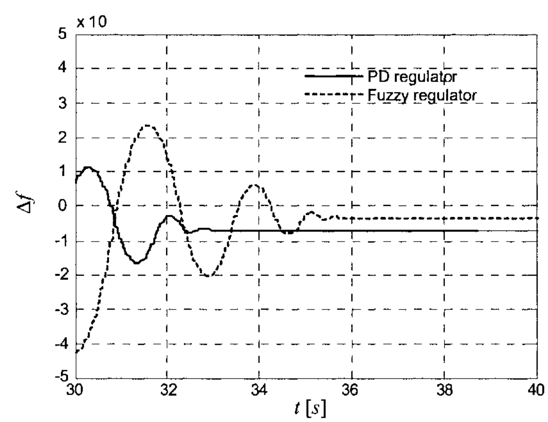


Figure 9.25: Frequency Δf response for $\Delta P_L = 0.5\%$.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

The stability boundary for given parameter values is shown in Fig. 9.24.

Figure 9.25 shows system responses when a conventional PD regulator and a fuzzy regulator was used. The parameters k_p and k_d are determined using the Ziegler-Nichols method. From the obtained results (Figs. 9.23, 9.24 and 9.25), we can conclude that it is justifiable to use tuning methods for conventional regulators to preliminarily determine parameters k_p and k_d of the SISO equivalent.

It can be seen in Fig. 9.25 that in this example it is better to use the fuzzy regulator. A much larger static error Δf is accomplished when a conventional PD regulator is used. ■

From the latter discussions, it can be concluded that these preliminary results on using the describing function in analysis of a specific class of nonlinear systems justifies further research of application of harmonic linearization in nonlinear systems with fuzzy logic and neural networks.

9.4 Influence of Fuzzy Regulator on Resonance Jump

In Sections 1.3 and 6.2.3 we have analyzed the occurrence of resonance jump in nonlinear systems. The procedure to determine frequency properties of the class of nonlinear systems that conform to the filter hypothesis was presented. In Example 6.7 the procedure to determine frequency characteristics of the system in Fig. 6.26 was illustrated. In Section 9.3, Example 9.3, we have shown the procedure of using a describing function in stability analysis of a system with a fuzzy regulator. In the following example, the influence of a fuzzy control algorithm on establishment of resonance jump will be analyzed (Kuljača et al., 2002).

EXAMPLE 9.4
For the system with fuzzy regulator shown in Fig. 9.26, determine the frequency characteristic of signal $X_m(\omega_v)$.

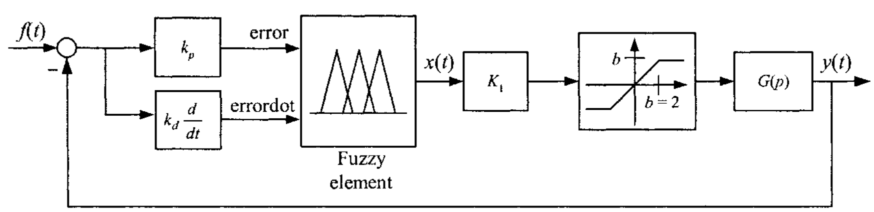


Figure 9.26: System with fuzzy regulator (Example 9.4).

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

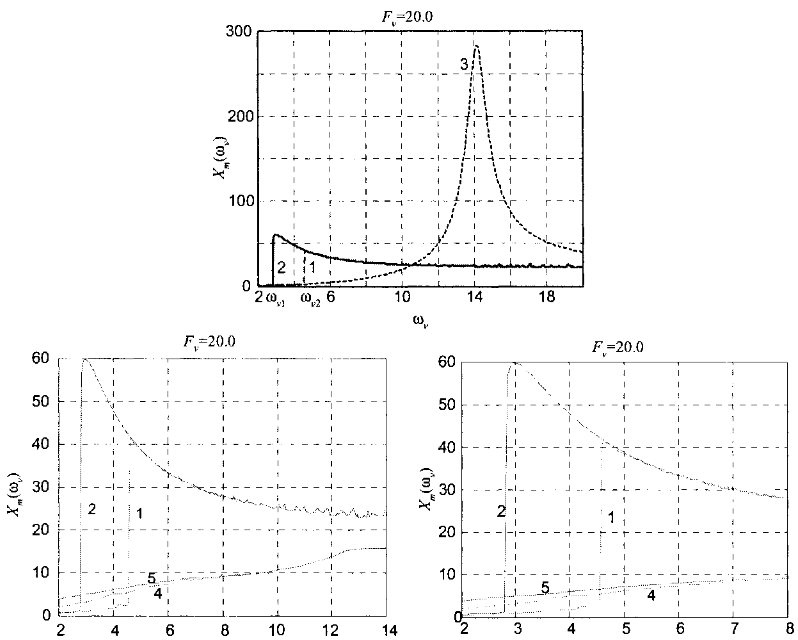


Figure 9.27: Frequency response of the system in Example 9.4. Legend: 1–increasing frequency (without fuzzy regulator); 2–decreasing frequency (without fuzzy regulator); 3–linear system; 4–increasing frequency (with fuzzy regulator); 5–decreasing frequency (with fuzzy regulator).

The quantities in Fig. 9.26 are $K_1 = 2$; $p = d/dt$; $k_p = 0.05$, $k_d = 0.9$, *coefficients of proportional, i.e. derivative action of equivalent SISO element, and* $G(p) = 200/p(p+1)$, *transfer function of linear part of the system.*

The fuzzy regulator is determined with Fig. 9.22 and Table 9.3 from the example in Section 9.3.

Solution. *The frequency characteristic* $X_m(\omega_v)$ *was determined by simulation for different values of the amplitude* F_v *of external excitation signal* $f(t) = F_v \sin \omega t$.

Figure 9.27 shows frequency characteristics obtained by simulation for the amplitude $F_v = 20$. *From these results it can be concluded that the application of a fuzzy controller decreased the system gain, increased the linear range of the nonlinear element, increased the system's sensitivity at the low-frequency range and eliminated the establishment of resonance jump in the possible frequency range.*

■

Based on the results from the previous example, we can conclude that further research of influence of the fuzzy control algorithms on nonlinear system dynamic properties is justifiable.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Appendix A

Harmonic Linearization

Table A.1: Describing functions of standard nonlinear elements, $G_N(A) = P_N(A) + jQ_N(A)$, $A = X_m/x_a$.

