control systems of variable structure

U. Itkis

control systems of variable structure

U. Itkis

A HALSTED PRESS BOOK

JOHN WILEY & SONS, New York · Toronto
ISRAEL UNIVERSITIES PRESS, Jerusalem

ISRAEL UNIVERSITIES PRESS
is a publishing division of
KETER PUBLISHING HOUSE JERUSALEM LTD.
P.O.Box 7145, Jerusalem, Israel

Published in the Western Hemisphere by
HALSTED PRESS, a division of
JOHN WILEY & SONS, INC., NEW YORK

Library of Congress Cataloging in Publication Data

Itkis, U.
Control systems of variable structure.

"A Halsted Press book."
Bibliography: p.
1. Feedback control systems. 2. Control theory.
I. Title.
TJ216.184 1976 629.8′312 76–4870
ISBN 0–470–15072–6

Distributors for the U.K., Europe, Africa and the Middle East
JOHN WILEY & SONS, LTD., CHICHESTER

Distributors for Japan, Southeast Asia and India
TOPPAN COMPANY, LTD., TOKYO AND SINGAPORE

Distributed in the rest of the world by
KETER PUBLISHING HOUSE JERUSALEM LTD.
IUP cat. no. 24203 2
ISBN 0 7065 1552 8

Set, printed and bound by Keterpress Enterprises, Jerusalem
PRINTED IN ISRAEL

PREFACE

This book was written at the recommendation of Professor M. Horowitz of the Weizmann Institute, Rehovoth, Israel and the University of Colorado, U.S.A. It is based on a lecture course on the theory of automatic control systems with variable structure, or variable-structure systems (VSS), presented by the author at the Feinberg School of the Weizmann Institute. The purpose of the book is to acquaint English-speaking control theorists and engineers with the elements of the theory of VSS, which has developed over the last fifteen years almost exclusively in the USSR, its main exponents there being Professors S. Emel'yanov (Moscow) and E. Barbashin (Sverdlovsk) and their co-workers.

As evidenced by their name, variable-structure systems differ from traditional automatic control systems in that their structure is purposefully changed, in jump-wise fashion, during the transient process, depending on the current value of the error signal and its derivatives. Research has shown that VSS possess several essential advantages, among these high speed, insensitivity to variations in the plant parameters and to external disturbances, and simplicity of physical realization.

The bulk of the book is concerned with the theoretical study of VSS (first and foremost, their stability), but we also describe block-diagrams of physical devices implementing the various control laws.

Though we do not claim to provide an exhaustive account of all results achieved in the theory of VSS, the material should give a fairly complete picture of its fundamental ideas and methods. The interested reader may draw additional information from the papers and books listed in the bibliography, which is fairly exhaustive.

The author owes his original acquaintance with the theory of VSS to Professor S. Emel'yanov. He is deeply indebted to Professor L. Segal of the Weizmann Institute for his assistance in publishing this book. A. Leibovich rendered invaluable help in the preparation of the manuscript.

U. Itkis

INTRODUCTION

The main distinctive feature of VSS, setting them apart as an independent class of control systems, is that changes can occur in the structure of the system during the transient process (this is indicated in the term "VSS"). The structure of a VSS is changed intentionally, in accordance with some preassigned algorithm or law of structural change; the times at which these changes occur (and the type of structure formed) are determined not by a fixed program but in accordance with the current value of the error signal and its derivatives (this distinguishes VSS from programmed controllers).

The changes that can be introduced in the structure of the system during the transient open up wide new vistas for the control designer, primarily because by changing the structure of the system he can resolve the conflict, typical for automatic control systems, between static accuracy (stability, noise immunity) and speed of response (dynamic accuracy).

