Rafail Khasminskii

Stochastic Stability of Differential Equations

Second Edition

Springer

Rafail Khasminskii

Stochastic Stability of Differential Equations

With contributions by G.N. Milstein and M.B. Nevelson

Completely Revised and Enlarged 2nd Edition

Springer

Rafail Khasminskii
Mathematics Department
1150 Faculty/Administration Building
Wayne State University
W. Kirby 656
Detroit, MI 48202
USA
rafail@math.wayne.edu
and
Institute for Information Transmission
Problems
Russian Academy of Scienses
Bolshoi Karetny per. 19, Moscow
Russia

Managing Editors

Boris Rozovskiĭ
Division of Applied Mathematics
Brown University
182 George St
Providence, RI 02912
USA
rozovsky@dam.brown.edu

Peter W. Glynn
Institute of Computational
and Mathematical Engineering
Stanford University
Via Ortega 475
Stanford, CA 94305-4042
USA
glynn@stanford.edu

Cover design: deblik

Preface to the Second Edition

After the publication of the first edition of this book, stochastic stability of differential equations has become a very popular theme of recent research in mathematics and its applications. It is enough to mention the Lecture Notes in Mathematics, Nos 294, 1186 and 1486, devoted to the stability of stochastic dynamical systems and Lyapunov Exponents, the books of L. Arnold [3], A. Borovkov [35], S. Meyn and R. Tweedie [196], among many others.

Nevertheless I think that this book is still useful for those researchers who would like to learn this subject, to start their research in this area or to study properties of concrete mechanical systems subjected to random perturbations. In particular, the method of Lyapunov functions for the analysis of qualitative behavior of stochastic differential equations (SDEs), the exact formulas for the Lyapunov exponent for linear SDEs, which are presented in this book, provide some very powerful instruments in the study of stability properties for concrete stochastic dynamical systems, conditions of existence the stationary solutions of SDEs and related problems.

The study of exponential stability of the moments (see Sects. 5.7, 6.3, 6.4 here) makes natural the consideration of certain properties of the moment Lyapunov exponents. This very important concept was first proposed by S. Molchanov [204], and was later studied in detail by L. Arnold, E. Oeljeklaus, E. Pardoux [8], P. Baxendale [19] and many other researchers (see, e.g., [136]).

Another important characteristic for stability (or instability) of the stochastic systems is the stability index, studied by Arnold, Baxendale and the author. For the reader's convenience I decided to include the main results on the moment Lyapunov exponents and the stability index in the Appendix B to this edition. The Appendix B was mainly written by G. Milstein, who is an accomplished researcher in this area. I thank him whole-heartily for his generous help and support.

I have many thanks to the Institute for the Problems Information Transmission, Russian Academy of Sciences, and to the Wayne State University, Detroit, for their support during my work on this edition. I also have many thanks to B.A. Amosov for his essential help in the preparation of this edition.

In conclusion I will enumerate some other changes in this edition.

1. Derivation of the often used in the book Feynman–Kac formula is added to Sect. 3.6.
2. A much improved version of Theorem 4.6 is proven in Chap. 4.
3. The Arcsine Law and its generalization are added in 4.12.
4. Sect. A.4 in the Appendix B to the first edition is shortened.
5. New books and papers, related to the content of this book, added to the bibliography.
6. Some footnotes are added and misprints are corrected.

Moscow   Rafail Khasminskii
March 2011

Preface to the First English Edition

I am very pleased to witness the printing of an English edition of this book by Noordhoff International Publishing. Since the date of the first Russian edition in 1969 there have appeared no less than two specialist texts devoted at least partly to the problems dealt with in the present book [38, 211]. There have also appeared a large number of research papers on our subject. Also worth mentioning is the monograph of Sagirov [243] containing applications of some of the results of this book to cosmology.

In the hope of bringing the book somewhat more up to date we have written, jointly with M.B. Nevelson, an Appendix A containing an exposition of recent results. Also, we have in some places improved the original text of the book and have made some corrections. Among these changes, the following two are especially worth mentioning: A new version of Sect. 8.4, generalizing and simplifying the previous exposition, and a new presentation of Theorem 7.8.

Finally, there have been added about thirty new titles to the list of references. In connection with this we would like to mention the following. In the first Russian edition we tried to give as complete as possible a list of references to works concerning the subject. This list was up to date in 1967. Since then the annual output of publications on stability of stochastic systems has increased so considerably that the task of supplying this book with a totally up to date and complete bibliography became very difficult indeed. Therefore we have chosen to limit ourselves to listing only those titles which pertain directly to the contents of this book. We have also mentioned some more recent papers which were published in Russian, assuming that those will be less known to the western reader.

I would like to conclude this preface by expressing my gratitude to M.B. Nevelson for his help in the preparation of this new edition of the book.

Moscow   Rafail Khasminskii
September 1979

Preface to the Russian Edition

This monograph is devoted to the study of the qualitative theory of differential equations with random right-hand side. More specifically, we shall consider here problems concerning the behavior of solutions of systems of ordinary differential equations whose right-hand sides involve stochastic processes. Among these the following questions will receive most of our attention.

1. When is each solution of the system defined with probability 1 for all $t > 0$ (i.e., the solution does not "escape to infinity" in a finite time)?
2. If the function $X(t) \equiv 0$ is a solution of the system, under which conditions is this solution stable in some stochastic sense?
3. Which systems admit only bounded for all $t > 0$ (again in some stochastic sense) solutions?
4. If the right-hand side of the system is a stationary (or periodic) stochastic process, under which additional assumptions does the system have a stationary (periodic) solution?
5. If the system has a stationary (or periodic) solution, under which circumstances will every other solution converge to it?

The above problems are also meaningful (and motivated by practical interest) for deterministic systems of differential equations. In that case, they received detailed attention in [154, 155, 178, 188, 191, 228], and others.

Problems 3–5 have been thoroughly investigated for linear systems of the type $\dot{x} = Ax + \xi(t)$, where A is a constant or time dependent matrix and $\xi(t)$ a stochastic process. For that case one can obtain not only qualitative but also quantitative results (i.e., the moment, correlation and spectral characteristics of the output process $x(t)$) in terms of the corresponding characteristics of the input process $\xi(t)$. Methods leading to this end are presented e.g., in [177, 233], etc. In view of this, we shall concentrate our attention in the present volume primarily on non-linear systems, and on linear systems whose parameters (the elements of the matrix A) are subjected to random perturbations.

In his celebrated memoir Lyapunov [188] applied his method of auxiliary functions (Lyapunov functions) to the study of stability. His method proved later to be

applicable also to many other problems in the qualitative theory of differential equations. Also in this book we shall utilize an appropriate modification of the method of Lyapunov functions when discussing the solutions to the above mentioned problems.

In Chaps. 1 and 2 we shall study problems 1–5 without making any specific assumptions on the form of the stochastic process on the right side of the special equation. We shall be predominantly concerned with systems of the type $\dot{x} = F(x, t) + \sigma(x, t)\xi(t)$ in Euclidean l-space. We shall discuss their solutions, using the Lyapunov functions of the truncated system $\dot{x} = F(x, t)$. In this we shall try to impose as few restrictions as possible on the stochastic process $\xi(t)$; e.g., we may require only that the expectation of $|\xi(t)|$ be bounded. It seems convenient to take this approach, first, because sophisticated methods are available for constructing Lyapunov functions for deterministic systems, and second, because the results so obtained will be applicable also when the properties of the process $\xi(t)$ are not completely known, as is often the case.

Evidently, to obtain more detailed results, we shall have to restrict the class of stochastic processes $\xi(t)$ that may appear on the right side of the equation. Thus in Chaps. 3 through 7 we shall study the solutions of the equation $\dot{x} = F(x, t) + \sigma(x, t)\xi(t)$ where $\xi(t)$ is a white noise, i.e. a Gaussian process such that $\mathbf{E}\xi(t) = 0$, $\mathbf{E}[\xi(s)\xi(t)] = \delta(t - s)$. We have chosen this process, because:

1. In many real situations physical noise can be well approximated by white noise.
2. Even under conditions different from white noise, but when the noise acting upon the system has a finite memory interval τ (i.e., the values of the noise at times t_1 and t_2 such that $|t_2 - t_1| > \tau$ are virtually independent), it is often possible after changing the time scale to find an approximating system, perturbed by the white noise.
3. When solutions of an equation are sought in the form of a process, continuous in time and without after-effects, the assumption that the noise in the system is "white" is essential. The investigation is facilitated by the existence of a well developed theory of processes without after-effects (Markov processes).

Shortly after the publication of Kolmogorov's paper [144], which laid the foundations for the modern analytical theory of Markov processes, Andronov, Pontryagin and Vitt [229] pointed out that actual noise in dynamic systems can be replaced by white noise, thus showing that the theory of Markov processes is a convenient tool for the study of such systems.

Certain difficulties in the investigation of the equation $\dot{x} = F(x, t) + \sigma(x, t)\xi(t)$, where $\xi(t)$ is white noise are caused by the fact that, strictly speaking, "white" noise processes do not exist; other difficulties arise because of the many ways of interpreting the equation itself. These difficulties have been largely overcome by the efforts of Bernshtein, Gikhman and Itô. In Chap. 3 we shall state without proof a theorem on the existence and uniqueness of the Markov process determined by an equation with the white noise. We shall assume a certain interpretation of this equation. For a detailed proof we refer the reader to [56, 64, 92].

However, we shall consider in Chap. 3 various other issues in great detail, such as sufficient conditions for a sample path of the process not to "escape to infinity"

in a finite time, or to reach a given bounded region with probability 1. It turns out that such conditions are often conveniently formulated in terms of certain auxiliary functions analogous to Lyapunov functions. Instead of the Lyapunov operator (the derivative along the path) one uses the infinitesimal generator of the corresponding Markov process.

In Chap. 4 we examine conditions under which a solution of a differential equation where $\xi(t)$ is white noise, converges to a stationary process. We show how this is related to the ergodic theory of dynamic systems and to the problem of stabilization of the solution of a Cauchy problem for partial differential equations of parabolic type.

Chapters 5–8 I contain the elements of stability theory of stochastic systems without after-effects. This theory has been created in the last few years for the purpose of studying the stabilization of controlled motion in systems perturbed by random noise. Its origins date from the 1960 paper by Kac and Krasovskii [111] which has stimulated considerable further research. More specifically, in Chap. 5 we generalize the theorems of Lyapunov's second method; Chapter 6 is devoted to a detailed investigation of linear systems, and in Chap. 7 we prove theorems on stability and instability in the first approximation. We do this, keeping in view applications to stochastic approximation and certain other problems.

Chapter 8 is devoted to application of the results of Chaps. 5 to 7 to optimal stabilization of controlled systems. It was written by the author in collaboration with M.B. Nevelson. In preparing this chapter we have been influenced by Krasovskii's excellent Appendix IV in [191].

As far as we know, there exists only one other monograph on stochastic stability. It was published in the U.S.A. in 1967 by Kushner [168], and its translation into Russian is now ready for print. Kushner's book contains many interesting theorems and examples. They overlap partly with the results of Sect. 3.7 and Sects. 5.1–5.5 of this book.

Though our presentation of the material is abstract, the reader who is primarily interested in applications should bear in mind that many of the results admit a directly "technical" interpretation. For example, problem 4, stated above, concerning the question of the existence of a stationary solution, is equivalent to the problem of determining when stationary operating conditions can prevail within a given, generally non-linear, automatic control system, whose parameters experience random perturbations and whose input process is also stochastic. Similarly, the convergence of each solution to a stationary solution (see Chap. 4) means that each output process of the system will ultimately "settle down" to stationary conditions.

In order not to deviate from the main purpose of the book, we shall present without proof many facts from analysis and from the general theory of stochastic process. However, in all such cases we shall mention either in the text or in a footnote where the proof can be found. For the reader's convenience, such references will usually be not to the original papers but rather to more accessible textbooks and monographs. On the other hand, in the rather narrow range of the actual subject matter we have tried to give precise references to the original research. Most of the references appear in footnotes.

Part of the book is devoted to the theory of stability of solutions of stochastic equations (Sects. 1.5–1.8, Chaps. 5–8). This appears to be an important subject which has recently been receiving growing attention. The volume of the relevant literature is increasing steadily. Unfortunately, in this area various authors have published results overlapping significantly with those of others. This is apparently due to the fact that the field is being studied by mathematicians, physicists, and engineers, and each of these groups publishes in journals not read by the others. Therefore the bibliography given at the end of this book lists, besides the books and papers cited in the text, various other publications on the stability of stochastic systems known to the author, which appeared prior to 1967. For the reason given above, this list is far from complete, and the author wishes to apologize to authors whose research he might have overlooked.