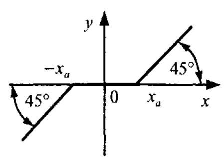
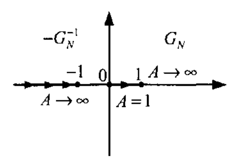
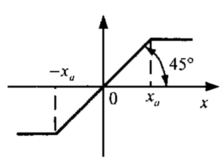
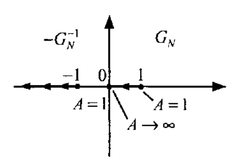
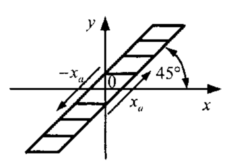
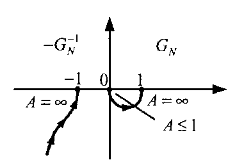
Name	Characteristic	$G_N(A)$	Diagram
Dead zone		$P_N = 1 - \dfrac{2\alpha + \sin 2\alpha}{\pi}$ $Q_N = 0$ $\sin\alpha = 1/A$	
Saturation		$P_N = \dfrac{2\alpha + \sin 2\alpha}{\pi}$ $Q_N = 0$ $\sin\alpha = \dfrac{1}{A}$	
Backlash		$P_N = \dfrac{1}{2} - \dfrac{2\alpha + \sin 2\alpha}{2\pi}$ $Q_N = -\dfrac{\cos^2\alpha}{\pi}$ $\sin\alpha = \dfrac{1}{A} - 1$	

Table A.1: Continued.

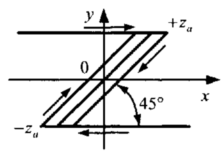
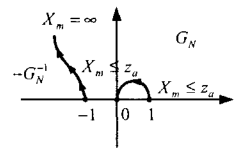
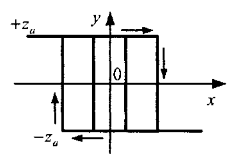
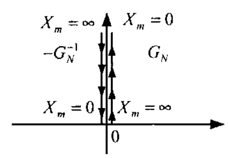
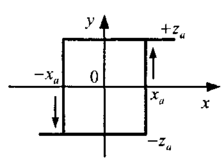
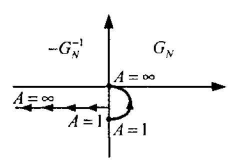
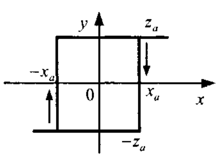
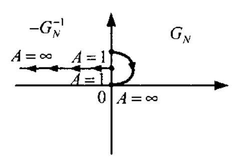
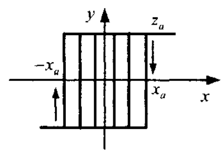
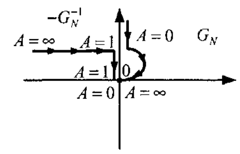
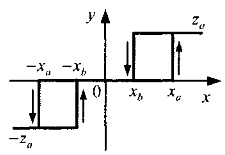
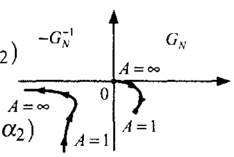
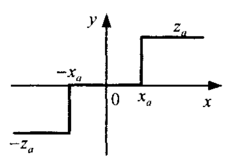
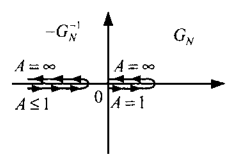
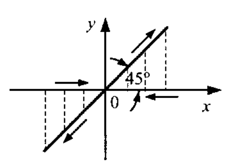
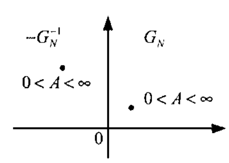
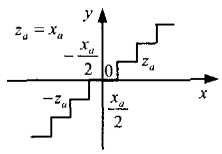
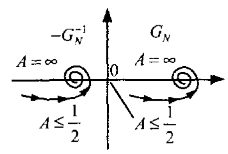
Name	Characteristic	$G_N(A)$	Diagram
Stop-type element		$P_N = \frac{1}{2} + \frac{2\alpha - \sin 2\alpha}{2\pi}$ $Q_N = \frac{\cos^2 \alpha}{\pi}$ $\sin \alpha = \frac{2z_a}{X_m} - 1$	
Dry friction		$P_N = 0$ $Q_N = \frac{4z_a}{\pi X_m}$	
Two-position relay with positive hysteresis		$P_N = \frac{4z_a}{\pi X_m} \cos \alpha$ $Q_N = \frac{4z_a}{\pi X_m} \sin \alpha$ $\sin \alpha = \frac{1}{A}$	
Two-position relay of constant width		$P_N = \frac{4z_a}{\pi X_m} \cos \alpha$ $Q_N = \frac{4z_a}{\pi X_m} \sin \alpha$ $\sin \alpha = \frac{1}{A}$	

Table A.1: Continued.

Name	Characteristic	$G_N(A)$	Diagram
Two-position relay with negative hysteresis of variable width		$\lvert x\rvert > x_a : P_N = 0$ $Q_N = \frac{4z_a}{\pi X_m}$ $\lvert x\rvert < x_a : P_N = \frac{4z_a}{\pi X_m}\cos\alpha$ $Q_N = \frac{4z_a}{\pi X_m}\sin\alpha$ $\sin\alpha = \frac{1}{A}$	
Three-position relay with positive hysteresis		$P_N = \frac{2z_a}{\pi X_m}(\cos\alpha_1 + \cos\alpha_2)$ $Q_N = -\frac{2z_a}{\pi X_m}(\sin\alpha_1 - \sin\alpha_2)$ $\sin\alpha_1 = 1/A; \sin\alpha_2 = \lambda/A$ $\lambda = x_b/x_a$	
Three-position relay without hysteresis		$P_N = \frac{4z_a}{\pi X_m}\cos\alpha$ $Q_N = 0$ $\sin\alpha = 1/A$	
Variable structure		$P_N = \frac{1}{2}$ $Q_N = \frac{1}{\pi}$	

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

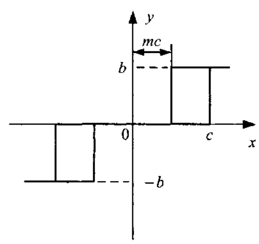
Table A.1: Continued.

Name	Characteristic	$G_N(A)$	Diagram
A/D converter		$P_N = \frac{2}{\pi A}\sum_{i=1}^{k}\sqrt{4-(1/A)^2(2i-1)^2}$ $Q_N = 0$ $k = E(A+0.5)$ E – integer number	

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

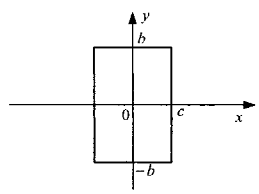
Table A.2: Coefficients of harmonic linearization of asymmetric nonlinear elements.