A simple example will illustrate the situation. It is well known that if the control law employs integral control the system has no steady-state error in response to a step input; besides, this improves the speed of response of the system, which is undoubtedly desirable. On the other hand, if the gain of the integrator is sufficiently high, overshoot will occur, increasing sharply as a function of the gain; this is highly undesirable. In the absence of integral control, one can (as a rule) sharply increase the gain of the closed-loop system and thereby improve the system response. However, the system will then display a steady-state error. Thus the control designer is obliged to reach some compromise, making the control (correcting effort) a linear combination:

$$u(t) = K_1 x(t) + K_2 \int_0^t x(t)\, dt, \tag{0.1}$$

where $x(t)$ is the error signal and the gains K_1, K_2 are given some "average" value which guarantees satisfactory response and not too large oscillation (overshoot) in the system. Of course, this compromise does not eliminate the conflict between the static and dynamic accuracy of the system; it merely enables the designer to "reconcile himself" to it.

On the other hand, the conflict may be resolved ("circumvented") by employing the principle of variable structure. It is intuitively clear that if the control law applied at the first stage of the transient (as long as the error is sufficiently large) is chosen as

$$u_1(t) = K'x(t) \quad \text{for} \quad |x(t)| > \varepsilon, \tag{0.2}$$

where $\varepsilon > 0$ is some constant, but at the final stage (when the error is small) the control law is

$$u_2(t) = K'' \int_{t_\varepsilon}^{t} x(t)\, dt, \tag{0.3}$$

where $|x(t)| < \varepsilon$ for $t \geqq t_\varepsilon$, then, if the parameters K', K'', ε are suitably selected, one can ensure a high-quality transient response, distinguished by good dynamic and steady-state characteristics. Indeed, taking K' sufficiently large, we make sure that the speed of the system is high. Thus the error $x(t)$ in response to a step input rapidly "enters the tube" $|x(t)| \leqq \varepsilon$. At the instant t_ε when the error has fallen to ε, the structure of the system is changed* by switching to an integral control (0.3), which eliminates the steady-state error remaining in the system (Figure 0.1).†

One could cite a host of other examples illustrating the potentials of variable structure as a means for improving the transient (see, e.g., Chapter I). But it is intuitively clear *a priori* that variable-structure systems offer far more possibilities than systems with fixed structure, if only because the latter constitute a special case (subclass) of VSS.

In principle, there are many different possible control laws governing the changes of system structure. One can switch on (or cut off) primary feedback, local feedbacks, additional derivative connections, cross couplings, and so on. Historically speaking, however, one specific type of control law from this almost unlimited supply has received most attention in the theory of VSS: laws producing in the system what is

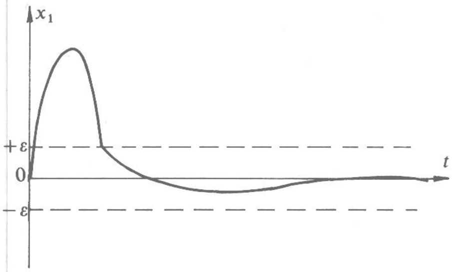

Fig. 0.1

* The inclusion of integral control action is equivalent to inserting an additional component into the system, with transfer function $1/s$, and hence to changing the structure of the system.

† This assertion can be proved analytically.

known as a sliding regime, in which the structure of the system is changed at infinite frequency.* The reason for the over-riding interest in sliding regimes lies in their many advantages, foremost among these being that motion in a sliding regime is insensitive to variations in the plant parameters and to external disturbances. The motion of a VSS in a sliding regime is equivalent to the motion of a certain new system, with a fixed structure differing from any of the structures on which the design of the original VSS is based† (it is even possible to obtain a high-quality stable equivalent structure in a VSS synthesized from several unstable structures). In most cases, organization of a sliding regime resolves the conflict between the static and dynamic accuracy of the system, for it enables one to split the transient into two independent stages: a brief motion up to the beginning of the sliding regime (known as "hitting"), characteristic of which is a high rate of decrease in the absolute value of the error, and an unlimited period of motion in the sliding regime, characterized by rapidly damped oscillations (and independence of the plant parameters and the external disturbances). It is only natural that this book too is concerned primarily with the question of whether a sliding regime can be guaranteed in a VSS and with the study of its properties, insofar as our goal is to present the main achievements of the theory as they stand today.‡