The book is intended for mathematicians and physicists. It may be of particular interest to those who specialize in mechanics, in particular in the applications of the theory of stochastic processes to problems in oscillation theory, automatic control and related fields. Certain sections may appeal to specialists in the theory of stochastic processes and differential equations. The author hopes that the book will also be of use to specialized engineers interested in the theoretical aspects of the effect of random noise on the operation of mechanical and radio-engineering systems and in problems relating to the control of systems perturbed by random noise.

To study the first two chapters it is sufficient to have an acquaintance with the elements of the theory of differential equations and probability theory, to the extent generally given in higher technical schools (the requisite material from the theory of stochastic processes is given in the text without proofs).

The heaviest mathematical demands on the reader are made in Chaps. 3 and 4. To read them, he will need an acquaintance with the elements of the theory of Markov processes to the extent given, e.g., in Chap. VIII of [92].

The reader interested only in the stability of stochastic systems might proceed directly from Chap. 2 to Chaps. 5–7, familiarizing himself with the results of Chaps. 3 and 4 as the need arises.

The origin of this monograph dates back to some fruitful conversations which the author had with N.N. Krasovskii. In the subsequent research, here described, the author has used the remarks and advice offered by his teachers A.N. Kolmogorov and E.B. Dynkin, to whom he is deeply indebted.

This book also owes much to the efforts of its editor, M.B. Nevelson, who not only took part in writing Chap. 8 and indicated several possible improvements, but also placed some of his yet unpublished examples at the author's disposal. I am grateful to him for this assistance. I also would like to thank V.N. Tutubalin, V.B. Kolmanovskii and A.S. Holevo for many critical remarks, and to R.N. Stepanova for her work on the preparation of the manuscript.

Moscow   Rafail Khasminskii
September, 1967

Contents

Basic Notation

I_T $= \{t : 0 \leq t < T\}$, set of points t such that $0 \leq t < T$, p. 1

I $= I_\infty$, p. 1

U_R $= \{x : |x| < R\}$, p. 4

$\mathbb{R}^l$ Euclidean l-space, p. 2

E $= \mathbb{R}^l \times I$, p. 4

L   class of functions $f(t)$ absolutely integrable on every finite interval, p. 4

C₂   class of functions $V(t, x)$ twice continuously differentiable with respect to x and once continuously differentiable with respect to t, p. 72

$\mathbf{C}_2^0(U)$ class of functions $V(t, x)$ twice continuously differentiable with respect to $x \in U$ and once continuously differentiable with respect to $t \in I$ everywhere except possibly at the point $x = 0$, p. 146

C   class of functions $V(t, x)$ absolutely continuous in t and satisfying a local Lipschitz condition, p. 6

C₀   class of functions $V(t, x) \in \mathbf{C}$ satisfying a global Lipschitz condition, p. 6

$\mathfrak{A}$ σ-algebra of Borel sets in the initial probability space, p. 1

$\mathfrak{B}$ σ-algebra of Borel sets in Euclidean space, p. 47

V_R $= \inf_{t \geq t_0, x \geq R} V(t, x)$, p. 7

$V^{(\delta)}$ $= \sup_{t \geq t_0, |x| < \delta} V(t, x)$, p. 28

A^c complement to the set A, p. 1

$\frac{d^0 V}{dt}$ Lyapunov operator for ODE, p. 6

$\tilde{U}_\delta(\Gamma)$ δ-neighborhood of the set Γ, p. 149

J identity matrix, p. 97

$\mathcal{N}_s$ family of σ-algebras defined on the p. 60

$\widetilde{\mathcal{N}}_t$ family of σ-algebras defined on the p. 68

$\mathbb{1}_A(\cdot)$ indicator function of the set A, p. 62

Chapter 1
Boundedness in Probability and Stability of Stochastic Processes Defined by Differential Equations

1.1 Brief Review of Prerequisites from Probability Theory

Let $\Omega = \{\omega\}$ be a space with a family of subsets $\mathfrak{A}$ such that, for any finite or countable sequence of sets $A_i \in \mathfrak{A}$, the intersection $\bigcap_i A_i$, union $\bigcup_i A_i$ and complement A_i^c (with respect to Ω) are also in $\mathfrak{A}$. Suppose moreover that $\Omega \in \mathfrak{A}$. A family of subsets possessing these properties is known as a σ-algebra. If a probability measure $\mathbf{P}$ is defined on the σ-algebra $\mathfrak{A}$ (i.e. $\mathbf{P}$ is a non-negative countably additive set function on $\mathfrak{A}$ such that $\mathbf{P}(\Omega) = 1$), then the triple $(\Omega, \mathfrak{A}, \mathbf{P})$ is called a probability space and the sets in $\mathfrak{A}$ are called random events. (For more details, see [56, 64, 185].)

The following standard properties of measures will be used without any further reference:

1. If $A \in \mathfrak{A}$, $B \in \mathfrak{A}$, $A \subset B$, then $\mathbf{P}(A) \leq \mathbf{P}(B)$.

2. For any finite or countable sequence A_n in $\mathfrak{A}$,

$$\mathbf{P}\left(\bigcup_n A_n\right) \leq \sum_n \mathbf{P}(A_n).$$

3. If $A_n \in \mathfrak{A}$ and $A_1 \subset A_2 \subset \cdots \subset A_n \subset A_{n+1} \subset \cdots$, then

$$\mathbf{P}\left(\bigcup_n A_n\right) = \lim_{n \to \infty} \mathbf{P}(A_n).$$

4. If $A_n \in \mathfrak{A}$ and $A_1 \supset A_2 \supset A_3 \supset \cdots \supset A_n \supset \cdots$, then

$$\mathbf{P}\left(\bigcap_n A_n\right) = \lim_{n \to \infty} \mathbf{P}(A_n).$$

Proofs of these properties may be found in any textbook on probability theory, such as [95, §8]; or [92, Sect. 1.1].

A random variable is a function $\xi(\omega)$ on Ω which is $\mathfrak{A}$ measurable and almost everywhere finite.[1] In this book we shall consider only random variables which take on values in Euclidean l-space $\mathbb{R}^l$ i.e., such that $\xi(\omega) = (\xi_1(\omega), \ldots, \xi_l(\omega))$ is a vector in $\mathbb{R}^l$ ($l = 1, 2, \ldots$). A vector-valued random variable $\xi(\omega)$ may be defined by its joint distribution function $F(x_1, \ldots, x_l)$, that is, by specifying the probability of the event $\{\xi_1(\omega) < x_1; \ldots; \xi_l(\omega) < x_l\}$. Given any vector $x \in \mathbb{R}^l$ or a $k \times l$ matrix $\sigma = ((\sigma_{ij}))$ ($i = 1, \ldots, k$; $j = 1, \ldots, l$) we shall denote, as usual,

$$|x| = (x_1^2 + \cdots + x_l^2)^{1/2}, \qquad \|\sigma\| = \left(\sum_{i=1}^{k} \sum_{j=1}^{l} \sigma_{ij}^2 \right)^{1/2}.$$

Then we have the well-known inequalities $|\sigma x| \leq \|\sigma\| |x|$, $\|\sigma_1 \sigma_2\| \leq \|\sigma_1\| \|\sigma_2\|$.

The expectation of a random variable $\xi(\omega)$ is defined to be the integral

$$\mathbf{E}\xi = \int_\Omega \xi(\omega) \, P(d\omega),$$

provided the function $|\xi(\omega)|$ is integrable.

Let $\mathcal{B}$ be a σ-algebra of Borel subsets of a closed interval $[s_0, s_1]$, $\mathcal{B} \times \mathfrak{A}$ the minimal σ-algebra of subsets of $I \times \Omega$ containing all subsets of the type $\{t \in \Delta, \omega \in A\}$, where $\Delta \in \mathcal{B}$, $A \in \mathfrak{A}$. A function $\xi(t, \omega) \in \mathbb{R}^l$ is called a measurable stochastic process (random function) defined on $[s_0, s_1]$ with values in $\mathbb{R}^l$ if it is $\mathcal{B} \times \mathfrak{A}$-measurable and $\xi(t, \omega)$ is a random variable for each $t \in [s_0, s_1]$. For fixed ω, we shall call the function $\xi(t, \omega)$ a trajectory or sample function of the stochastic process. In the sequel we shall consider only separable stochastic processes, i.e., processes whose behavior for all $t \in [s_0, s_1]$ is determined up to an event of probability zero by its behavior on some dense subset $\Lambda \in [s_0, s_1]$. To be precise, a process $\xi(t, \omega)$ is said to be *separable* if, for some countable dense subset $\Lambda \in [s_0, s_1]$, there exists an event A of probability 0 such that for each closed subset $C \subset \mathbb{R}^l$ and each open subset $\Delta \subset [s_0, s_1]$ the event

$$\{\xi(t_j, \omega) \in C; \, t_j \in \Lambda \cap \Delta\}$$

implies the event

$$A \cup \{\xi(t, \omega) \in C; \, t \in \Delta\}.$$

A process $\xi(t, \omega)$ is stochastically continuous at a point $s \in [s_0, s_1]$ if for each $\varepsilon > 0$

$$\lim_{t \to s} \mathbf{P}\{|\xi(t, \omega) - \xi(s, \omega)| > \varepsilon\} = 0.$$

The definitions of right and left stochastic continuity are analogous.

It can be proved (see [56, Chap. II, Theorem 2.6]) that for each process $\xi(t, \omega)$ which is stochastically continuous throughout $[s_0, s_1]$, except for a countable subset

---

[1] Sometimes (see Chap. 3), but only when this is explicitly mentioned, we shall find it convenient to consider random variables which can take on the values $\pm\infty$ with positive probability.

of $[s_0, s_1]$, there exists a separable measurable process $\tilde{\xi}(t, \omega)$ such that for every $t \in [s_0, s_1]$

$$\mathbf{P}\{\xi(t, \omega) = \tilde{\xi}(t, \omega)\} = 1 \qquad (\xi(t, \omega) = \tilde{\xi}(t, \omega) \text{ almost surely}).$$

If $\xi(t, \omega)$ is a measurable stochastic process, then for fixed ω the function $\xi(t, \omega)$, as a function of t, is almost surely Lebesgue-measurable. If, moreover, $\mathbf{E}\xi(t, \omega) = m(t)$ exists, then $m(t)$ is Lebesgue-measurable, and the inequality

$$\int_A \mathbf{E}|\xi(t, \omega)| \, dt < \infty$$

implies that the process $\xi(t, \omega)$ is almost surely integrable over A [56, Chap. II, Theorem 2.7].

On the σ-algebra $\mathcal{B} \times \mathfrak{A}$ there is defined the direct product $\mu \times \mathbf{P}$ of the Lebesgue measure μ and the probability measure $\mathbf{P}$. If some relation holds for $(t, \omega) \in A$ and $\mu \times \mathbf{P}(A^c) = 0$, the relation will be said to hold for almost all t, ω. Let $A_1, \dots, A_n$ be Borel sets in $\mathbb{R}^l$, and $t_1, \dots, t_n \in [s_0, s_1]$; the probabilities

$$\mathbf{P}(t_1, \dots, t_n, A_1, \dots, A_n) = \mathbf{P}\{\xi(t_1, \omega) \in A_1, \dots, \xi(t_n, \omega) \in A_n\}$$

are the values of the n-dimensional distributions of the process $\xi(t, \omega)$. Kolmogorov has shown that any compatible family of distributions $\mathbf{P}(t_1, \dots, t_n, A_1, \dots, A_n)$ is the family of the finite-dimensional distributions of some stochastic process.

The following theorem of Kolmogorov will play an important role in the sequel.

Theorem 1.1 *If α, β, k are positive numbers such that whenever $t_1, t_2 \in [s_0, s_1]$,*

$$\mathbf{E}|\xi(t_2, \omega) - \xi(t_1, \omega)|^\alpha < k|t_1 - t_2|^{1+\beta}$$

and $\xi(t, \omega)$ is separable, then the process $\xi(t, \omega)$ has continuous sample functions almost surely (a.s.).