Nonlinear characteristic	Coefficients of harmonic linearization



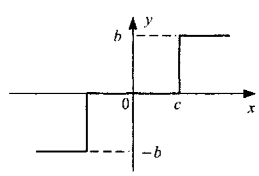
$$\begin{aligned} F^0(A,x^0) &= \tfrac{b}{2\pi}\Big(\sin^{-1}\tfrac{c+x^0}{A}-\sin^{-1}\tfrac{c-x^0}{A} \\ &\quad +\sin^{-1}\tfrac{mc+x^0}{A}-\sin^{-1}\tfrac{mc-x^0}{A}\Big) \\ P(A,x^0) &= \tfrac{b}{\pi A}\Bigg[\sqrt{1-\left(\tfrac{c+x^0}{A}\right)^2}+\sqrt{1-\left(\tfrac{c-x^0}{A}\right)^2} \\ &\quad +\sqrt{1-\left(\tfrac{mc+x^0}{A}\right)^2}+\sqrt{1-\left(\tfrac{mc-x^0}{A}\right)^2}\Bigg] \\ Q(A) &= -\tfrac{2bc}{\pi A^2}(1-m) \end{aligned}$$

$$A \geq c+|x^0|$$



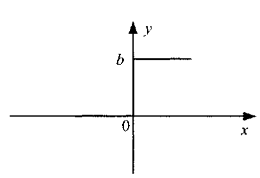
$$\begin{aligned} F^0(X_m,x^0) &= \tfrac{b}{2\pi}\left(\sin^{-1}\tfrac{c+x^0}{X_m}-\sin^{-1}\tfrac{c-x^0}{X_m}+\right) \\ P(X_m,x^0) &= \tfrac{b}{\pi X_m}\left[\sqrt{1-\left(\tfrac{c+x^0}{X_m}\right)^2}+\sqrt{1-\left(\tfrac{c-x^0}{X_m}\right)^2}\right] \\ Q(X_m) &= -\tfrac{4bc}{\pi X_m^2} \end{aligned}$$

$$A \geq c+|x^0|$$



$$\begin{aligned} F^0(X_m,x^0) &= \tfrac{b}{\pi}\left(\sin^{-1}\tfrac{c+x^0}{X_m}-\sin^{-1}\tfrac{c-x^0}{X_m}+\right) \\ P(X_m,x^0) &= \tfrac{b}{\pi A}\left[\sqrt{1-\left(\tfrac{c+x^0}{X_m}\right)^2}+\sqrt{1-\left(\tfrac{c-x^0}{X_m}\right)^2}\right] \\ Q(X_m) &= -\tfrac{4bc}{\pi X_m^2} \end{aligned}$$

$$A \geq c+|x^0|$$

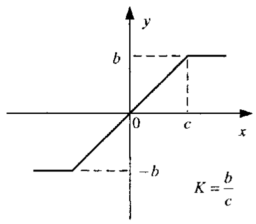
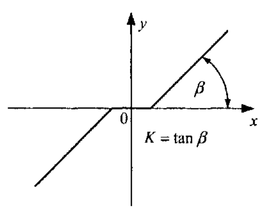
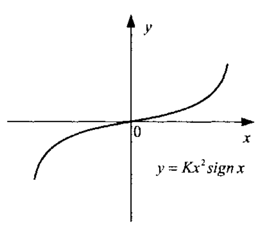


$$\begin{aligned} F^0(A,x^0) &= \tfrac{b}{2}+\tfrac{b}{\pi}\sin^{-1}\tfrac{x^0}{X_m} \\ P(A,x^0) &= \tfrac{2b}{\pi X_m}\sqrt{1-\left(\tfrac{x^0}{X_m}\right)^2} \end{aligned}$$

$$A \geq c+|x^0|$$

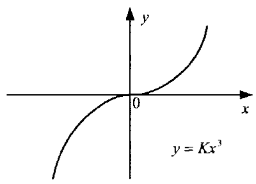
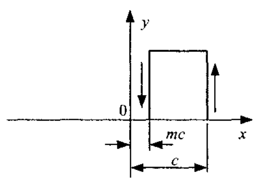
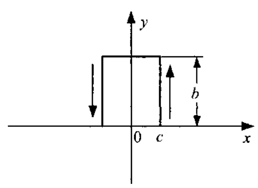
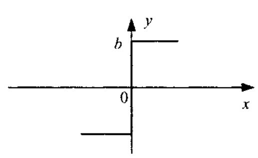
Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Table A.2: Continued.

Nonlinear characteristic	Coefficients of harmonic linearization
	$\begin{aligned} F^0(X_m,x^0) &= \frac{KX_m}{\pi}\left[\sqrt{1-\left(\frac{c+x^0}{X_m}\right)^2}+\sqrt{1-\left(\frac{c-x^0}{X_m}\right)^2}\right] \\ &+\frac{Kx^0}{\pi}\left[\sin^{-1}\frac{c+x^0}{X_m}-\sin^{-1}\frac{c-x^0}{X_m}\right] \\ &+\frac{Kb}{\pi}\left[\sin^{-1}\frac{c+x^0}{X_m}-\sin^{-1}\frac{c-x^0}{X_m}\right] \\ P(X_m,x^0) &= \frac{K}{X_m}\left[\sin^{-1}\frac{c+x^0}{X_m}+\sin^{-1}\frac{c-x^0}{X_m}\right. \\ &+\frac{c-x^0}{X_m}\sqrt{1-\left(\frac{c+x^0}{X_m}\right)^2} \\ &\left.+\frac{c+x^0}{X_m}\sqrt{1-\left(\frac{c+x^0}{X_m}\right)^2}\right] \end{aligned}$ $A \geq c+\lvert x^0\rvert$
	$\begin{aligned} F^0(X_m,x^0) &= \frac{KX_m}{\pi}\left[\sqrt{1-\left(\frac{c+x^0}{X_m}\right)^2}+\sqrt{1-\left(\frac{c-x^0}{X_m}\right)^2}+Kx^0\right] \\ &+\frac{K}{\pi}\left[C\left(\sin^{-1}\frac{c+x^0}{X_m}-\sin^{-1}\frac{c-x^0}{X_m}\right)\right. \\ &\left.-x^0\left(\sin^{-1}\frac{c+x^0}{X_m}-\sin^{-1}\frac{c-x^0}{X_m}\right)\right] \\ P(X_m,x^0) &= K-\frac{K}{\pi}\left[\sin^{-1}\frac{c+x^0}{X_m}+\sin^{-1}\frac{c-x^0}{X_m}+\right. \\ &+\frac{c-x^0}{X_m}\sqrt{1-\left(\frac{c+x^0}{X_m}\right)^2} \\ &\left.+\frac{c+x^0}{X_m}\sqrt{1-\left(\frac{c+x^0}{X_m}\right)^2}\right] \end{aligned}$ $A \geq c+\lvert x^0\rvert$
	$\begin{aligned} F^0(X_m,x^0) &= \frac{2K}{\pi}\left\{\left[\left(x^0\right)^2+\frac{A^2}{2}\right]\sin^{-1}\frac{x^0}{X_m}\right. \\ &\left.+\frac{3}{2}x^0X_m\sqrt{1-\left(\frac{x^0}{X_m}\right)^2}\right\} \\ P(X_m,x^0) &= \frac{4K}{\pi}\left\{x^0\sin^{-1}\frac{x^0}{X_m}+\right. \\ &\left.+\left[\frac{2X_m}{3}+\frac{(x^0)^2}{3X_m}\right]\sqrt{1-\left(\frac{x^0}{X_m}\right)^2}\right\} \end{aligned}$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Table A.2: Continued.