The guiding principle in our exposition is gradually to increase the complexity of the structure-control laws discussed. Thus, in Chapter I we present simple examples of second-order VSS, illustrate the increase in the quality of control achieved by varying structure, either in a sliding regime or in a low-frequency switching regime, and of course define the concept of a sliding regime in mathematical terms. In Chapter II a detailed study is made of the dynamics of VSS of arbitrary order with constant and variable plant parameters, when the control law provides for changing the sign and magnitude of the primary feedback in the system. In Chapter III we describe methods for improving the quality of control in the systems investigated in Chapter II by switching derivative connections in addition to primary feedback. Chapter IV presents control laws which implement simultaneous switching in three types of coupling: primary feedback, derivative connections and local feedbacks in the actuator; because of this, these laws produce a transient which is not only insensitive to changes in the plant parameters (for plants of arbitrary order), but is also invariant under external disturbances. Chapter VI describes efficient laws for control of cross couplings in multiconnection systems and centralized-control systems, and also

* Theoretically, of course; in practice, the frequency must be made sufficiently high, depending directly on the response characteristics of the switching devices.

† The type and parameters of this new equivalent structure may be chosen by suitable selection of the control algorithm.

‡ The present-day theory of VSS displays a tendency to intensify research into "ordinary" regimes, in which the frequency of structural changes is fairly low.

for control of connections of general (arbitrary) type in variable-structure multivariable systems. The seventh and last chapter considers the organization of an optimal transient process in a VSS by adaptation (self-adjustment) of the controller parameters, based on incoming information about the sequence (order) and frequency of structural switchings in the system.

In regard to the dynamics of all the VSS considered in Chapters I–IV and VI, VII, we shall assume that the controller is ideal and not subject to noise, so that there may exist in the system an ideal sliding regime with infinite frequency of structural switchings. In a real VSS an ideal sliding regime cannot exist. Nevertheless, it will be shown in Chapter V that, despite imperfections (such as hysteresis, insensitive zones, time lag, gain instability, pulse-amplitude and pulse-width modulation of the control signal) and noise in real VSS, there exists a high-frequency switching regime (called a quasi-sliding regime) whose properties approximate those of an ideal sliding regime, this approximation being better, the smaller are the imperfections and the lower the noise level.

CONTENTS

Chapter One
INCREASE IN CONTROL PERFORMANCE DUE TO VARIABLE STRUCTURE

1.1. EXAMPLES OF SYNTHESIS OF FAST-RESPONSE STABLE VSS BASED ON UNSTABLE STRUCTURES

Automatic control systems of variable structure (VSS = variable-structure systems) constitute a special class of nonlinear control systems. As evidenced by their name, these systems differ from other control systems mainly in that their structure is not constant but is varied during the control process. Unfortunately, it is at present impossible to give a rigorous definition of the concept "VSS," not only because the theory is still comparatively young but also because there is no really rigorous definition of "structure." Nevertheless, there seems to be unanimous agreement that the structure of a system with positive feedback is not the same as that of a system with negative feedback (even when the components of the systems are the same, as are all other couplings). Similarly, the structure of a system with negative feedback is different from that of an open-loop control system.

We shall therefore assume throughout this book that two systems which differ in the sign of the coupling between at least two of their elements, or in which there are two elements connected in one system and unconnected in the other, have different structure. Later we shall make the term "different structure" more specific.