Let $\xi(t, \omega)$ be a stochastic process defined for $t \geq t_0$. The process is said to satisfy the law of large numbers if for each $\varepsilon > 0$, $\delta > 0$ there exists a $T > 0$ such that for all $t > T$

$$\mathbf{P}\left\{\left|\frac{1}{t}\int_{t_0}^{t_0+t} \xi(s, \omega) \, ds - \frac{1}{t}\int_{t_0}^{t_0+t} \mathbf{E}\xi(s, \omega) \, ds\right| > \delta\right\} < \varepsilon. \qquad (1.1)$$

A stochastic process $\xi(t, \omega)$ satisfies the *strong* law of large numbers if

$$\mathbf{P}\left\{\frac{1}{t}\int_{t_0}^{t_0+t} \xi(s, \omega) \, ds - \frac{1}{t}\int_{t_0}^{t_0+t} \mathbf{E}\xi(s, \omega) \, ds \xrightarrow[t \to \infty]{} 0\right\} = 1. \qquad (1.2)$$

The most important characteristics of a stochastic process are its expectation $m(t) = \mathbf{E}\xi(t, \omega)$ and covariance matrix

$$K(s, t) = \text{cov}(\xi(s), \xi(t)) = ((\mathbf{E}[(\xi_i(s) - m_i(s))(\xi_j(t) - m_j(t))])).$$

In particular, all the finite-dimensional distributions of a Gaussian process can be reconstructed from the function $m(t)$ and $K(s,t)$. A Gaussian process is stationary if

$$m(t) = \text{const}, \qquad K(s,t) = K(t-s). \tag{1.3}$$

A stochastic process $\xi(t,\omega)$ satisfying condition (1.3) is said to be stationary in the wide sense. The Fourier transform of the matrix $K(\tau)$ is called the spectral density of the process $\xi(t,\omega)$. It is clear that the spectral density $f(\lambda)$ exists and is bounded if the function $\|K(\tau)\|$ is absolutely integrable.

1.2 Dissipative Systems of Differential Equations

In this section we prove some theorems from the theory of differential equations that we shall need later. We begin with a few definitions.

Let I_T denote the set $0 < t < T$, $I = I_\infty$, $E = \mathbb{R}^l \times I$; U_R the ball $|x| < R$ and U_R^c its complement in $\mathbb{R}^l$. If $f(t)$ is a function defined on I, we write $f \in \mathbf{L}$ if $f(t)$ is absolutely integrable over every finite interval. The same notation $f \in \mathbf{L}$ will be retained for a stochastic function $f(t,\omega)$ which is almost surely absolutely integrable over every finite interval.

Let $F(x,t) = (F_1(x,t), \dots, F_l(x,t))$ be a Borel-measurable function defined for $(x,t) \in E$. Let us assume that for each $R > 0$ there exist functions $M_R(t) \in \mathbf{L}$ and $B_R(t) \in \mathbf{L}$ such that

$$|F(x,t)| \leq M_R(t), \tag{1.4}$$

$$|F(x_2,t) - F(x_1,t)| \leq B_R(t)|x_2 - x_1| \tag{1.5}$$

for $x, x_i \in U_R$.

We shall say that a function $x(t)$ is a solution of the equation

$$\frac{dx}{dt} = F(x,t), \tag{1.6}$$

satisfying the initial condition

$$x(t_0) = x_0 \quad (t_0 \geq 0) \tag{1.7}$$

on the interval $[t_0, t_1]$, if for all $t \in [t_0, t_1]$

$$x(t) = x_0 + \int_{t_0}^{t} F(x(s), s)\, ds. \tag{1.8}$$

In cases where solutions are being considered under varying initial conditions, we shall denote this solution by $x(t, x_0, t_0)$.

The function $x(t)$ is evidently absolutely continuous, and at all points of continuity of $F(x,t)$ it also satisfies (1.6).

Theorem 1.2 *If conditions* (1.4) *and* (1.5) *are satisfied, then the solution* $x(t)$ *of problem* (1.6), (1.7) *exists and is unique in some neighborhood of* t_0. *Suppose moreover that for every solution* $x(t)$ *(if a solution exists) and some function* τ_R *which tends to infinity as* $R \to \infty$, *we have the following "a priori estimate":*

$$\inf\{t : t \geq t_0; \; |x(t)| > R\} \geq \tau_R. \tag{1.9}$$

Then the solution of the problem (1.6), (1.7) *exists and is unique for all* $t \geq t_0$ *(i.e., the solution can be unlimitedly continued for* $t \geq t_0$*).*

Proof We may assume without loss of generality that the function $M_R(t)$ in (1.4) satisfies the inequality

$$|M_R(t)| > 1. \tag{1.10}$$

Therefore we can find numbers R and $t_1 > t_0$ such that $|x_0| \leq R/2$ and

$$\Phi(t_0, t_1) = \int_{t_0}^{t_1} M_R(s)\, ds \, \exp\left\{\int_{t_0}^{t_1} B_R(s)\, ds\right\} = \frac{R}{2}. \tag{1.11}$$

Applying the method of successive approximations to (1.8) on the interval $[t_0, t_1]$,

$$x^{(n+1)}(t) = x_0 + \int_{t_0}^{t} F(x^{(n)}(s), s)\, ds, \quad x^0(t) \equiv x_0,$$

and using (1.4), (1.5) and (1.11), we get the estimates

$$|x^{(1)}(t) - x_0| \leq \int_{t_0}^{t} M_R(s)\, ds \leq \frac{R}{2},$$

$$|x^{(n+1)}(t) - x^{(n)}(t)| \leq \int_{t_0}^{t} B_R(s)|x^{(n)}(s) - x^{(n-1)}(s)|\, ds.$$

Together with (1.11), these imply the inequality

$$|x^{(n+1)}(t) - x^{(n)}(t)| \leq \int_{t_0}^{t} M_R(s)\, ds \frac{[\int_{t_0}^{t} B_R(s)\, ds]^n}{n!}. \tag{1.12}$$

It follows from (1.12) that $\lim_{n \to \infty} x^{(n)}(t)$ exists and that it satisfies (1.8). The proof of uniqueness is similar.

Now consider an arbitrary $T > t_0$ and choose R so that, besides the relations $|x_0| < R/2$ and (1.11), we also have $\tau_{R/2} > T$. Then by (1.9), it follows that $x(t_1) \leq R/2$ and thus the solutions can be continued to a point t_2 such that $\Phi(t_1, t_2) = R/2$. Repeating this procedure, we get $t_n \geq T$ for some n, since the functions $M_R(t)$ and $L_R(t)$ are integrable over every finite interval. This completes the proof. $\qquad \square$

If the function $M_R(t)$ is independent of t and its rate of increase in R is at most linear, i.e.,

$$|F(x, t)| \leq c_1|x| + c_2, \tag{1.13}$$

we get the following estimate for the solution of problem (1.6), (1.7), valid for $t \geq t_0$ and some $c_3 > 0$:

$$|x(t)| \leq |x_0| c_3 e^{c_1(t-t_0)}.$$

We omit the proof now, since we shall later prove a more general theorem. But if condition (1.13) fails to hold, the solution will generally "escape to infinity" in a finite time. (As for example, the solution $x = (1-t)^{-1}$ of the problem $dx/dt = x^2$, $x(0) = 1$.) Since condition (1.13) fails to cover many cases of practical importance, we shall need a more general condition implying that the solution can be unlimitedly continued. We present first some definitions.

The Lyapunov operator associated with (1.6) is the operator d^0/dt defined by

$$\frac{d^0 V(x,t)}{dt} = \overline{\lim_{h \to +0}} \frac{1}{h} [V(x(t+h,x,t),t+h) - V(x,t)]. \qquad (1.14)$$

It is obvious that if $V(x,t)$ is continuously differentiable with respect to x and t, then for almost all t the action of the Lyapunov operator

$$\frac{d^0 V}{dt} = \frac{\partial V}{\partial t} + \sum_{i=1}^{l} \frac{\partial V}{\partial x_i} F_i(x,t) = \frac{\partial V}{\partial t} + \left(\frac{\partial}{\partial x} V, F \right) \qquad (1.15)$$

is simply a differentiation of the function V along the trajectory of the system (1.6).

In his classical work [188], Lyapunov discussed the stability of systems of differential equations by considering non-negative functions for which $d^0 V/dt$ satisfies certain inequalities.

These functions will be called *Lyapunov functions* here.

In Sects. 1.5, 1.6, 1.8, and also in Chaps. 5 to 7 we shall apply Lyapunov's ideas to stability problems for random perturbations.

In this and the next sections we shall use method of Lyapunov functions to find conditions under which the solution can be continued for all $t > 0$ and to conditions of boundedness solution. All Lyapunov functions figuring in the discussion will be henceforth assumed to be absolutely continuous in t, uniformly in x in the neighborhood of every point. Moreover we shall assume a Lipschitz condition with respect to x:

$$|V(x_2,t) - V(x_1,t)| < B|x_2 - x_1| \qquad (1.16)$$

in the domain $U_R \times I_T$, with a Lipschitz constant which generally depends on R and T. We shall write $V \in \mathbf{C}$ in this case. If the function V satisfies condition (1.16) with a constant B not depending on R and T, we shall write $V \in \mathbf{C}_0$.

If $V \in \mathbf{C}$ and the function $y(t)$ is absolutely continuous, then it is easily verified that the function $V(y(t),t)$ is also absolutely continuous. Hence, for almost all t,

$$\frac{d^0 V(x,t)}{dt} = \frac{d}{dt} V(x(t),t) \bigg|_{x(t)=x},$$

where $x(t)$ is the solution of (1.6). We shall use this fact frequently without further reference.

Theorem 1.3 [2] *Assume that there exists a Lyapunov function $V \in \mathbf{C}$ defined on the domain $\mathbb{R}^l \times \{t > t_0\}$ such that for some $c_1 > 0$*

$$V_R = \inf_{(x,t) \in U_R^c \times \{t > t_0\}} V(x,t) \to \infty \quad as\ R \to \infty, \tag{1.17}$$

$$\frac{d^0 V}{dt} \leq c_1 V, \tag{1.18}$$

and let the function F satisfy conditions (1.4), (1.5).
 Then the solution of problem (1.6), (1.7) *can be extended for all $t \geq t_0$.*

The proof of this theorem employs the following well-known lemma, which will be used repeatedly.

Lemma 1.1 *Let the function $y(t)$ be absolutely continuous for $t \geq t_0$ and let the derivative dy/dt satisfy the inequality*

$$\frac{dy}{dt} < A(t)y + B(t) \tag{1.19}$$

for almost all $t \geq t_0$, where $A(t)$ and $B(t)$ are almost everywhere continuous functions integrable over every finite interval. Then for $t > t_0$

$$y(t) < y(t_0) \exp\left\{ \int_{t_0}^t A(s)\,ds \right\} + \int_{t_0}^t \exp\left\{ \int_s^t A(u)\,du \right\} B(s)\,ds. \tag{1.20}$$

Proof It follows from (1.19) that for almost all $t \geq t_0$

$$\frac{d}{dt}\left(y(t) \exp\left\{ -\int_{t_0}^t A(s)\,ds \right\} \right) < B(t) \exp\left\{ -\int_{t_0}^t A(s)\,ds \right\}.$$

Integration of this inequality yields (1.20).   □

Proof of Theorem 1.3   It follows from (1.18) that for almost all t we have $dV(x(t),t)/dt \leq c_1 V(x(t),t)$. Hence, by Lemma 1.1, it follows that for $t > t_0$

$$V(x(t),t) \leq V(x_0,t_0) \exp\{c_1(t - t_0)\}.$$

If τ_R denotes a solution of the equation

$$V(x_0,t_0) \exp\{c_1(\tau_R - t_0)\} = V_R,$$

then condition (1.9) is obviously satisfied. Thus all assumptions of Theorem 1.2 are now satisfied. This completes the proof.   □

---

[2]General conditions for every solution to be unboundedly continuable have been obtained by Okamura and are described in [178]. These results imply Theorem 1.3.

Let us now consider conditions under which the solutions of (1.6) are bounded for $t > 0$. There exist in the literature various definitions of boundedness. We shall adopt here only one which is most suitable for our purposes, referring the reader for more details to [285], [178], and [51, 52].

The system (1.6) is said to be *dissipative* for $t > 0$ if there exists a positive number $R > 0$ such that for each $r > 0$, beginning from some time $T(r, t_0) \geq t_0$, the solution $x(t, x_0, t_0)$ of problem (1.6), (1.7), $x_0 \in U_r$, $t_0 > 0$, lies in the domain U_R. (Yoshizawa [285] calls the solutions of such a system equi-ultimately bounded.)

Theorem 1.4 [3] *A sufficient condition for the system* (1.6) *to be dissipative is that there exist a nonnegative Lyapunov function* $V(x, t) \in \mathbf{C}$ *on E with the properties*

$$V_R = \inf_{(x,t)\in U_R^c \times I} V(x, t) \to \infty \quad as \ R \to \infty, \tag{1.21}$$

$$\frac{d^0 V}{dt} < -cV \quad (c = \text{const} > 0). \tag{1.22}$$

Proof It follows from Lemma 1.1 and from (1.22) that for $t > t_0$, $x_0 \in U_r$,

$$V(x(t), t) \leq V(x_0, t_0) e^{-c(t-t_0)} \leq e^{-c(t-t_0)} \sup_{|x_0| < r} V(x_0, t_0).$$

Therefore $V(x(t), t) < 1$ for $t > T(t_0, r)$. This inequality and (1.21) imply the statement of the theorem. □

Remark 1.1 The converse theorem is also valid: Yoshizawa [285] proves that for each system which is dissipative in the above sense there exists a nonnegative function V with properties (1.21), (1.22), provided $F(x, t)$ satisfies a Lipschitz condition in every bounded subset of E.