Nonlinear characteristic	Coefficients of harmonic linearization
	$F^0(X_m,x^0) = K\left[(x^0)^3 + \frac{3}{2}x^0X_m^2\right]$ $P(X_m,x^0) = 3K\left[(x^0)^3 + \frac{X_m^2}{4}\right]$
	$F^0(X_m,x^0) = \frac{b}{2} - \frac{b}{2\pi}\left[sin^{-1}\frac{c-x^0}{X_m} + sin^{-1}\frac{mc-x^0}{X_m}\right]$ $P(X_m,x^0) = \frac{b}{\pi X_m}\left[\sqrt{1-\left(\frac{c-x^0}{X_m}\right)^2} + \sqrt{1-\left(\frac{mc-x^0}{X_m}\right)^2}\right]$ $Q(X_m,x^0) = \frac{bc}{X_m^2}(1-m)$ $X_m \geq \left\|c-x^0\right\|;\ X_m \geq \left\|x^0-mc\right\|$
	$F^0(X_m,x^0) = \frac{b}{2} - \frac{b}{2\pi}\left[sin^{-1}\frac{c+x^0}{X_m} - sin^{-1}\frac{c+x^0}{X_m}\right]$ $P(X_m,x^0) = \frac{b}{\pi X_m}\left[\sqrt{1-\left(\frac{c-x^0}{X_m}\right)^2} + \sqrt{1-\left(\frac{c+x^0}{X_m}\right)^2}\right]$ $Q(X_m,x^0) = -\frac{bc}{X_m^2}(1-m)$ $X_m \geq \left\|c+x^0\right\|$
	$F^0(X_m,x^0) = \frac{2b}{\pi}\sin^{-1}\frac{x^0}{X_m}$ $P(X_m,x^0) = \frac{4b}{\pi X_m}\sqrt{1-\left(\frac{x^0}{X_m}\right)^2}$ $X_m \geq \left\|x^0\right\|$

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Table A.2: Continued.

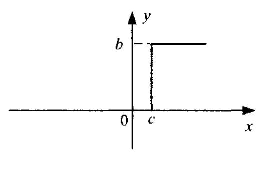
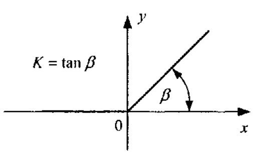
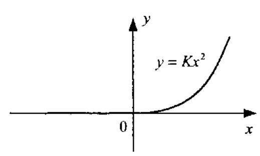
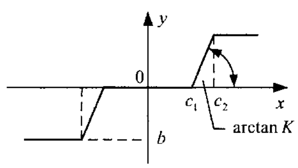
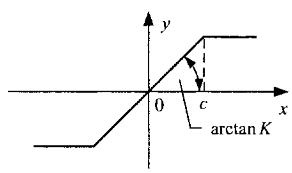
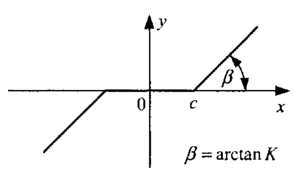
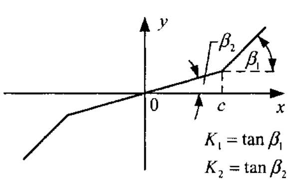
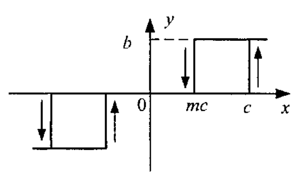
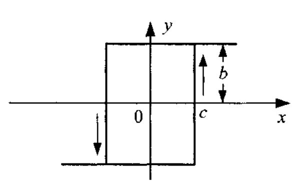
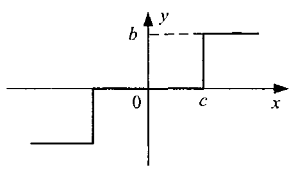
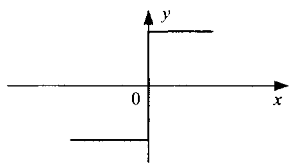
Nonlinear characteristic	Coefficients of harmonic linearization
	$F^0(X_m,x^0) = \frac{b}{2} - \frac{b}{\pi}\sin^{-1}\frac{c-x^0}{X_m}$ $P(X_m,x^0) = \frac{2b}{\pi X_m}\sqrt{1-\left(\frac{c-x^0}{X_m}\right)^2}$ $X_m \geq \lvert c-x^0 \rvert$
	$F^0(X_m,x^0) = \frac{Kx^0}{2} + \frac{K}{\pi}\left[x^0\sin^{-1}\frac{x^0}{X_m} + X_m\sqrt{1-\left(\frac{x^0}{X_m}\right)^2}\right]$ $P^0(X_m,x^0) = \frac{Kx}{2} + \frac{K}{\pi}\left[\sin^{-1}\frac{x^0}{X_m} + \frac{x^0}{X_m}\sqrt{1-\left(\frac{x^0}{X_m}\right)^2}\right]$
	$F^0(X_m,x^0) = \frac{K}{\pi}\left\{\left[(x^0)^2 + \frac{X_m^2}{2}\right]\left(\pi + \sin^{-1}\frac{x^0}{X_m}\right) + \frac{3}{2}x^0X_m\sqrt{1-\left(\frac{x^0}{X_m}\right)^2}\right\}$ $P^0(X_m,x^0) = \frac{2K}{\pi}\left\{\left[x^0 left(\frac{\pi}{2} + \sin^{-1}\frac{x^0}{X_m}\right) + \left[\frac{2X_m}{3} + \frac{(x^0)^2}{3X_m}\right]\sqrt{1-\left(\frac{x^0}{X_m}\right)^2}\right\}$

Table A.3: Linearization coefficients for standard nonlinear elements ($K_N = \Phi(x^0)/x^0$).