Generally speaking, the structure of a control system may change either accidentally (for example, owing to sudden breakdown of the actuating mechanism the feedback loop may be broken) or regularly, in conformance with a definite rule. The VSS considered in this chapter fall into the category of systems with changeable structure—systems whose structure is intentionally changed during the transient in accordance with a preset structure-control law. The control laws developed in the theory of VSS usually provide for changes in the structure of the system whenever the representative point crosses certain surfaces (hypersurfaces) in the phase space of the system. The form of these surfaces depends essentially on the type of plant. Essentially, the theory of VSS is the theory selecting rational (for a specific plant) switching surfaces and structures in the regions of the phase space that they define. It is only natural that a systematic exposition of the theory should begin with the simplest case: second-order phase space, i.e., a phase space adequate for describing the dynamics of second-order systems.

We briefly recall the fundamental types of phase portraits of second-order systems, limiting ourselves to linear systems with constant parameters, whose free motion is described by an ordinary differential equation:

$$\ddot{x} + a_2\dot{x} + a_1 x = 0, \tag{1.1.1}$$

where a_2 is the damping coefficient and a_1 the gain of the error signal x. There are several possible phase-portrait types for equation (1.1.1), depending on the relative positions of the roots of the characteristic equation:

$$p^2 + a_2 p + a_1 = 0. \tag{1.1.2}$$

The possible phase portraits are as follows:

1. Degenerate case: one of the roots of the characteristic equation, say λ_1, is zero, the other (λ_2) negative (it is clear that in this case $a_1 = 0$, $a_2 > 0$). The phase trajectories (Figure 1.1.1a) are straight lines with slope $-a_2$, and there is a unique line through the origin along which the motion of the representative point is asymptotically stable. If the representative point is not on this line at the initial point of time, the system will display a steady-state error.

2. $\lambda_1 = 0$, $\lambda_2 > 0$. The phase trajectories (Figure 1.1.1b) are again straight lines, but of positive slope. The system is unstable.

3. The roots are real, negative and distinct, say $\lambda_1 < \lambda_2 < 0$. The phase trajectories (Figure 1.1.1c) are parabolas, with two asymptotes: $\sigma_1 = \dot{x} + \lambda_1 x = 0$, $\sigma_2 = \dot{x} + \lambda_2 x = 0$. The system is globally asymptotically stable (whatever the initial position of the representative point in the phase plane).

4. The roots are real, negative and equal, $\lambda_1 = \lambda_2 = \lambda < 0$. The phase trajectories (Figure 1.1.1d) are again convergent parabolas, but there is only one asymptote, $\sigma = \dot{x} + \lambda x = 0$.

5. The roots are complex, with negative real parts: $\operatorname{Re}\lambda_1, \operatorname{Re}\lambda_2 < 0$. The phase trajectory (Figure 1.1.1e) is a spiral. The system is globally asymptotically stable (oscillatory).

6. The roots are pure imaginary, $\operatorname{Re}\lambda_1 = \operatorname{Re}\lambda_2 = 0 = a_2 < a_1$. The phase trajectories (Figure 1.1.1f) are ellipses. If $a_1 > 1$, the ellipses are elongated along the vertical axis (Figure 1.1.1f), if $a_1 < 1$—along the horizontal axis (Figure 1.1.1g); if $a_1 = 1$, the phase trajectories are circles (Figure 1.1.1h). The system is conservative.

7. The roots are complex, with positive real parts: $\operatorname{Re}\lambda_1 > 0 < \operatorname{Re}\lambda_2$. The phase trajectories (Figure 1.1.1i) are expanding spirals. The system is oscillatorily unstable.

8. The roots are real, positive and distinct: $\lambda_1 > \lambda_2 > 0$. The trajectories are parabolas (Figure 1.1.1j). There are two asymptotes, $\dot{x} + \lambda_1 x = 0$, $\dot{x} + \lambda_2 x = 0$. The system is aperiodically unstable.

9. The roots are real, positive and equal: $\lambda_1 = \lambda_2 > 0$. Same as the preceding case, except that there is only one asymptote (Figure 1.1.1k).