Remark 1.2 It is easy to show that the conclusion of Theorem 1.4 remains valid if it is merely assumed that (1.22) holds in a domain U_R^c for some $R > 0$, and in the domain U_R the functions V and $d^0 V/dt$ are bounded above. To prove this, it is enough to apply Lemma 1.1 to the inequality

$$\frac{d^0 V}{dt} < -cV + c_1,$$

which is valid under the above assumptions for some positive constant c_1 and for $(x, t) \in E$.

In the sequel we shall need a certain frequently used estimate; its proof, analogous to the proof of Lemma 1.1, may be found, e.g., in [23].

---

[3] See [285].

Lemma 1.2 (Gronwall–Bellman Lemma) *Let $u(t)$ and $v(t)$ be nonnegative functions and let k be a positive constant such that for $t \geq s$*

$$u(t) \leq k + \int_s^t u(t_1)v(t_1) \, dt_1.$$

Then for $t \geq s$

$$u(t) \leq k \exp\left\{ \int_s^t v(t_1) \, dt_1 \right\}.$$

1.3 Stochastic Processes as Solutions of Differential Equations

Let $\xi(t, \omega)$ $(t \geq 0)$ be a separable measurable stochastic process with values in $\mathbb{R}^k$, and let $G(x, t, z)$ $(x \in \mathbb{R}^l, t \geq 0, z \in \mathbb{R}^k)$ be a Borel-measurable function of (x, t, z) satisfying the following conditions:

1. There exists a stochastic process $B(t, \omega) \in \mathbf{L}$ such that for all $x_i \in \mathbb{R}^l$

$$|G(x_2, t, \xi(t, \omega)) - G(x_1, t, \xi(t, \omega))| \leq B(t, \omega)|x_1 - x_2|. \tag{1.23}$$

2. The process $G(0, t, \xi(t, \omega))$ is in $\mathbf{L}$, i.e., for every $T > 0$,

$$\mathbf{P}\left\{ \int_0^T |G(0, t, \xi(t, \omega))| \, dt < \infty \right\} = 1. \tag{1.24}$$

We shall show presently that under these assumptions the equation

$$\frac{dx}{dt} = G(x, t, \xi(t, \omega)) \tag{1.25}$$

with initial condition

$$x(t_0) = x_0(\omega) \tag{1.26}$$

determines a new stochastic process in $\mathbb{R}^l$ for $t \geq t_0$.

Theorem 1.5 *If conditions (1.23) and (1.24) are satisfied, then problem (1.25), (1.26) has a unique solution $x(t, \omega)$, determining a stochastic process which is almost surely absolutely continuous for all $t \geq t_0$. For each $t \geq t_0$, this solution admits the estimate*

$$|x(t, \omega) - x_0(\omega)| \leq \int_{t_0}^t |G(x_0(\omega), s, \xi(s, \omega))| \, ds \, \exp\left\{ \int_{t_0}^t B(s, \omega) \, ds \right\}. \tag{1.27}$$

The proof is analogous to that of Theorem 1.2.

Example 1.1 Consider the linear system

$$\frac{dx}{dt} = A(t, \omega)x + b(t, \omega), \qquad x(0) = x_0(\omega).$$

If $\|A(t, \omega)\|, |b(t, \omega)| \in L$, then it follows from Theorem 1.5 that this system has a solution which is a continuous stochastic process for all $t > 0$.

The global Lipschitz condition (1.23) fails to hold in many important applications. Most frequently the following local Lipschitz condition holds: For each $R > 0$, there exists a stochastic process $B_R(t, \omega) \in \mathbf{L}$ such that if $x_i \in U_R$, then

$$|G(x_2, t, \xi(t, \omega)) - G(x_1, t, \xi(t, \omega))| \leq B_R(t, \omega)|x_2 - x_1|. \qquad (1.28)$$

As we have already noted in Sect. 1.2, condition (1.28) does not prevent the sample function escaping to infinity in a finite time, even in the deterministic case. However, we have the following theorem which is a direct corollary of Theorem 1.2.

Theorem 1.6 *Let $\tau(R, \omega)$ be a family of random variables such that $\tau(R, \omega) \uparrow \infty$ almost surely as $R \to \infty$. Suppose that these random variables satisfy almost surely for each solution $x(t, \omega)$ of problem (1.25), (1.26) (if a solution exists) the following inequality*:

$$\inf\{t : |x(t, \omega)| \geq R\} \geq \tau(R, \omega). \qquad (1.29)$$

Assume moreover that conditions (1.24) and (1.28) are satisfied. Then the solution of problem (1.25), (1.26) is almost surely unique and it determines an absolutely continuous stochastic process for all $t \geq t_0$ (unboundedly continuable for $t \geq t_0$).

Assume that the function G in (1.25) depends linearly on the third variable, i.e.,

$$\frac{dx}{dt} = F(x, t) + \sigma(x, t)\xi(t, \omega). \qquad (1.30)$$

(Here σ is a $k \times l$ matrix, ξ a vector in $\mathbb{R}^k$ and k a positive integer.) Then the solution of (1.30) can be unboundedly continued if there exists a Lyapunov function of the truncated system

$$\frac{dx}{dt} = F(x, t). \qquad (1.31)$$

Let us use $d^{(1)}/dt$ to denote the Lyapunov operator of the system (1.30), retaining the notation d^0/dt for the Lyapunov operator of the system (1.31).

Theorem 1.7 *Let $\xi(t, \omega) \in \mathbf{L}$ be a stochastic process, F a vector and σ a matrix satisfying the local Lipschitz condition (1.16), where $F(0, t) \in \mathbf{L}$ and*

$$\sup_{\mathbb{R}^l \times \{t > t_0\}} \|\sigma(x, t)\| < c_2. \qquad (1.32)$$

Assume that a Lyapunov function $V(x,t) \in \mathbf{C}_0$ of the system (1.31) exists with

$$V_R = \inf_{U_R^c \times \{t > t_0\}} V(x,t) \to \infty \quad as \ R \to \infty, \tag{1.33}$$

$$\frac{d^0 V}{dt} < c_1 V. \tag{1.34}$$

Then the solution of problem (1.30), (1.26) exists and determines an absolutely continuous stochastic process for all $t \geq t_0$.

To prove this theorem we need the following lemma.

Lemma 1.3 *If $V(x,t) \in \mathbf{C}_0$, then for almost all t the following relation holds almost surely:*

$$\frac{d^{(1)} V(x,t)}{dt} \leq \frac{d^0 V(x,t)}{dt} + B \|\sigma(x,t)\| |\xi(t,\omega)|, \tag{1.35}$$

where B is the constant in the condition (1.16).

Proof It can be easily verified that the difference $x(t+h, \omega, x, t) - x(t+h, x, t)$ between solutions of (1.30) and (1.31) with the initial condition $x(t) = x$, satisfies for almost all t, ω the inequality

$$|x(t+h, \omega, x, t) - x(t+h, x, t)| \leq h \|\sigma(x,t)\| |\xi(t,\omega)| + o(h) \quad (h \to 0).$$

This inequality, together with (1.16), implies (1.35). $\qquad\square$

Proof of Theorem 1.7 We shall show that the assumptions of Theorem 1.6 are satisfied. Since conditions (1.24) and (1.28) are obviously satisfied, it will suffice to prove (1.29). Let $x(t,\omega)$ be a solution of problem (1.30), (1.26). It follows from the assumptions of the theorem and from Lemma 1.3 that the function $V(x(t,\omega), t)$ is absolutely continuous, and for almost all t, ω

$$\frac{dV(x(t,\omega), t)}{dt} \leq \frac{d^0 V(x(t,\omega), t)}{dt} + B \|\sigma(x(t,\omega), t)\| |\xi(t,\omega)|$$

$$\leq c_1 V(x(t,\omega), t) + B c_2 |\xi(t,\omega)|.$$

Combining this with Lemma 1.1 we get that almost surely

$$V(x(t,\omega), t) \leq e^{c_1(t-t_0)} \left[V(x_0(\omega), t_0) + B c_2 \int_{t_0}^{t} |\xi(s,\omega)| \, ds \right]. \tag{1.36}$$

Let $\tau_R(\omega)$ denote a solution of the equation

$$e^{c_1(\tau_R - t_0)} \left[V(s_0(\omega), t_0) + B c_2 \int_{t_0}^{\tau_R} |\xi(s,\omega)| \, ds \right] = V_R. \tag{1.37}$$

It now follows from the relation $\xi(t, \omega) \in \mathbf{L}$ and from (1.33) that $\tau_R \uparrow \infty$ almost surely as $R \to \infty$. (1.29) follows now from (1.36) and (1.37). Thus all assumptions of Theorem 1.6 are satisfied. $\qquad \square$

Remark 1.3 If the relation $|\xi(t, \omega)|^{(1+\varepsilon)/\varepsilon} \in \mathbf{L}$ holds for some $\varepsilon > 0$, condition (1.32) can be slightly weakened and replaced by the condition

$$\|\sigma(x, t)\|^{1+\varepsilon} \le c_3 V(x, t). \tag{1.38}$$

To prove this, we need only use Young's inequality

$$|ab| < \frac{|a|^p}{p} + \frac{|b|^q}{q} \quad \left(\frac{1}{p} + \frac{1}{q} = 1, \ p, q > 0 \right) \tag{1.39}$$

for estimating $\|\sigma\| |\xi|$. In particular, if for each $T > 0$, there is a constant c such that the process $\xi(t, \omega)$ satisfies the condition

$$\mathbf{P}\left\{ \sup_{0 \le t \le T} |\xi(t, \omega)| < c \right\} = 1,$$

then it is enough to require that inequality (1.38) holds for sufficiently small $\varepsilon > 0$.

Remark 1.4 The conditions of Theorem 1.7 guarantee that the solutions of (1.30) are unboundedly continuable, uniformly in the following sense: For all initial conditions $x_0(\omega)$ which satisfy the relation

$$\mathbf{P}\{|x_0(\omega)| < K\} = 1 \tag{1.40}$$

for some K, one can find a family of random variables $\tau(R, \omega)$ satisfying condition (1.29). Since

$$\mathbf{P}\left\{ \max_{0 \ge t \ge T} |x(t, \omega, x_0(\omega))| > R \right\} \le \mathbf{P}\{\tau_R < T\},$$

this implies in particular that for every $\varepsilon > 0$, $T > 0$ and $K > 0$ there exists an $R > 0$ such that

$$\mathbf{P}\left\{ \max_{0 \ge t \ge T} |x(t, \omega, x_0(\omega))| > R \right\} > \varepsilon$$

for all $x_0(\omega)$ satisfying condition (1.40).

Example 1.2 In the one-dimensional case with the Lyapunov function $V(x, t) = |x| + 1$ we get the following result. If $F \in \mathbf{C}$, $\sigma \in \mathbf{C}$, σ satisfies the condition (1.32), while $\xi(t, \omega)$, $F(0, t) \in \mathbf{L}$, then a sufficient condition for the solutions of problem (1.30), (1.26) to be unboundedly extendable is that $F(x, t) \operatorname{sign} x < c(|x| + 1)$ for some $c > 0$.

Example 1.3 Consider the equation

$$x'' + f(x)x' + g(x) = \sigma(x, x')\xi(t, \omega). \tag{1.41}$$

This equation describes the process "at the output" of many mechanical systems driven by a stochastic process. In particular, for $f(x) = x^2 - 1$, $g(x) = x$ and $\sigma(x, x') = 1$, the output process is that of a system described by a Van der Pol equation. Let the function $f(x)$ be bounded from below and assume that

$$|\sigma(x, x')| < c_1, \qquad \left|\frac{g(x)}{x}\right| < c_2.$$

Then

$$V(x, y) = (x^2 + y^2)^{1/2} \in \mathbf{C}_0$$

is obviously a Lyapunov function for the system

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = -f(x)y - g(x).$$

Moreover V satisfies conditions (1.33) and (1.34). Applying Theorem 1.7, we see that the process in (1.41) exists for all $t \geq t_0$ provided that $\xi(t, \omega) \in \mathbf{L}$.