Nonlinear characteristic	Gain K_N, $X_m = A$
	$$K_N = \frac{2K}{\pi}\left(\sin^{-1}\frac{c_2}{X_m} - \sin^{-1}\frac{c_1}{X_m}\right)$$
	$$K_N = \frac{2K}{\pi}\sin^{-1}\frac{c}{X_m}$$
	$$K_N = K - \frac{2K}{\pi}\sin^{-1}\frac{c}{X_m}$$
	$$K_N = K_2 - \frac{2(K_2 - K_1)}{\pi}\sin^{-1}\frac{c}{X_m}$$
	$$K_N = \frac{b}{\pi X_m}\left[\frac{1}{\sqrt{1-\left(\frac{c}{X_m}\right)^2}} + \frac{1}{\sqrt{1-\left(\frac{mc}{X_m}\right)^2}}\right]$$
	$$K_N = \frac{2b}{\pi X_m}\frac{1}{\sqrt{1-\left(\frac{c}{X_m}\right)^2}}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Table A.3: Continued.

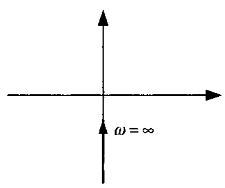
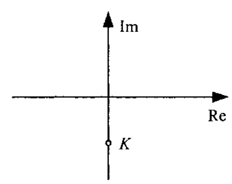
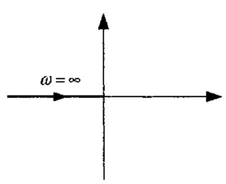
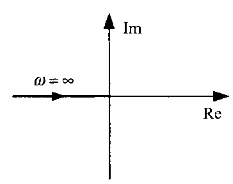
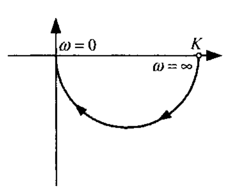
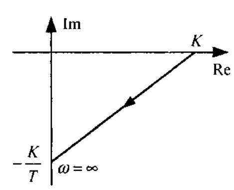
Nonlinear characteristic	Gain K_N, $X_m = A$
	$$K_N = \frac{2b}{\pi X_m} \frac{1}{\sqrt{1 - \left(\frac{c}{X_m}\right)^2}}$$
	$$K_N = \frac{2b}{\pi X_m}$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Appendix B

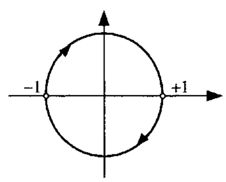
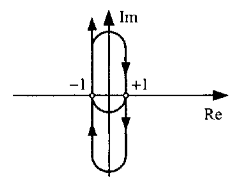
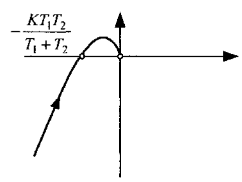
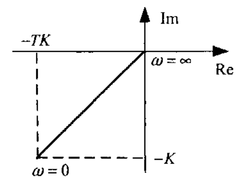
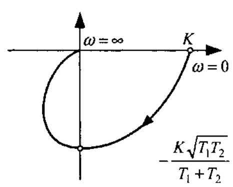
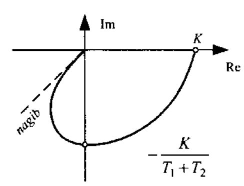
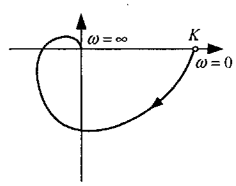
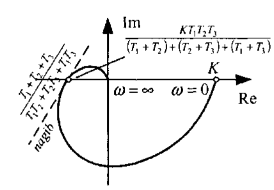
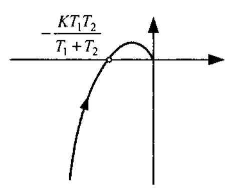
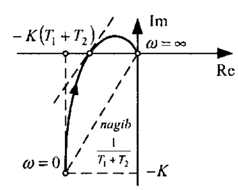
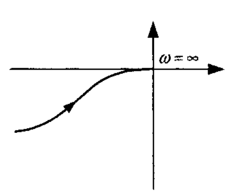
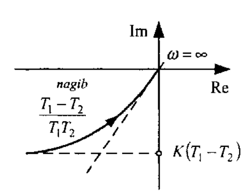
Popov Diagrams

Table B.1: Nyquist ($G(j\omega)$) and Popov ($G_P(j\omega)$) diagrams for characteristic linear part transfer functions.

Transfer function	$G(j\omega)$	$G_P(j\omega)$
$\frac{K}{s}$		
$\frac{K}{s^2}$		
$\frac{K}{Ts+1}$		

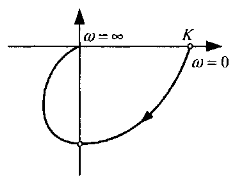
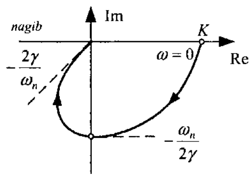
Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Table B.1: Continued.

Transfer function	$G(j\omega)$	$G_P(j\omega)$
$e^{-\tau p}$		
$\frac{K}{s(Ts+1)}$		
$\frac{K}{(T_1s+1)(T_2s+1)}$		
$\frac{K}{(T_1s+1)(T_2s+1)(T_3s+1)}$		
$\frac{K}{s(T_1s+1)(T_2s+1)}$		
$\frac{K(T_1s+1)}{s^2(T_2s+1)}$		

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Table B.1: Continued.

Transfer function	$G(j\omega)$	$G_P(j\omega)$
$\dfrac{\omega_n^2}{s^2+2\gamma\omega_n s+\omega_n^2}$		

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Bibliography

Airy, G. B. (1840). On the regulator of the clock-work for effecting uniform movement of equatorials, *Memoirs of the Royal Astronomical Society* **11**: 249–267.

Aizerman, I. (1946). About convergence of control systems following large initial disturbances, *Avtomatika & telemekhanika* (2-3): 148–167. (in Russian).

Aizerman, M. A. (1949). On a problem concerning stability in the large of dynamical systems, *Uspehi Mat. Nauk* **4**: 187–188.

Aizerman, M. A. & Gantmacher, F. R. (1964). *Absolute Stability of Regulator Systems*, Holden Day, San Francisco.

Andronov, A. A. & Bautin, N. N. (1944). Flight of the neutral airplane with the autopilot and the theory of pointwise mapping of the area., *Papers of the Academy of Science of the USSR T. XLIII, No.5*, Academy of Science of the USSR (in Russian), Moscow.

Andronov, A. A., Witt, A. A. & Haykin, S. (1959). *Theory of Oscillations*, Fizmatgiz, Moscow. (in Russian).