10. The roots are real and of unlike signs, $\lambda_1 < 0 < \lambda_2$. The phase trajectories are hyperbolas (Figure 1.1.1l) with two asymptotes, $\sigma_1 = \dot{x} + \lambda_1 x = 0$, $\sigma_2 = \dot{x} +$

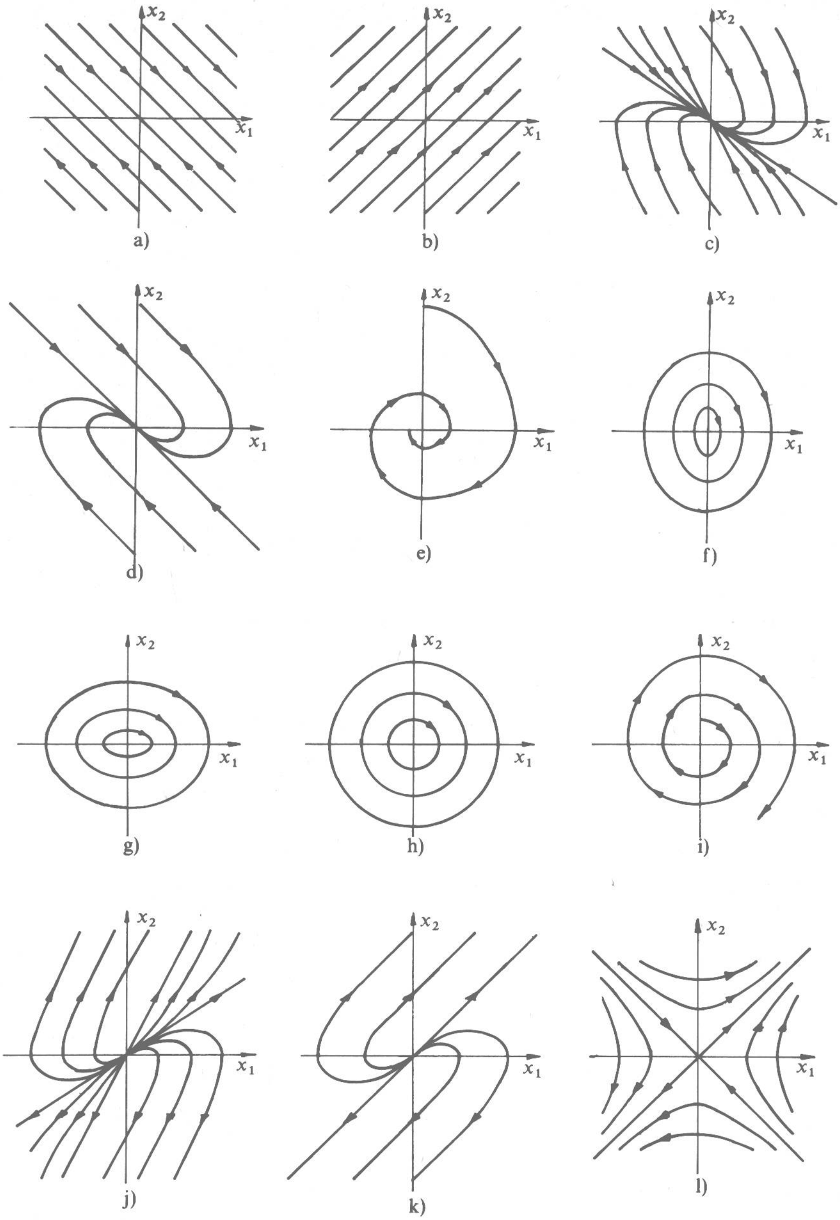

Fig. 1.1.1

$\lambda_2 x = 0$. The slopes are different, depending on the relationship between λ_1 and λ_2. The system is unstable, but there is a unique phase trajectory ($\sigma_1 = \dot{x} + \lambda_1 x = 0$) along which the representative point moves asymptotically to the origin. This distinctive existence of a single stable trajectory should be emphasized; we shall use it essentially in various VSS.