1.4   Boundedness in Probability of Stochastic Processes Defined by Systems of Differential Equations

A stochastic process $\xi(t, \omega)$ $(t \geq 0)$ is said to be *bounded in probability* if the random variables $|\xi(t, \omega)|$ are bounded in probability uniformly in t, i.e.

$$\sup_{t>0} \mathbf{P}\{|\xi(t, \omega)| > R\} \to 0 \quad \text{as } R \to \infty.$$

We shall say that a random variable $x_0(\omega)$ is in the class A_{R_0} if

$$\mathbf{P}\{|x_0(\omega)| < R_0\} = 1. \tag{1.42}$$

The system (1.25) will be called *dissipative* if the random variables $|x(t, \omega, x_0, t_0)|$ are bounded in probability, uniformly in $t \geq t_0$ whenever $x_0(\omega) \in A_R$ for some $R > 0$.

It is readily seen that this definition agrees with that of a deterministic dissipative system (see Sect. 1.2).

Theorem 1.8 *Let $V(x, t) \in \mathbf{C}_0$ be a non-negative Lyapunov function, defined on the domain E which satisfies condition (1.33) and the condition*

$$\frac{d^0 V}{dt} \leq -c_1 V \quad (c_1 = \text{const} > 0). \tag{1.43}$$

Let F and σ satisfy a local Lipschitz condition (1.16), and let σ also satisfy condition (1.32).

Then the system (1.30) *is dissipative for every stochastic process* $\xi(t, \omega)$ *such that*

$$\sup_{t>0} \mathbf{E}|\xi(t, \omega)| < \infty. \tag{1.44}$$

Before proving this theorem, we shall prove a lemma which yields a convenient for further applications form of Chebyshev's inequality.

Lemma 1.4 *Let $V(x, t)$ be a nonnegative function and $\eta(t, \omega)$ a stochastic process such that $\mathbf{E}V(\eta(t, \omega), t)$ exists.*
    Then

$$\mathbf{P}\{|\eta(t, \omega)| > R\} \leq \frac{\mathbf{E}V(\eta(t, \omega), t)}{\inf_{U_R^c \times \{s>t_0\}} V(x, s)}. \tag{1.45}$$

The proof follows from the following chain of inequalities:

$$\mathbf{E}V(\eta(t, \omega), t) \geq \int_{|\eta(t, \omega)|>R} V(\eta(t, \omega), t)\,\mathbf{P}(d\omega)$$

$$\geq \inf_{U_R^c \times \{s>t_0\}} V(x, s)\mathbf{P}\{|\eta(t, \omega)| > R\}. \qquad \square$$

Proof of Theorem 1.8 Let $x(t, \omega)$ be a solution of problem (1.30), (1.26). Then the function $V(x(t, \omega), t)$ is differentiable for almost all t, ω. By Lemma 1.3 and by (1.43),

$$\frac{dV(x(t, \omega), t)}{dt} \leq \frac{d^0 V(x(t, \omega), t)}{dt} + Bc_2|\xi(t, \omega)|$$

$$\leq -c_1 V(x(t, \omega), t) + Bc_2|\xi(t, \omega)|.$$

Hence, by Lemma 1.1,

$$V(x(t, \omega), t) \leq V(x_0(\omega), t_0)e^{c_1(t-t_0)} + Bc_2 \int_{t_0}^{t} e^{c_1(s-t)}|\xi(s, \omega)|\,ds.$$

Calculating the expectation of both sides of this inequality and using (1.44), we see that the function $\mathbf{E}V(x(t, \omega), t)$ is bounded uniformly for $t \geq t_0$ and for all $x_0(\omega)$ satisfying condition (1.42). Together with (1.45), this implies the theorem. $\qquad \square$

Remark 1.5 It is clear from Remark 1.1 that the existence of a function V satisfying conditions (1.33), (1.43) is not only sufficient but also necessary for the system (1.30) to be dissipative for each stochastic process $\xi(t, \omega)$ satisfying (1.44).

Remark 1.6 If for some $\varepsilon > 0$

$$\sup_{t>0} \mathbf{E}|\xi(t, \omega)|^{(1+\varepsilon)/\varepsilon} < \infty,$$

then using (1.39) it is easy to show that one may replace condition (1.32) in the formulation of Theorem 1.8 by condition (1.38). Another modification of this theorem is obtained by requiring that condition (1.43) only holds in some U_R, where $R > 0$, and that V and $d^0 V/dt$ are bounded in the domain U_R (see Remark 1.2, and also [121]).

Remark 1.7 Let the conditions of Theorem 1.8 are valid and moreover there exist positive constants c_3 and c_4 such that

$$V(x, t) > c_3|x| - c_4, \tag{1.46}$$

then it follows from Theorem 1.8 that

$$\sup_{t>0} \mathbf{E}|x(t, \omega)| < \infty.$$

The following theorem generalizes this observation.

Theorem 1.9 *Let the functions V, F and σ satisfy the assumptions of Theorem 1.8 and assume moreover that V satisfies also (1.46). Suppose further that for some $\alpha > 1$*

$$\sup_{t>0} \mathbf{E}|\xi(t, \omega)|^\alpha < \infty. \tag{1.47}$$

Then every solution $x(t, \omega)$ of problem (1.30), (1.26) satisfies the inequality

$$\sup_{t>0} \mathbf{E}|x(t, \omega)|^\alpha < \infty.$$

Moreover, there exist constants c and $T = T(R_0, t_0)$ such that for every initial condition $x_0(\omega)$ which satisfies the equality (1.42) for some R_0, the solution $x(t, \omega)$ satisfies for all $t > T(R_0, t_0)$ the following inequality:

$$\mathbf{E}|x(t, \omega)|^\alpha < c.$$

Proof Consider the Lyapunov function $W(x, t) = [V(x, t)]^\alpha$. The assumptions of the theorem, Lemma 1.3 and (1.39) imply that

$$\frac{dW(x(t, \omega), t)}{dt} \leq -c_5 W(x(t, \omega), t) + c_6[V(x(t, \omega), t)]^{\alpha-1}|\xi(t, \omega)|$$

$$\leq -c_7 W(x(t, \omega), t) + c_8|\xi(t, \omega)|^\alpha$$

for some $c_5 > 0, \ldots, c_8 > 0$. Further, as in the proof of Theorem 1.8, we see that

$$\sup_{t>0} \mathbf{E}W(x(t, \omega), t) < \infty.$$

The first part of the theorem follows now from this inequality and the inequality

$$W(x, t) \geq c_9|x|^\alpha - c_{10},$$

which is a consequence of (1.46). The proof of the second part is analogous. $\square$

By considering various narrower classes of stochastic processes $\xi(t, \omega)$, we can derive various dissipativity conditions under less stringent restrictions on the Lyapunov functions. The following theorem is an example.

Theorem 1.10 *Let the process $\xi(t, \omega)$ be such that for some $c_1 > 0$, $c_2 > 0$, $A > 0$ and all $0 \leq s \leq t$*

$$\mathbf{E} \exp\left\{ c_1 \int_s^t |\xi(u, \omega)| \, du \right\} \leq A \exp\{c_2(t - s)\}. \tag{1.48}$$

Assume that there exists a non-negative function $V(x, t) \in \mathbf{C}_0$ defined on E, satisfying condition (1.33) and the conditions

$$\sup_{t > 0} V(0, t) < \infty, \qquad \frac{d^0 V}{dt} < c,$$

$$\frac{d^0 V}{dt} < -c_2 - \varepsilon \quad \text{for } |x| > R_0 \text{ and some } \varepsilon > 0,$$

$$\lim_{R \to \infty} \sup_{x_i \in U_R^c, \, t \geq 0} \frac{|V(x_2, t) - V(x_1, t)|}{|x_2 - x_1|} = B_1.$$

Further let F and σ satisfy the condition (1.16) and the condition $\|\sigma\| \leq K$, where $B_1 K < c_1$. Then the system (1.30) is dissipative.

Proof Let $V(x, t)$ be a function satisfying the assumptions of the theorem. Assume moreover that $R > R_0$ is large enough, so that for $|x_i| > R$ we have

$$|V(x_2, t) - V(x_1, t)| < \frac{c_1}{K} |x_2 - x_1|.$$

Set $W(x, t) = \exp\{V(x, t)\}$. It follows from the assumptions of the theorem that for almost all $t \geq t_0$ and for (t, ω) such that $|x(t, \omega)| > R$ we have

$$\frac{dW(x(t, \omega), t)}{dt} \leq W\left[\frac{d^0 V}{dt} + c_1 |\xi(t, \omega)| \right] \leq W[-(c_2 + \varepsilon) + c_1 |\xi(t, \omega)|].$$

Since $V \in \mathbf{C}_0$ and both V and $d^0 V/dt$ are bounded for $|x| > R$, this implies that there exist constants c_3, c_4, such that the following estimate is valid for almost all t, ω:

$$\frac{dW(x(t, \omega), t)}{dt} \leq W[-(c_2 + \varepsilon) + c_1 |\xi(t, \omega)|] + c_3 + c_4 |\xi(t, \omega)|.$$

Applying Lemma 1.1, we see that almost surely

$$W(x(t, \omega), t) \le W(x_0(\omega), t_0) \exp\left\{\int_{t_0}^t (-c_2 - \varepsilon + c_1|\xi(s, \omega)|)\, ds\right\}$$

$$+ \int_{t_0}^t \exp\left\{\int_s^t (-c_2 - \varepsilon + c_1|\xi(u, \omega)|)\, du\right\}(c_3 + c_4|\xi(s, \omega)|)\, ds$$

$$\le (W(x_0(\omega), t_0) + c_5) \exp\left\{\int_{t_0}^t (-c_2 + c_1|\xi(s, \omega)|)\, ds\right\}$$

$$+ c_6 \int_{t_0}^t \exp\left\{\int_s^t (-c_2 - \varepsilon + c_1|\xi(u, \omega)|)\, du\right\} ds.$$

Taking the expectation of each of the sides of this inequality and using (1.42) and (1.48), we easily see that $\mathbf{E}W(x(t, \omega), t) < c_7$. This together with (1.45) implies the assertion of Theorem 1.10. $\square$

The following example shows that the assertion of Theorem 1.8 fails to hold if we replace condition (1.44) in this theorem by the condition

$$\sup \mathbf{E}|\xi(t, \omega)|^\alpha < \infty \quad (\alpha < 1).$$

Example 1.4 Let $x(t, \omega)$ be a solution in $\mathbb{R}^l$ of the problem

$$\frac{dx}{dt} = -x + \xi(t, \omega), \quad x(1) = 0. \tag{1.49}$$

Define the stochastic process $\xi(t, \omega)$ by

$$\xi(t, \omega) = \begin{cases} 2^{k/a} \exp\{2^k - t + \gamma(2^k - \tau_k)\} & \text{for } t \in [\tau_k, \tau_k + 2^{-k}], \\ 0 & \text{otherwise.} \end{cases} \tag{1.50}$$

Here $\tau_1, \tau_2, \ldots$ are independent random variables such that τ_k is distributed on the interval $[2^{k-1}, 2^k - 2^{-k}]$ with density

$$p_k(s) = \lambda_k \exp\{-\gamma(2^k - s)\}, \tag{1.51}$$

where $\gamma > 0$ and λ_k is determined by the normalization requirement (it is clear that $\lambda_k \to \gamma$ as $k \to \infty$).
From (1.49)–(1.51) we readily get the estimate

$$x(2^k, \omega) = \int_1^{2^k} \exp(s - 2^k)\xi(s, \omega)\, ds > \int_{2^{k-1}}^{2^k} \exp(s - 2^k)\xi(s, \omega)\, ds$$

$$= \int_{\tau_k}^{\tau_k + 2^{k-1}} 2^{k/\alpha} e^{\gamma(2^k - \tau_k)}\, ds \ge 2^{k(1/\alpha - 1)} \to \infty \quad \text{as } k \to \infty$$

which holds almost surely.

On the other hand, if $\gamma > \alpha/(1 - \alpha)$, $2^{k-1} \le t \le 2^k$, then

$$\mathbf{E}|\xi(t, \omega)|^\alpha \le \lambda_k \int_{t-2^{-k}}^t e^{-\gamma(2^k-s)} 2 e^{\alpha(2^k-t)+\alpha\gamma(2^k-s)}\,ds$$

$$\le \lambda_k e^{-(2^k-t)[\gamma(1+\alpha)-\alpha]} < \infty.$$

Thus, if (1.44) does not hold, we cannot assert that the system (1.30) is dissipative, even when the unperturbed system (1.31) is an asymptotically stable linear system.

The next example will show that condition (1.48) in Theorem 1.10 cannot be replaced by (1.44), or even by the stronger condition that $\mathbf{E}|\xi(t, \omega)| \to 0$ as $t \to \infty$.