Banks, S. P. (1981). The circle theorem for nonlinear parabolic systems, *Int. J. Control* **24**: 843–850.

Barbalat (1959). Systemes d'equations differentielles d'oscillations non lineaires, *Revue de Mathematiques Pures et Appliquees* **4**(2): 267–270. Academie de la Republique Populaire Roumaine.

Bennett, S. (1979). *A History of Control Engineering: 1800-1930*, Peter Peregrinus, London.

Bennett, S. (1993). *A History of Control Engineering: 1930-1970*, second edn, Peter Peregrinus, Stevenage.

Besekerskii, V. A. & Popov, E. P. (1975). *Theory of Automatic Control Systems*, Nauka, Moscow. (in Russian).

Campos, J., Lewis, F. L., Davis, L. & Ikenaga, S. (2000). Backstepping based fuzzy logic control of active vehicle suspension systems, *American Control Conference*, ACC, Chicago, Illinois, pp. 4030–4035.

Chen, C. T. (1984). *Linear System Theory and Design*, Holt, Rinehart and Winston.

Chua, L. O., Komuro, M. & Matsumoto, T. (1986). The double scroll family. part i: Rigorous proof of chaos, part II: Rigorous analysis of bifurcation phenomena, *IEEE Trans. on Circuits and Systems* **CAS-33**(11): Part I: 1073–1097, Part II: 1097–1118.

Cook, P. A. (1986). *Nonlinear Dynamical Systems*, Prentice Hall, London.

Csáki, F. (1972). *Nonlinear, Optimal and Adaptive Systems*, Akadémiai Kiado, Budapest.

Desoer, C. A. (1965). A generalization of the Popov criterion, *IEEE Trans. on Automatic Control* **AC-10**: 182–184.

Desoer, C. A. & Vidyasagar, M. (1975). *Feedback Systems: Input-Output Properties.*, Academic Press, New York.

Fradkov, A. & Hill, D. (1998). Exponential feedback passivity and stabilizability of nonlinear systems, *Automatica* **34**: 697–703.

Freedman, M. I., Falb, P. L. & Zames, G. (1969). A hilbert space stability theorem, *SIAM J. Control* **3**: 479–483.

Freeman, R. A. & Kokotović, P. V. (1996). *Robust Nonlinear Control Design, State-Space and Lyapunov Techniques*, Birkhauser, Boston.

Fuller, A. T. (1976a). The early development of control theory: I, *Trans. ASME, J. Dynamic Systems Measurement and Control* **98G**(3): 109–118.

Fuller, A. T. (1976b). The early development of control theory: II, *Trans. ASME, J. Dynamic Systems Measurement and Control* **98G**(3): 224–235.

Fuller, A. T. (1992). General problem of stability of motion, *Int. J. Control* **55**: 581–773. (English translation from French of the A. M. Lypunov dissertation defended September 30th 1892 in Moscow).

Fuller, A. T. (ed.) (1975). *Stability of Motion: A Collection of Early Scientific Papers by Routh, Clifford, Sturm and Bocher*, Taylor & Francis Ltd., London.

Gantmacher, F. R. (1959). *The Theory of Matrices*, Vol. I & II, Chelsea Publishing Co.

Garcia-Cerezo, A., Ollero, A. & Aracil, J. (1992). Stability of fuzzy control systems by using nonlinear system theory, *Artificial Intelligence in Real-Time Control*, IFAC, Delft, pp. 121–126.

Gelb, A. & Vander Velde, W. E. (1968). *Multiple-Input Describing Functions and Nonlinear System Design*, McGraw-Hill, New York.

Goldfarb, L. S. (1965). *Theory of Automatic Control*, MEI, Moscow. (in Russian).

Golub, G. H. & Van Loan, G. F. (1989). *Matrix Computations*, Johns Hopkins Series in the Mathematical Sciences, The Johns Hopkins University Press, Baltimore and London.

Goodwin, G. C. & Payne, R. L. (1977). *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press.

Guckenheimer, J. & Holmes, P. J. (1983). *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Springer-Verlag, New York.

Hartley, T. T., Beale, G. O. & Chicatelli, S. P. (1994). *Digital Simulation of Dynamic Systems: A Control Theory Approach*, Prentice Hall Inc., Englewood Cliffs.

Hurwitz, A. (1895). Über die Bedienungen unter welchen eine Gleichung nur Wurzeln mit negativen realen Teilen besitzt, *Mathematische Annalen* **46**: 273–284.

Isidori, A. (1995). *Nonlinear Control Systems*, third edn, Springer-Verlag, Berlin.

Ivashchenko, N. N. (1973). *Automatic Control*, Mashinostroenie, Moscow. (in Russian).

Jacubczyk, B. & Respondek, W. (1980). On linearization of control systems, *Bull. Acad. Polonaise Sci., Ser. Sci. Math.* **28**: 517–522.

Kalman, R. E. (1957). Physical and mathematical mechanisms of instability in nonlinear automatic control systems, *Trans. ASME* **79**: 553–566.

Kalman, R. E. (1963). Lyapunov functions for the problem of Lur'e in automatic control, *Proc. Natl. Acad. Sci.* **49**: 201–205.

Karnopp, D. C., Margolis, D. L. & Rosenberg, R. C. (1990). *System Dynamics: A Unified Approach*, John Wiley.

Karnopp, D. & Rosenberg, R. (1983). *Introduction to Physical System Dynamics*, McGraw-Hill.

Khalil, H. K. (1996). *Nonlinear Systems*, Prentice Hall Inc., Englewood Cliffs.

Kharitonov, V. L. (1979). Asymptotic stability of an equilibrium position of a family of systems of linear differential equations, *Diff. Equations* **14**: 1483–1485. (Original article published in Russian in journal Diff. Uravnenia, Vol.14, 1978, pp. 2086-2088).

Kharitonov, V. L. (1994). Robust stability of nested polynomial families, *Technical Report 306*, Institute for Dynamic Systems, Univ. Bremen, Bremen.

Kharitonov, V. L. & Zhabko, A. P. (1994). Robust stability of time delay systems, *IEEE Trans. on Automatic Control* **AC-39**(12): 2388–2397.

Kiendl, H. (1993). Harmonic balance for fuzzy control systems, *EUFIT'93 - First European Congress on Fuzzy and Intelligent Technologies*, ELITE-Foundation, Aachen, pp. 137–141.

Kiendl, H. & Rüger, J. J. (1993). Verfahren zum entwurf und stabilitätsnachweis von regelungssystemen mit fuzzy-reglern, *Automatisierungstechnik* **41**: 138–144.

Kokotović, P. (1992). The joy of feedback: Nonlinear and adaptive, *IEEE Control Systems Magazine* (12): 7–17.