Each of the above portraits corresponds to a certain system structure. It is generally agreed that the elliptic structure of Figure 1.1.1f and the hyperbolic structure of Figure 1.1.1l are distinct. However, opinions differ as to whether elliptic structures in which the ellipses are "compressed" to different degrees are to be considered distinct. The same question may be posed for hyperbolic structures with differently situated asymptotes, for parabolic structures, etc. We shall see below that as far as the theory of VSS is concerned it is often convenient to stipulate different structures for "elliptic" systems whose ellipses have different eccentricities; one reason for this is that by using the difference between the eccentricities of the ellipses one can significantly improve the quality of the transient by stepwise changes in the coefficient a_1 (i.e., in the structure of the system).

Variable-structure systems offer the control designer new possibilities for improving the quality of control in comparison with fixed-structure systems, since the wider range of control actions implies a larger range of admissible transient processes in the system. In fact, VSS may have transients which are quite unrealizable in fixed-structure systems. This includes the possibility of synthesizing high-quality stable VSS which combine unstable structures in a certain way.*

As an illustration, we consider a VSS controlling a conservative plant by switching the sign of the feedback (Letov, 1957).† The unchangeable part of the system consists of two integrators connected in series. If the plant is part of a stabilizing negative feedback loop (Figure 1.1.2a), the resulting control system will be conservative, and its free motion is described by a second-order system of differential equations:

$$\begin{cases} \dfrac{dx_1}{dt} = x_2, \\[2ex] \dfrac{dx_2}{dt} = -a_1 x_1, \end{cases} \tag{1.1.3}$$

where x_1 is the difference between the reference input r and the output signal φ of the system; $a_1 > 0$ is the gain (this system clearly corresponds to the differential equation (1.1.1) with $a_2 = 0$, $x_1 = x$, so that the structure of system (1.1.3) is elliptic; see Figure 1.1.2b). However, if the plant is included in a positive feedback loop

* There is an analogy here with reliability theory—synthesis of reliable systems from unreliable elements.

† Letov does not use the term VSS, but "conditionally stable optimal system"; essentially, however, this is a typical representative of the class of VSS.

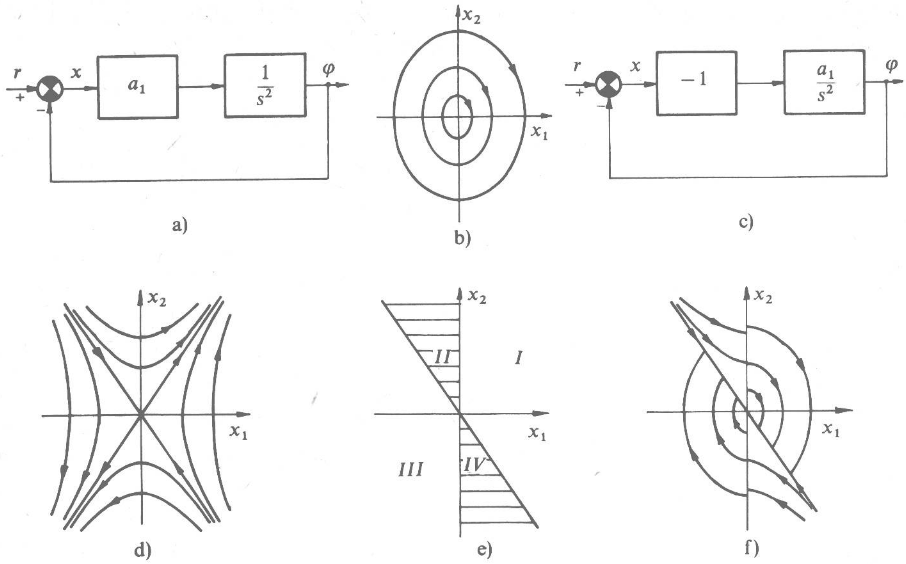

Figure 1.1.2

(Figure 1.1.2c), the system is aperiodically unstable and is described by the equations

$$\begin{cases} \dfrac{dx_1}{dt} = x_2, \\ \dfrac{dx_2}{dt} = +a_1 x_1 \end{cases} \tag{1.1.4}$$