Example 1.5 Consider in $\mathbb{R}^1$ the problem

$$\frac{dx}{dt} = -\operatorname{sign} x \cdot \ln(|x| + 1) + \eta(t, \omega), \qquad x(0) = 0, \qquad (1.52)$$

where $\eta(t, \omega)$ is a stochastic step process satisfying the conditions

$$\eta(t, \omega) = \begin{cases} 2(2^{n+1} - \tau_n) \ln(2^{n+1} - \tau_n) & \text{for } \tau_n < t < \tau_n + 1 \ (n = 2, 3, \dots), \\ 0 & \text{otherwise.} \end{cases}$$

Let $\tau_2, \dots, \tau_n, \dots$ be independent random variables, τ_n being distributed on the interval $[2^n, 2^{n+1} - 1]$ with density

$$p_n(s) = \frac{c_n}{(2^{n+1} - s) \ln(2^{n+1} - s)}, \qquad c_n = (\ln \ln(2^n - 2) - \ln \ln 2)^{-1}.$$

It is clear that for $2^n \le t \le 2^{n+1}$ we have

$$\mathbf{E}\eta(t, \omega) = \mathbf{E}|\eta(t, \omega)| \le \int_{t-1}^t 2 p_n(s)(2^{n+1} - s) \ln(2^{n+1} - s)\,ds = 2c_n$$

and therefore $\mathbf{E}|\eta(t, \omega)| \to 0$ as $t \to \infty$.

Let $\tilde{x}(t, \omega)$ denote the solution of (1.52) satisfying the initial condition $\tilde{x}(\tau_n, \omega) = 0$. Then, by the uniqueness of the solution of the Cauchy problem for (1.52) and by the definition of the process $\eta(t)$, we have the inequality

$$x(2^{n+1}, \omega) \ge \tilde{x}(2^{n+1}, \omega) = \int_{\tau_n}^{2^{n+1}} [\eta(t, \omega) - \ln(\tilde{x}(t, \omega) + 1)]\,dt$$

$$\ge 2(2^{n+1} - \tau_n) \ln(2^{n+1} - \tau_n)$$

$$- (2^{n+1} - \tau_n) \ln[2(2^{n+1} - \tau_n) \ln(2^{n+1} - \tau_n)]. \qquad (1.53)$$

By the definition of $p_n(s)$,

$$\mathbf{P}\{2^{n+1} - \tau_n > n\} > \frac{1}{2}.$$

Hence, using (1.53), we see that for sufficiently large n

$$\mathbf{P}\left\{x(2^{n+1},\omega) > \frac{n}{2}\ln n\right\} > \frac{1}{2}.$$

This means that the system (1.52) is non-dissipative.

It is readily seen that the Lyapunov function $V(x) = |x|$ for this system satisfies

$$\frac{d^0 V}{dt} = -\ln(V+1) \to -\infty \quad \text{as } |x| \to \infty.$$

Using the same method, one easily constructs examples of non-dissipative systems satisfying all assumptions of Theorem 1.8 except (1.43), instead of which we have

$$\frac{d^0 V}{dt} < -\Phi(V), \tag{1.54}$$

where $\Phi(V)$ is any function such that

$$\int_1^\infty \frac{dx}{x\Phi(x)} = \infty. \tag{1.55}$$

(Examples of this type were constructed by the author in [121].) It is as yet an open question whether condition (1.43) in Theorem 1.8 may be replaced by condition (1.54) with a function $\Phi(V)$ such that the integral (1.55) is convergent. We do not even know the answer to the following more specific question: Do there exist non-dissipative systems of the type

$$\frac{dx}{dt} = \begin{cases} -|x|^\alpha \operatorname{sign} x + \xi(t,\omega) & \text{for } |x| > 1, \\ -x + \xi(t,\omega) & \text{for } |x| \leq 1, \end{cases}$$

where $0 < \alpha < 1$ and $\xi(t,\omega)$ satisfies condition (1.44)?

We now apply the theorems of this section to one-dimensional systems.

Example 1.6 Consider (1.30) in $\mathbb{R}^1$, and assume that $|\xi(t,\omega)| < c$ almost surely. Assume further that the necessary smoothness conditions hold and that $|\sigma| \leq k$. Set $V(x) = |x|$, $c_1 = k + \varepsilon$, $c_2 = c(k + \varepsilon)$. Condition (1.48) is obviously valid if the constants c_1, c_2 are chosen in this way. If moreover

$$\frac{d^0 V}{dt} = F(x,t)\operatorname{sign} x \leq -ck - \varepsilon_1 \quad (|x| > R_0) \tag{1.56}$$

for some $\varepsilon_1 > 0$ and all sufficiently large $|x|$, then also all the other assumptions of Theorem 1.10 hold. Thus we may conclude from this theorem the following corollary.

Corollary 1.1 *A sufficient condition for* (1.30) *to be dissipative in* $\mathbb{R}^1$ *is the existence of positive constants* c, k, ε_1 *such that* (1.56) *and*

$$\mathbf{P}\{|\xi(t,\omega)| \leq c\} = 1, \qquad |\sigma| \leq k$$

hold.

On the other hand, it is clear that if

$$F(x,t) > -ck$$

holds for all $x > R_0$, then the equation is non-dissipative for $\sigma = k, \xi(t,\omega) \equiv c$.

Example 1.7 Suppose that for some positive constant c_1

$$\frac{F(x,t)}{x} < -c_1, \tag{1.57}$$

whenever $|x| > R_0$ and assume that the process $\xi(t,\omega)$ satisfies condition (1.44). Considering the Lyapunov function

$$V(x,t) = \begin{cases} |x| - R_0, & |x| > R_0, \\ 0, & |x| \leq R_0, \end{cases}$$

we see that all assumptions of Theorem 1.8 are satisfied. Thus, relations (1.57) and (1.44) are sufficient conditions for the one-dimensional system (1.30) to be dissipative.

Note that the above Lyapunov function satisfies inequality (1.46). Thus, applying Theorem 1.9, we get the following result: If condition (1.57) is satisfied and condition (1.47) holds for some $\alpha > 1$, then the solution $x(t,\omega)$ of problem (1.30), (1.26) in $\mathbb{R}^1$ has a bounded α-th moment.

Example 1.8 [4] Let us again consider the equation

$$x'' + f(x)x' + g(x) = \sigma(x,x')\xi(t,\omega). \tag{1.58}$$

Assume that the process $|\xi(t,\omega)|$ has a bounded expectation, that $|\sigma|$ is bounded and that there exists an $x_0 > 0$ such that

$$0 < c_1 < \frac{g(x)}{x} < c_2, \qquad 0 < c_3 < f(x) < c_4 \tag{1.59}$$

for $|x| > x_0$. Then the process defined by the system (1.58) is dissipative. To prove this consider the system

$$x' = y, \qquad y' = -f(x)y - g(x) + \sigma(x,y)\xi(t,\omega) \tag{1.60}$$

---

[4]The author's exposition of this example in [121] contains an error. The following corrected version is due to Nevelson.

which is equivalent to (1.58), and set

$$F(x) = \int_0^x f(t)\,dt, \qquad G(x) = \int_0^x g(t)\,dt,$$

$$W(x, y) = (F(x) - \gamma x)y + G(x) + \int_0^x f(t)(F(t) - \gamma t)\,dt + 1 + \frac{y^2}{2},$$

$$V(x, y) = \begin{cases} [W(x, y)]^\alpha - c & \text{for } [W(x, y)]^\alpha > c, \\ 0 & \text{for } [W(x, y)]^\alpha \le c. \end{cases}$$

Regarding the function W was a quadratic form in y and using (1.59), we easily see that for a certain $\gamma > 0$ we have that $W \to \infty$ as $r = (x^2 + y^2)^{1/2} \to \infty$. Next, we can choose an $\alpha > 0$ in such a way that $V(x, y) \in \mathbf{C}_0$. Using the equality

$$\frac{d^0 W}{dt} = -[\gamma y^2 + g(F(x) - \gamma x)]$$

and (1.59) we see that, for sufficiently small $\gamma > 0$ and $\beta > 0$, condition (1.43) holds whenever $r > r_0$. Hence it follows that for a suitable choice of c inequality (1.43) is valid for $V(x, y)$ everywhere. It now follows from Theorem 1.8 that our process is dissipative.

For the general system (1.25) one can prove the following result which is analogous to a theorem of Demidovich [51, 52] for the deterministic case.

Theorem 1.11 *Let the following conditions hold*:

1. $\mathbf{E}|G(0, t, \xi(t, \omega))| < c < \infty$ $(t \ge t_0)$.
2. *There exists a symmetric positive definite matrix $D = ((d_{ij}))$ such that the Jacobian $J(x, t, z) = ((\partial G(x, t, z)/\partial x))$, symmetrized by the matrix D, is negative definite uniformly in x, t and z, i.e., all roots of the symmetric matrix $DJ + J^*D$ satisfy the inequality $\lambda(x, z, t) < -\lambda_U < 0$. Then the system (1.25) is dissipative.*

Proof Set $V(x) = (Dx, x)^{1/2}$. Obviously,

$$\frac{dV(X(t, \omega))}{dt} = (\text{grad } V, G) = \frac{(DG(X(t, \omega), t, \xi(t, \omega)), X(t, \omega))}{V(X(t, \omega))}.$$

It follows from the assumptions of the theorem and from the fundamental lemma in [51, 52] that

$$(DG(x, t, \zeta) - DG(0, t, z), x) < -\lambda_0(x, x).$$

Thus we get the inequality

$$\frac{dV(X(t, \omega))}{dt} \le -\lambda_0 \frac{(X(t, \omega), X(t, \omega))}{V(X(t, \omega))} + \frac{(DG(0, t, \xi(t, \omega), X(t, \omega)))}{V(X(t, \omega))}$$

$$\leq -c_1 V + c_2 G(0, t, \xi(t, \omega)).$$

By Lemma 1.1, this inequality implies the desired conclusion.   □

1.5 Stability[5]

In this section we shall study conditions ensuring the stability of a particular solution $y = y(t, \omega)$ of the equation

$$\frac{dx}{dt} = G(x, t, \xi(t, \omega)). \tag{1.61}$$

Following the usual procedure of introducing new variables, equal to the deviations of the corresponding coordinates of the "perturbed" motion from their "unperturbed" values, we see that we only need to consider the stability of the solution $x(t) \equiv 0$ of an equation of type (1.61) in which the function G satisfies the condition

$$G(0, t, \xi(t, \omega)) \equiv 0. \tag{1.62}$$

Even in the deterministic case the concept of stability of the trivial solution $x(t) \equiv 0$ can be given various meanings. For example, one distinguishes between local stability and stability in the large, also between asymptotic and nonasymptotic stability. The diversity is even greater in the presence of "randomness". We shall not list here all the possible definitions, but we shall confine ourselves to those which are in our view of greatest practical interest. Accordingly, we introduce the following definitions.

The solution $x(t) \equiv 0$ is said to be

1. *(Weakly) stable in probability* (for $t \geq t_0$) if, for every $\varepsilon > 0$ and $\delta > 0$, there exists an $r \geq 0$ such that if $t > t_0$ and $|x_0| < r$, then

$$\mathbf{P}\{|x(t, \omega, t_0, x_0)| > \varepsilon\} < \delta. \tag{1.63}$$

2. *(Weakly) asymptotically stable in probability*[6] if it is stable in probability and, for each $\varepsilon > 0$, there exists an $r = r(\varepsilon)$ such that for $t \to \infty$

$$\mathbf{P}\{|x(t, \omega, t_0, x_0)| > \varepsilon\} \to 0, \quad \text{if } |x_0| < r.$$

3. *p-stable*, if for each $\varepsilon > 0$, there exists an $r > 0$ such that

$$\mathbf{E}|x(t, \omega, t_0, x_0)|^p < \varepsilon \quad (p > 0),$$

---

[5]Almost sure stability has been considered by Kozin [150] and Caughey and Gray [43] for less general systems. Mean and mean square stability has been considered by Bertram and Sarachik [31], Malakhov [189] and others.

[6]Throughout this chapter we shall consider stability and asymptotic stability in the weak sense (compare Chap. 5, where stability in the strong sense will be discussed).

whenever $t \geq t_0$ and $|x_0| < r$.

4. *Asymptotically p-stable* if it is p-stable and for sufficiently small values of $|x_0|$

$$\mathbf{E}|x(t, \omega, t_0, x_0)|^p \to 0 \quad \text{as } t \to \infty.$$

5. *Stable in probability in the large* if it is stable in probability and if furthermore for every $x_0, \varepsilon > 0$ and $\delta > 0$, there exists a $T = T(x_0, \varepsilon, \delta)$ such that (1.63) is valid for all $t > T$. A similar definition obtains for asymptotic stability in probability and p-stability in the large.