Kokotović, P. V. & Arcak, M. (2001). Constructive nonlinear control: a historical perspective, *Automatica* **37**(5): 637–662.

Kosko, B. (1992). *Neural Networks and Fuzzy Systems*, Prentice Hall, Englewood Cliffs, New Jersey.

Krasovskii, N. N. (1963). *Problems of the Theory of Stability of Motion*, Stanford University Press, Stanford.

Krasovskii, N. N. & Pospelov, G. S. (1962). *Fundamentals of Control Engineering and Technical Cybernetics*, Gosenergoizdat, Moscow. (in Russian).

Krstić, M. & Deng, H. (1998). *Stabilization of Nonlinear Uncertain Systems*, Springer-Verlag, New York.

Krstić, M., Kanellakopoulos, I. & Kokotović, P. (1995). *Nonlinear and Adaptive Control Design*, John Wiley & Sons, Inc., New York.

Kuljača, Lj. (1970). Theory of automatic control - nonlinear systems, University of Zagreb. (in Croatian).

Kuljača, Lj., Djonlagić, D., Vukić, Z. & Tešnjak, S. (1998). *Nonlinear Systems of Automatic Control*, FERI, Univerza v Mariboru, Maribor. (in Slovenian).

Kuljača, Lj. & Vukić, Z. (1985). *Automatic Control Systems*, Skolska knjiga, Zagreb. (in Croatian).

Kuljača, Lj., Wang, W. Y. & Vukić, Z. (1983). Dynamic analysis of the ship rudder servo system, *Brodogradnja* **31**(3): 139–152. (in Croatian).

Kuljača, O., Kuljača, Lj., Vukić, Z. & Strah, B. (2002). Fuzzy controller for elimination of the nonlinear resonance phenomenon, *Proceedings of the Mediterranean Conference on Control and Automation*, MED2002, Portugal.

Kuljača, O., Tešnjak, S. & Vukić, Z. (1999). Describing function of Mamdani type fuzzy controller with input signals derived from single system input and singleton output membership functions, *IEEE Hong Kong Symposium on Robotics and Control*, Vol. I, IEEE, Hong Kong, pp. 327–331.

La Salle, J. P. & Lefschetz, S. (1961). *Stability by Lyapunov's Direct Method with Applications*, Academic Press, New York.

Lagrange, J. L. (1788). *Mecanique Celeste*, Dunod, Paris.

Lawrence, D. A. & Rugh, W. J. (1994). Input-output pseudolinearization of nonlinear systems, *IEEE Trans. on Automatic Control* **AC-39**(11): 2207–2218.

Léauté, H. (1885). Mémoire sur les oscillations á longnes périods dans les machines actionnées per des moteurs hydrauliques, *Jornal de L'Ecole Polytechnique* **LV**(1): 1–126.

Lecornu, L. A. (1898). *Regularisation du Mouvement dans les Machines*, Villors et Fils, Paris.

Ljung, L. & Söderström, T. (1983). *Theory and Practice of Recursive Identification*, The MIT Press, Massachusetts and London.

Lur'e, A. I. (1957). Some nonlinear problems in the theory of automatic control, Her Majesty's Stationery Office, London.

Lur'e, A. I. & Postnikov, V. N. (1944). On the theory of stability of control systems, *Prikl. Mat. & Mekh.* **8**: 246–248.

Lyapunov, A. M. (1892). *The General Problem of the Stability of Motion*, PhD thesis. (In Russian. French translation in Ann. Fac. Sci., Univ.Toulouse, 9, 203-474 (1907): English transl. in Ann. Math. Studies, Princeton Univ. Press, N.J. (1949), and also in Int. J. Control (1992)).

Maciejowski, J. (1989). *Multivariable Feedback Design*, Addison-Wesley, Wokingham, U.K.

Maciejowski, J. (1999). *Predictive Control with Constraints*, Addison-Wesley, Wokingham, U.K.

Maeda, M. & Murakami, S. (1991). Stability analysis of fuzzy control systems using phase planes, *Japanese Journal of Fuzzy Theory and Systems* 3(2): 149–160.

Massera, J. L. (1949). On Liapunoff's conditions of stability, *Ann. Math.* **50**: 705–721.

Matsumoto, T. (1987). Chaos in electronic circuits, *Proceedings of the IEEE* **75**(8): 1033–1057.

Maxwell, J. C. (1868). On governors, *Proc. Royal Society London* **16**: 270–283.

Minorsky, N. (1969). *Theory of Nonlinear Control Systems*, McGraw-Hill, New York.

Mohler, R. R. (1991). *Nonlinear Systems*, Prentice Hall Inc., New Jersey, Englewood Cliffs.

Mutter, V. M. (1973). *Absolute Stability of Nonlinear Ship Control Systems*, Sudostroenie, Leningrad. (in Russian).

Nakajama, N. & Gupta, M. M. (1988). A generalized fuzzy set and its representations, *in* T. Yamava (ed.), *Fuzzy Computing, Theory, Hardware and Applications*, Elsevier, Amsterdam.

Naumov, B. N. (1972). *Theory of Nonlinear Control Systems*, Nauka, Moscow. (in Russian).

Naumov, B. N. & Tsypkin, Y. Z. (1964). A frequency criterion for absolute process stability in nonlinear automatic control, *Automation & Remote Control* **25**: 765–772.

Naylor, A. W. & Sell, G. R. (1962). *Linear Operator Theory in Engineering and Science*, Springer-Verlag, New York.

Nelepin, P. A. (1971). *Exact Methods for Nonlinear Control Systems*, Mashinostroenie, Moscow. (in Russian).

Netushil, A. (1983). *Theory of Automatic Control*, Mir, Moscow.

Nijmeijer, H. & der Schaft, A. V. (1990). *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York.

Nyquist, H. (1932). Regeneration theory, *Bell Systems Technical Journal* **11**: 126–147.

Opitz, H. P. (1993). Fuzzy-control and stability criteria, *EUFIT'93 - First European Congress on Fuzzy and Intelligent Technologies*, ELITE-Foundation, Aachen, pp. 130–136.

Pal'tov, L. P. (1975). *Nonlinear Control Systems Process Performances and Compensator Design*, Nauka, Moscow. (in Russian).

Petrov, B. N., Solodovnikov, V. V. & Topcheev, Y. I. (1967). *Modern Methods of Designing Control Systems*, Mashinostroenie, Moscow. (in Russian).

Poincaré, H. (1928). *Sur les Courbes Désiniés par une équation Différentielle*, Oeuvres T.I., Paris.

Ponomaryev, V. M. & Litvinov, A. I. (1974). *Fundamentals of Automatic Control Systems*, Vishaya Shkola, Moscow. (in Russian).