(hyperbolic structure: Figure 1.1.2d). Neither of systems (1.1.3) or (1.1.4) is satisfactory as far as the quality of the transient is concerned (if only for the fact that they are unstable). Nevertheless, certain parts of the phase trajectories of both systems are quite satisfactory. For example, the error of system (1.1.3) decreases rapidly in the first quadrant of the phase plane (the region in which $x_1, x_2 > 0$), and system (1.1.4) has a "good" phase trajectory in the fourth quadrant (counting anticlockwise)—the asymptote of the family of hyperbolas. Let us try to combine the advantages of both systems (at the same time eliminating their shortcomings) by suitable choice of their structure in the appropriate parts of the phase plane.

Divide the phase plane $\mathscr{X}$ into four pairwise symmetric subregions:

$$\begin{aligned} &\text{Region I:} && x_1 \geqq 0, \quad x_2 + \sqrt{a_1} x_1 > 0, \\ &\text{Region II:} && x_1 < 0, \quad x_2 + \sqrt{a_1} x_1 \geqq 0, \\ &\text{Region III:} && x_1 \leqq 0, \quad x_2 + \sqrt{a_1} x_1 < 0, \\ &\text{Region IV:} && x_1 > 0, \quad x_2 + \sqrt{a_1} x_1 \leqq 0. \end{aligned} \tag{1.1.5}$$

These regions are separated from one another by the straight lines $x_1 = 0$ and $x_2 + \sqrt{a_1}x_1 = 0$ (Figure 1.1.2e). Suppose we can stipulate the structure of the system (more precisely, exactly one of two possible structures: elliptic or hyperbolic) at our discretion, at each point of the phase space.* We should then proceed as follows. At each point of region I, we stipulate that the system include a negative feedback; more precisely, we introduce negative feedback at the instant the representative point (RP) enters this region, crossing the axis $x_1 = 0$, and do not change it until the RP leaves the region, crossing the line $x_2 + \sqrt{a_1}x_1 = 0$. At each point of region IV we introduce positive feedback. The structure in regions II and III is chosen by symmetry considerations: hyperbolic in region II, elliptic in region III.

Suppose that the RP is in region I at the initial time $t_0 = 0$; thus a negative feedback is switched on at time t_0 and the RP will move along an ellipse in the clockwise direction (Figure 1.1.2f). At a certain time $t_1 > t_0$ the RP will "hit" the line $x_2 + \sqrt{a_1}x_1 = 0$. At this time the RP is in region IV, so that the positive feedback must come into action. Consequently, at time t_1 the RP begins to move along a hyperbolic trajectory; however, since its position at time t_1 is on the asymptote of the family of hyperbolas (the line $x_2 + \sqrt{a_2}x_1 = 0$), it must move along this asymptote. While it is doing so the structure of the system cannot change, since the RP remains throughout in region IV. Thus the RP will continue indefinitely to move along the straight line $x_2 + \sqrt{a_1}x_1 = 0$, approaching the origin asymptotically. In this case, therefore, we have $\lim_{t\to\infty} x_1(t) = \lim_{t\to\infty} x_2(t) = 0$.

Now consider the case that the RP is in region IV at time $t_0 = 0$. In accordance with the structural rule stipulated above, positive feedback must be switched in. Hence the RP will move along an arc of a hyperbola, crossing the ordinate axis at some time $t_2 > t_0$. At this time t_2 negative feedback is switched in, and from then on the process is similar to that considered above. Thus, in this case too we have $\lim_{t\to\infty} x_1(t) = \lim_{t\to\infty} x_2(t) = 0$, but the transient is now somewhat less satisfactory in quality, since we have a case of overshoot (the sign of the error changes); true, this occurs only once.