6. *Exponentially p-stable* if there exist constants $A > 0$ and $\alpha > 0$ such that

$$\mathbf{E}|x(t, \omega, x_0, t_0)|^p \leq A|x_0|^p \exp\{-\alpha(t - t_0)\}.$$

7. *Almost surely stable* in any of the above senses if almost all sample functions i.e. all, except those from some set of probability 0, are stable in the appropriate sense.

It follows from Chebyshev's inequality that (asymptotic) p-stability of the trivial solution for any value of $p > 0$ implies its (asymptotic) p-stability for every smaller value of $p > 0$ and stability in probability. On the other hand, one can easily show by an example that a solution could be (asymptotically) p-stable for some p and not (asymptotically) p-stable for $p_1 > p$ (see below, Sect. 1.6).

The case most often discussed in the literature is asymptotic p-stability for $p = 2$. Henceforth we shall refer to it as mean square stability.

Unless certain restrictive assumptions are made concerning a given system, it is not likely that non-trivial and effective stability conditions can be found. For example, in [31] stability conditions are given in terms of a Lyapunov function $V(x, t) \geq 0$ such that $\mathbf{E}\dot{V}(x, t) < 0$ where $\dot{V}$ denotes the derivative with respect to (1.61). However, in order to calculate the expectation $\mathbf{E}\dot{V}(x, t)$ one must solve the system (1.61) with a suitable initial condition, and this limits the practical use of the criterion.

Here we shall limit ourselves to stability conditions for systems of the type $(x \in \mathbb{R}^l)$

$$\left. \begin{array}{l} \dfrac{dx}{dt} = F(x, t) + \sigma(x, t)\xi(t, \omega), \\[2mm] F(0, t) \equiv 0, \quad \sigma(0, t) \equiv 0. \end{array} \right\} \tag{1.64}$$

Sufficient conditions for stability will be given in terms of the existence of a Lyapunov function for the truncated system

$$\frac{dx}{dt} = F(x, t). \tag{1.65}$$

We shall assume throughout this section that all Lyapunov functions under consideration are positive definite uniformly in t, i.e.,

$$\inf_{t > 0, \, |x| > r} V(x, t) = V_r > 0 \quad \text{for } r > 0. \tag{1.66}$$

We set

$$B = \sup_{t>0,\, x_i \in \mathbb{R}^l} \frac{|V(x_2,t) - V(x_1,t)|}{|x_2 - x_1|}.$$

Theorem 1.12 *Suppose that there exists a Lyapunov function $V(x,t) \in \mathbf{C}_0$ for the system* (1.65) *satisfying condition* (1.66) *and the conditions*

$$V(0,t) \equiv 0,$$

$$\frac{d^0 V}{dt} \leq -c_1 V, \qquad \|\sigma\| \leq c_2 V \tag{1.67}$$

($c_1, c_2 > 0$ are constants).

 Suppose moreover that the process $|\xi(t,\omega)|$ satisfies the law of large numbers (1.1) *and the condition*

$$\sup_{t>0} \mathbf{E}|\xi(t,\omega)| < \frac{c_1}{Bc_2}. \tag{1.68}$$

Then the trivial solution of the system (1.64) *is asymptotically stable in probability in the large. If the process $|\xi(t,\omega)|$ satisfies the strong law of large numbers* (1.2), *while all the other assumptions remain unchanged, then the solution $x = 0$ is almost surely asymptotically stable in the large.*

Proof By Lemma 1.3, it follows from (1.67) that

$$\frac{dV(x(t,\omega),t)}{dt} \leq -c_1 V(x(t,\omega),t) + Bc_2|\xi(t,\omega)|V.$$

We may assume without loss of generality that $t_0 = 0$. Applying Lemma 1.1, we get the estimate

$$V(x(t,\omega),t) \leq V(x_0,0) \exp\left\{ \int_0^t (Bc_2|\xi(s,\omega)| - c_1)\,ds \right\}$$

$$\leq V(x_0,0) \exp\left\{ Bc_2\left(\frac{1}{t}\int_0^t |\xi(s,\omega)|\,ds - \frac{c_1}{Bc_2} \right)t \right\}. \tag{1.69}$$

Now let $\varepsilon > 0$ and $\delta > 0$ be arbitrary. Using (1.68) and the fact that the process $|\xi(t,\omega)|$ satisfies the law of large numbers, we see that there exists a number $T > 0$ such that for $t \geq T$

$$\mathbf{P}\left\{ \frac{1}{t}\int_0^t |\xi(s,\omega)|\,ds > \frac{c_1}{Bc_2} \right\} < \varepsilon. \tag{1.70}$$

We now choose a large enough number $M > 1$ such that

$$\mathbf{P}\left\{ Bc_2 \int_0^T |\xi(s,\omega)|\,ds > \ln M \right\} < \varepsilon. \tag{1.71}$$

Finally, we choose r small enough so that, for $|x_0| < r$,

$$V(x_0, 0)M < V_\delta. \tag{1.72}$$

It follows now from the inequalities (1.69)–(1.72), considered separately for $t < T$ and $t \geq T$, that for $|x_0| < r$ and all $t \geq 0$

$$\mathbf{P}\{|x(t, \omega)| > \delta\} \leq \mathbf{P}\{V(x(t, \omega), t) > V_\delta\} \leq \epsilon.$$

Hence, using the relation

$$\mathbf{P}\left\{\frac{1}{t}\int_0^t |\xi(s, \omega)|\, ds > \frac{c_1}{Bc_2}\right\} \to 0 \quad \text{as } t \to \infty,$$

we get the first part of the theorem.

The proof of the second part is analogous. □

Theorem 1.13 *Suppose that there exists a Lyapunov function $V(x, t) \in \mathbf{C}_0$ for the system (1.65), satisfying condition (1.67) and the inequality*

$$V(x, t) > c|x| \quad (c > 0). \tag{1.73}$$

Assume that the process $\xi(t, \omega)$ is such that for some positive constants k_1, k_2 and $t > 0$

$$\mathbf{E}\exp\left\{k_1 \int_0^t |\xi(s, \omega)|\, ds\right\} \leq \exp\{k_2 t\}, \tag{1.74}$$

where the constants k_i, c_i, B satisfy the inequality

$$Bk_2 c_2 \leq k_1 c_1. \tag{1.75}$$

Then the solution $x(t) \equiv 0$ of the system (1.64) is p-stable for $p \leq k_1/Bc_2$. If the strict inequality is valid,

$$Bk_2 c_2 < k_1 c_1, \tag{1.76}$$

then the solution is exponentially p-stable for $p \leq k_1/Bc_2$.

Proof The proof is based on the inequality (1.69). Raising both sides of this inequality to the power k_1/Bc_2 and then calculating the expectation of both sides, we see, using (1.73), that

$$c^{k_1/Bc_2}\mathbf{E}|\xi(t, \omega)|^{k_1/Bc_2} \leq \mathbf{E}[V(x(t, \omega), t)]^{k_1/Bc_2}$$

$$\leq [V(x_0, 0)]^{k_1/Bc_2}\mathbf{E}\exp\left\{k_1 \int_0^t |\xi(s, \omega)|\, ds - \frac{c_1 k_1}{Bc_2}t\right\}.$$

This, together with the inequalities (1.74)–(1.76), implies the assertion. □

Remark 1.8 It is clear from the proof of Theorem 1.12 that we can somewhat weaken the requirement that the process $|\xi(t, \omega)|$ should satisfy the law of large numbers. However, this condition cannot be completely dropped. This can be seen from the example

$$\frac{dx}{dt} = (-a + \xi)x,$$

where $a > 0$ and the random variable ξ can take arbitrarily large positive values. However small the expectation $\mathbf{E}|\xi|$ may be, the solution $x(t, \omega) = x_0 \exp\{(-a + \xi)t\}$ of this equation tends to infinity with probability $p = \mathbf{P}\{\xi > a\}$.

The same example shows that condition (1.74) of Theorem 1.13 cannot be essentially weakened.

Remark 1.9 Theorems 1.12 and 1.13 which furnish conditions for the occurrence of a stable equilibrium are not local in nature since conditions (1.67) imposed on the Lyapunov functions must hold everywhere, not only in the neighborhood of the origin. It is not hard to devise examples of stochastic processes $\xi(t, \omega)$ for which all assumptions of Theorems 1.12 and 1.13 hold locally, but the origin is nevertheless unstable. This is the case, for example, for the equation

$$\frac{dx}{dt} = \frac{-x + x^3}{1 + |x|^3} + x\xi(t, \omega),$$

where the process $\xi(t, \omega)$ vanishes everywhere except on intervals of length $\Delta_k \to 0$, on which it is equal to $1/\Delta_k^2$. Scattering the intervals Δ_k at random and sufficiently sparsely over the t-axis, we can ensure that the law of large numbers and condition (1.68) will hold. Nevertheless, $x(t) \to \infty$ almost surely if $x_0 > 0$.

This example shows that the existence of a local Lyapunov function is not sufficient for stability in probability. It is also clear that these systems do not satisfy the analog of Lyapunov's theorem on stability in the first approximation. As we shall see later, the situation changes radically if $\xi(t, \omega)$ is assumed to be a white Gaussian noise process.

1.6 Stability of Randomly Perturbed Deterministic Systems

The following problem has been considered by several authors. Let $x \equiv 0$ be a stable solution, in some sense, of the equation ($x \in \mathbb{R}^l$)

$$\frac{dx}{dt} = F(x, t) \quad (F(0, t) \equiv 0). \tag{1.77}$$

Will the solution of this system remain in a given neighborhood of the origin for all $t \geq t_0$ if the right-hand side $F(x, t)$ is perturbed, say, by sufficiently small random forces? More precisely, along with the system (1.77), we consider the system

$$\frac{dx}{dt} = F(x, t) + R(x, t) \tag{1.78}$$

and call the solution $x(t) \equiv 0$ of (1.61) stable under continually acting perturbations if, for each $\varepsilon > 0$, there exists a $\delta > 0$ such that if

$$|x_0| < \delta, \qquad |R(x, t)| < \delta,$$

then the solution $x(t, x_0, t_0)$ of the system (1.78) satisfies the inequality $|x(t, x_0, t_0)| < \varepsilon$ for all $t \geq t_0$. It is known [191] that a sufficient condition for stability under continually acting perturbations is that the trivial solution of the system (1.77) be asymptotically stable uniformly in x_0, t_0.

Sometimes, however, assumptions of this kind might be too restrictive. This is the case, for instance, when the right-hand side of the system (1.77) is subjected to random perturbations which are small only on the average, but sometimes, even if only rarely, experience quite significant "overshoots" which begin at a random time and extend over a period which is not necessarily short. It is clear that then restrictions on F only in the neighborhood of the point $x = 0$ will not imply stability of the trivial solution, since the solution may sometimes extend far beyond the origin. The only meaningful definition of stability in such a situation is that at any fixed time the sample function should lie in the neighborhood of the origin with sufficiently high probability.

We now present the rigorous definition.

Along with (1.77), we consider equation

$$\frac{dx}{dt} = F(x, t) + R(x, t, \omega), \tag{1.79}$$

where the function $R(x, t, \omega)$ is such that (1.79) satisfies the existence and uniqueness theorems of Sect. 1.2. We also assume that the stochastic process

$$\eta(t, \omega) = \sup_x |R(x, t, \omega)|$$

has finite expectation. The solution $x \equiv 0$ of the system (1.77) will be called stable for $t \geq t_0$, under continually acting random perturbations which are small *on the average* (briefly: *stable under small random perturbations*) if the solution of (1.79) satisfying the initial condition $x(t_0, \omega) = x_0$ tends to zero in probability uniformly for $t \geq t_0$ as

$$|x_0| + \sup_{t \geq t_0} \mathbf{E}\eta(t, \omega) \to 0. \tag{1.80}$$

In other words, the solution $x \equiv 0$ of the system (1.77) is stable under small random perturbations if, for each $\varepsilon > 0$ and $\Delta > 0$, there exists a $\gamma > 0$ such that, when

$$|x_0| + \sup_{t \geq t_0} \mathbf{E}\eta(t, \omega) < \gamma,$$

then the following inequality holds for $t \geq t_0$:

$$\mathbf{P}\{|x(t, \omega)| > \Delta\} < \varepsilon.$$

Theorem 1.14 *Suppose that there exists a Lyapunov function $V(x,t) \in \mathbf{C}_0$ on E with the following properties*:

1. $V(0,t) \equiv 0$, $V_\delta > 0$ for $\delta > 0$.
2. *For each $\delta > 0$, there exists a $c_\delta > 0$ such that*

$$\frac{d^0 V}{dt} < -c_\delta V \tag{1.81}$$

holds in the domain $\{|x| > \delta\} \times \{t > t_0\}$.