Popov, E. P. & Pal'tov, I. P. (1960). *Approximate Analysis Methods of Nonlinear Control Systems*, Fizmatgiz, Moscow. (in Russian).

Popov, V. M. (1959). Criterii suficiente de stabilitate asimptoticá in mare pentru sisteme le automate nelinaire cu mai multe organe de executie, *Studii se cercetári de energetica* **4**: 647–680.

Popov, V. M. (1960). Nouveaux critérium de stabilité pour les systémes automatiques non linéares, *Revue d'electrotehnique et d'energetique* **1**: 73–88.

Popov, V. M. (1961). About absolute stability of nonlinear control systems, *Avtomatika & telemekhanika* **8**(3): 961–979. (in Russian).

Popov, V. M. (1973). *Hyperstability of Control Systems*, Springer-Verlag, Berlin.

Proell, R. (1884). Ueber den indirektwirkenden regulierapparat patent proell, *Zeitschrift des vereines Deutsche ingenieure* **Bd. XXVIII, No. 24**.

Reid, R. E., Youhanaie, E., Blanke, M. & Thomsen, J. C. (1982). Energy losses due to steering gear installations on merchant ships: Theory and practice, *Proc. Of the SNAME Ship Costs and Energy '82 Symposium*, SNAME, New York.

Reiss, R. & Geiss, G. (1963). The construction of Lyapunov functions, *IEEE Trans. on Automatic Control* **AC-8**(4): 382–383.

Rosenbrock, H. H. (1972). Multivariable circle theorems, *Proc. 2nd IMA Conf. On Recent Mathematical Developments in Control*.

Rosenwasser, E. N. (1969). *Oscillations of Nonlinear Systems*, Nauka, Moscow. (in Russian).

Routh, E. J. (1877). *A Treatise on the Stability of a Given State of Motion*, MacMillan, London.

Sandberg, I. W. (1964). A frequency-domain condition for stability of feedback systems containing a single time-varying nonlinear element, *Bell Syst. Techn. J.* **43**: 1601–1608.

Sankovskii, E. A. (1973). *Handbook of Automatic Control Theory*, Vishaya Shkola, Moscow. (in Russian).

Schultz, D. & Gibson, J. E. (1962). The variable gradient method for generating lyapunov functions, *Trans. AIEE* **81**(11): 203–210.

Sepulchre, R., Janković, M. & Kokotović, P. V. (1997). *Constructive Nonlinear Control*, Springer-Verlag, New York.

Shcherbakov, P. S. (1992). Alexander Mikhailovitch Lyapunov: On the centenary of his doctoral dissertation on stability of motion, *Automatica* **28**(5): 865–871.

Slotine, J. J. E. & Li, W. (1991). *Applied Nonlinear Control*, Prentice Hall Inc., Englewood Cliffs.

Solodovnikov, V. V. (1969). *Technical Cybernetics*, Vol. 3, Mashinostroenie, Moscow. (in Russian).

Sontag, E. D. & Sussmann, H. J. (1988). Further comments on the stabilizability of the angular velocity of a rigid body, *Systems and Control Letters* (12): 437–442.

Stodola, A. (1893). Uber die regelung von turbinen, *Schweizerische Bauzeitung* **XXII**(17): 113–117. (See also No. 18, pp. 121-122, No. 19, pp. 126-128, No. 20, pp. 134-135).

Suevalov, L. F. (1977). *Handbook on Design of Ship Control Systems*, Sudostroenie, Leningrad. (in Russian).

Tanaka, K. & Sugeno, M. (1992). Stability analysis and design of fuzzy control systems, *Fuzzy Sets and Systems* **45**: 135–156.

Tovornik, B., Djonlagić, D. & Jurković, F. (1994). *Basics of the Design of Fuzzy Control Systems*, KoREMA, Zagreb.

Tsypkin, Y. Z. (1951). Control systems stability criteria, *in* V. V. Solodovnikov (ed.), *Theory of Automatic Control*, Mashgiz, Moscow. (in Russian).

Tsypkin, Y. Z. (1977). *Theory of Automatic Control Fundamentals*, Fizmatgiz, Moscow. (in Russian).

Vavilov, A. A. (1970). *Frequency Design Methods for Nonlinear Systems*, Energiya, Leningrad.

Vidyasagar, M. (1985). *Control Systems Synthesis: A Factorization Approach*, Vol. 8 of *North Holland System and Control Series*, MIT Press, Cambridge, MA.

Vidyasagar, M. (1993). *Nonlinear Systems Analysis*, second edn, Prentice Hall Inc., Englewood Cliffs.

Voronov, A. A. (1977). *Theory of Automatic Control*, Vishaya Shkola, Moscow. (in Russian).

Voronov, A. A. (1979). *Stability, Controllability and Observability*, Nauka, Moscow. (in Russian).

Šijak, T., Kuljača, O., Kuljača, LJ. & Tešnjak, S. (2002). Design of fuzzy regulator for power system secondary load frequency control, *Proceedings of the Mediterranean Conference on Control and Automation*, MED2002, Portugal.

Šijak, T., Kuljača, O. & Tešnjak, S. (2001). Stability analysis of fuzzy control system using describing function method, *Proceedings of the Mediterranean Conference on Control and Automation*, MED2001, Dubrovnik.

Šiljak, D. (1969). *Nonlinear Systems: Parametric Analysis and Design*, John Wiley & Sons, New York.

Willems, J. C. (1971). *The Analysis of Feedback Systems*, M.I.T. Press, Cambridge.

Wischnegradskii, I. A. (1876). Sur la theorie generale des regulateurs, *Comptes rendus hebdomodaires des sciences de l'Academie des sciences* **83**: 318–321.

Yakubovich, V. A. (1964). Solution of certain matrix inequalities encountered in nonlinear control theory, *Soviet Math. Doklady* **5**: 652–656.

Yakubovich, V. A. (1965). Matrix inequalities method in the theory of stability of nonlinear control systems, *Avtomatika & telemekhanika* **4**(4).

Zadeh, L. A. (1965). Fuzzy sets, *Information and Control* (8): 338–353.

Zames, G. (1966a). On the input-output stability of time-varying nonlinear feedback systems. part I, *IEEE Trans. on Automatic Control* **AC-11**: 228–238.

Zames, G. (1966b). On the input-output stability of time-varying nonlinear feedback systems. part II, *IEEE Trans. on Automatic Control* **AC-11**: 465–476.

Ziegler, J. G. & Nichols, N. B. (1942). Optimum settings for automatic controllers, *Trans. ASME* **64**: 759–768.

Zubov, V. I. (1961). *Mathematical Methods for the Study of Automatic Control Systems*, Academic Press, New York.

Index

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016