It is clear that if the initial position of the RP is in either of regions II or III the behavior of the transient is completely symmetric.

Thus, the VSS under consideration is globally asymptotically stable and the transient is either aperiodic or involves at most one overshoot—this for a system synthesized from two unstable systems! Of course, the transient just described is rather idealized, since in a real system unavoidable fluctuations may cause the RP to move off the asymptote $x_2 + \sqrt{a_1}x_1 = 0$, but it will subsequently move along a trajectory sufficiently close to the asymptote (the question of whether a given VSS is realizable will be studied in greater detail in the next section).

* We shall show later that this may indeed be done, using a relatively simple controller.

The next example will again demonstrate the synthesis of stable VSS from unstable structures. We again consider a conservative plant (two integrators connected in series, as part of a negative feedback loop), whose motion is described by system (1.1.3) with gain $a_1 = a_1' < 1$.

For control of the plant we use one of the simplest types of variable-structure controllers—the half-proportional controller of V. Ferner (1956), which operates on the following principle: the control is proportional to the error signal if the latter has the same sign as its rate of change (i.e., sign $\dot{x}$ = sign x); otherwise the control is zero. The control law implemented by this controller is:

$$u = \begin{cases} kx & \text{if} \quad \dot{x}x \geqq 0, \\ 0 & \text{if} \quad \dot{x}x < 0. \end{cases} \tag{1.1.6}$$

This law clearly divides the phase space into four regions, separated by the coordinate axes (Figure 1.1.3a). In the regions $\dot{x}, x \geqq 0$ (region I) and $\dot{x}, x \leqq 0$ (region III) the control signal is $u = kx$, and the free motion of the VSS is described by the system

$$\begin{cases} \dfrac{dx_1}{dt} = x_2, \\ \dfrac{dx_2}{dt} = -(a_1' + k)\,x_1, \end{cases} \tag{1.1.7}$$

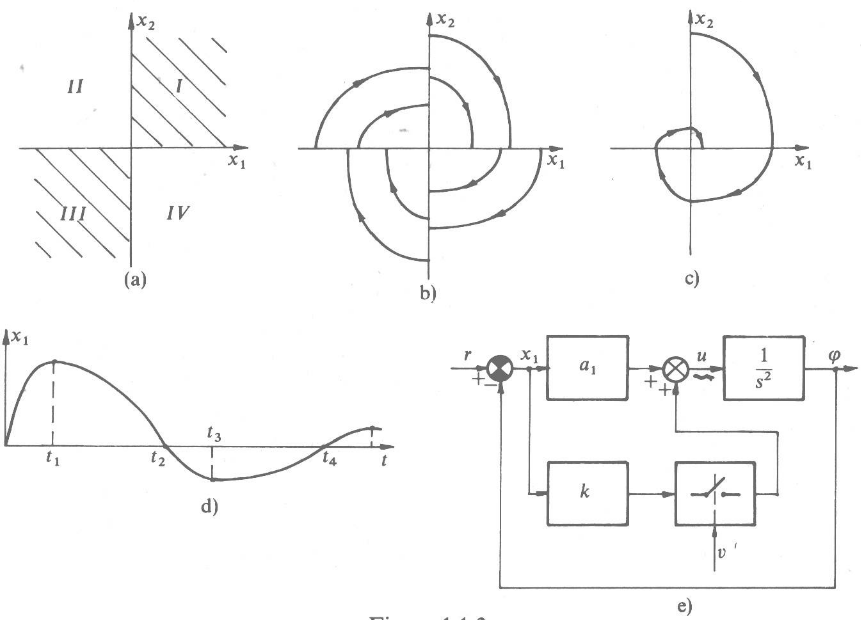

Figure 1.1.3