Then the solution $x \equiv 0$ of (1.77) is stable under small random perturbations for $t \geq t_0$.

Proof It follows from the assumptions of the theorem that in the domain $|x| > \delta$

$$\frac{dV(x(t,\omega),t)}{dt} \leq \frac{d^0 V}{dt} + c\eta(t,\omega). \tag{1.82}$$

Set $V^{(\delta)} = \sup_{t>t_0, |x|<\delta} V(x,t)$. Then $V^{(\delta)} \to 0$ as $\delta \to 0$. Moreover, the assumptions of the theorem imply that $d^0 V/dt < 0$ for $x \neq 0$. In view of this inequality and (1.81), it follows that for all x and almost all ω

$$\frac{dV}{dt} \leq -c_\delta V + c\eta(t,\omega) + c_\delta V^{(\delta)}. \tag{1.83}$$

Applying Lemma 1.1 and then taking the expectation of both sides of the inequality (see proof of Theorem 1.8), we easily see that

$$\mathbf{E}V(x(t,\omega),t) \leq V(x_0,t_0)e^{-c_\delta(t-t_0)} + \frac{c}{c_\delta} \sup_{t>t_0} \mathbf{E}\eta(t,\omega) + V^{(\delta)}. \tag{1.84}$$

Now let $\varepsilon > 0$ and $\Delta > 0$ be arbitrary. Taking δ, $|x_0|$ and $\mathbf{E}\eta(t,\omega)$ sufficiently small, we easily get the inequality

$$\mathbf{E}V(x(t,\omega),t) \leq \varepsilon \sup_{t\geq t_0, |x|>\delta} V(x,t). \tag{1.85}$$

This inequality and Lemma 1.4 imply the required assertion. □

Remark 1.10 Let the point $x = 0$ be exponentially stable for the system (1.77), i.e., assume that the solution $x(t,x_0,t_0)$ of the system (1.77) with initial condition $x(t_0) = x_0$ admits the estimate $|x(t,x_0,t_0)| < B|x_0| \exp\{-\alpha(t-t_0)\}$, where $B, \alpha > 0$ are constants independent of x_0 and t_0. Then (see [156, p. 72]) there exists a function $W(x,t)$ for the system (1.77) such that

$$c_1|x|^2 < W(x,t) < c_2|x|^2,$$

$$\frac{d^0 W}{dt} \leq -c_3|x|^2, \qquad \left|\frac{\partial W}{\partial x}\right| < c_4|x|,$$

provided $\|\partial F/\partial x\|$ is bounded in E. It follows from these estimates that the function $V(x, t) = [W(x, t)]^{1/2}$ satisfies all the assumptions of Theorem 1.14. Thus, if the point $x = 0$ is exponentially stable for the system (1.77), then it is also stable under small random perturbations.

Remark 1.11 If the function $V(x, t)$ satisfies the assumptions of Theorem 1.14 and moreover $V(x, t) > c_1|x|$ for some $c_1 > 0$, then it evidently follows from the proof of Theorem 1.14 that the system (1.77) is stable under small random perturbations in a stronger sense. Indeed, we have then

$$\sup_{t>0} \mathbf{E}|x(t, \omega)| \to 0 \quad \text{as } |x_0| + \sup_{t>0} \mathbf{E}\eta(t, \omega) \to 0.$$

The preceding remark shows that the above type of stability (mean stability) holds when the unperturbed system is exponentially stable. It is readily shown that in the latter case we also have mean square stability with respect to random perturbations which are small in mean square, i.e.,

$$\sup_{t>0} \mathbf{E}|x(t, \omega)|^2 \to 0 \quad \text{as } |x_0| + \sup_{t>0} \mathbf{E}\eta^2(t, \omega) \to 0.$$

The derivation of further, more general criteria for mean square stability presents no difficulties.

Remark 1.12 It is clear from Example 1.5 that condition (1.81) cannot be replaced by the condition: $d^0V/dt < -c_\delta$ in the domain $|x| > \delta$. By slightly modifying Example 1.5 one readily shows that even a linear asymptotically stable system may be unstable under perturbations such that only $\sup_{t>0} \mathbf{E}[\eta(t, \omega)]^\alpha \to 0$, if $\alpha < 1$. Thus, even the "best-behaved" stable systems may lose their stability if $\mathbf{E}\eta(t, \omega)$ does not tend to zero (we are not considering here the case of white noise, when $\mathbf{E}\eta(t, \omega)$ does not exist; for this case, see Chap. 5).

The assumptions of Theorem 1.14 may be slightly weakened by further restricting the range of admissible random perturbations. It seems that in most applications it is sufficient to consider random perturbations of the type

$$R(x, t, \omega) = \sigma(x, t)\xi(t, \omega). \tag{1.86}$$

We shall say that the solution $x(t) \equiv 0$ of the system (1.77) is stable under small random perturbations of type (1.86) if, for each $\epsilon > 0$ and $\Delta > 0$, there exists a $\kappa > 0$ such that, whenever

$$|x_0| + \sup_{x,t} \|\sigma(x, t)\| < \kappa, \tag{1.87}$$

then the following inequality holds for $t > t_0$:

$$\mathbf{P}\{|x(t, \omega)| > \Delta\} < \varepsilon.$$

(This definition is natural, since the matrix $\sigma(x, t)$ characterizes the intensity of the random perturbations at the point (x, t).)

Theorem 1.15 *Let $V(x, t) \in \mathbf{C}_0$ be a Lyapunov function in E satisfying assumption 1 of Theorem* 1.14 *and assumption* 2 *with inequality* (1.81) *replaced by*

$$\frac{d^0 V}{dt} < -c_\delta.$$

Assume further that the process $\xi(t, \omega)$ satisfies the following condition: For each $\varepsilon > 0$, there exists a $\gamma > 0$ such that for $t_0 \le s \le t$

$$\mathbf{E} \exp\left\{\gamma \int_s^t |\xi(u, \omega)| \, du\right\} \le e^{\varepsilon(t-s)}. \tag{1.88}$$

Then the solution $x \equiv 0$ of (1.77) *is stable for $t \ge t_0$ under random perturbations of type* (1.86).

Proof We set

$$W(x, t) = \exp\{V(x, t)\} - 1, \qquad W^{(\delta)} = \sup_{U_\delta \times \{t > t_0\}} W(x, t).$$

Proceeding as in the proof of (1.83) we get the estimate

$$\frac{dW}{dt} \le W(-c_\delta + \gamma |\xi(t, \omega)|) + \gamma |\xi(t, \omega)| + c_\delta W^{(\delta)}, \tag{1.89}$$

valid for every $\delta > 0$ and $\gamma > 0$, provided inequality (1.87) holds for $\kappa \le \kappa_0(\gamma)$. Using Lemma 1.1, we infer from (1.89) that

$$
\begin{aligned}
W(x(t, \omega), t) \le \; & W(x_0, t_0) \exp\left\{\int_{t_0}^t (-c_\delta + \gamma |\xi(s, \omega)|) \, ds\right\} \\
& + c_\delta W^{(\delta)} \int_{t_0}^t \exp\left\{\int_s^t (-c_\delta + \gamma |\xi(u, \omega)|) \, du\right\} ds \\
& + \gamma \int_{t_0}^t \exp\left\{\int_s^t (-c_\delta + \gamma |\xi(u, \omega)|) \, du\right\} |\xi(s, \omega)| \, ds. \tag{1.90}
\end{aligned}
$$

Let ε be any number such that $0 < \varepsilon < 1/2$. It follows from (1.88) that we can choose a number $\gamma_0(\varepsilon)$ such that, for $\gamma < \gamma_0(\varepsilon)$,

$$\mathbf{E} \exp\left\{\gamma \int_s^t |\xi(u, \omega)| \, du\right\} \le \exp\{\varepsilon c_\delta(t - s)\}. \tag{1.91}$$

Hence, using the fact that $W^{(\delta)} \to 0$ as $\delta \to 0$ (which follows directly from the assumptions of the theorem) we see that we can choose, first $\delta(\varepsilon)$, and then $\kappa(\varepsilon)$

and $\gamma_0(\varepsilon)$, so that inequality (1.91) will hold, and moreover

$$
\mathbf{E}\left[W(x_0,t_0)\exp\left\{\int_{t_0}^t(-c_\delta+\gamma|\xi(u,\omega)|)\,du\right\}\right.
$$
$$
\left.+c_\delta W^{(\delta)}\int_{t_0}^t\exp\left\{\int_s^t(-c_\delta+\gamma|\xi(u,\omega)|)\,du\right\}ds\right]<\varepsilon. \qquad (1.92)
$$

Next, from the equalities

$$
A=\gamma\int_{t_0}^t\exp\left\{\int_s^t(-c_\delta+\gamma|\xi(u,\omega)|)\,du\right\}|\xi(s,\omega)|\,ds
$$
$$
=\exp\left\{\int_{t_0}^t(-c_\delta+\gamma|\xi(u,\omega)|)\,du\right\}-1
$$
$$
+c_\delta\int_{t_0}^t\exp\left\{\int_s^t(-c_\delta+\gamma|\xi(u,\omega)|)\,du\right\}ds
$$

and from (1.91) it follows that

$$
\mathbf{E}A\leq\exp\{-c_\delta(1-\varepsilon)(t-t_0)\}-1+c_\delta\int_{t_0}^t\exp\{-c_\delta(1-\varepsilon)(t-s)\}\,ds
$$
$$
\leq\frac{2\varepsilon}{1-\varepsilon}\leq 4\varepsilon. \qquad (1.93)
$$

It follows from (1.90), (1.92) and (1.93) that for all $t>t_0$ we have $\mathbf{E}W(x(t,\omega),t)<5\varepsilon$, whenever $|x_0|+\sup_{t>t_0}\mathbf{E}\eta(t,\omega)<\kappa(\varepsilon)$. Now it suffices to apply Lemma 1.4 in order to derive the assertion of the theorem. □

Example 1.9 In the one-dimensional case we may use the Lyapunov function $V(x)=x$ to get the following result. Let $F(x,t)/x<-c_\delta$ hold in the domain $|x|>\delta$ for some $c_\delta>0$. Then the point $x=0$ is stable under small random perturbations. But if, instead, we assume that $F(x,t)\operatorname{sign}x<-c_\delta$ for $|x|>\delta$, then the point $x=0$ is stable under small random perturbations of type (1.86), provided $\xi(t,\omega)$ satisfies condition (1.88).

1.7 Estimation of a Certain Functional of a Gaussian Process

One sees from Theorem 1.15 that the following estimate plays a major role in the theory of stability of stochastic systems:

$$
\mathbf{E}\exp\left\{k_1\int_{t_0}^{t_1}|\xi(s,\omega)|\,ds\right\}\leq\exp\{k_2(t_1-t_0)\},\quad t_1\geq t_0. \qquad (1.94)
$$

Stability of the system under continually acting random perturbations requires that for each $\varepsilon > 0$ there exists a $\gamma > 0$ such that for $t_1 > t_0$

$$\mathbf{E}\exp\left\{\gamma \int_{t_0}^{t_1} |\xi(s,\omega)|\, ds\right\} \leq \exp\{\varepsilon(t_1 - t_0)\}. \tag{1.95}$$

In this section we shall derive simple conditions for these estimates to hold for Gaussian processes.

We shall consider a Gaussian stochastic process $\xi(t,\omega)$ in $\mathbb{R}^l$, i.e., a process all of whose finite-dimensional distributions are Gaussian. Assume in addition that this process is measurable and that its kernel $K(s,t)$ is continuous.[7]

As usual, we define the *trace* of a square matrix $A = ((a_{ij}))$ by

$$\operatorname{tr} A = \sum_{i=1}^{l} a_{ii}.$$

Theorem 1.16 [8] *Assume that the Gaussian process $\xi(t,\omega)$ with zero expectation satisfies*

$$\operatorname{tr} K(s,s) = \mathbf{E}|\xi(s,\omega)|^2 \leq c_1, \tag{1.96}$$

$$\int_0^\infty \|K(s,u)\|\, du = \int_0^\infty \|K(u,s)\|\, du \leq c_2, \tag{1.97}$$

for some $c_i > 0$ and all $s > 0$.

Then the following estimate holds:

$$\mathbf{E}\exp\left\{k_1 \int_{t_0}^{t_1} |\xi(s,\omega)|\, ds\right\} < \exp\left\{k_1\left(\sqrt{c_1} + \frac{k_1 c_2}{2}\right)(t_1 - t_0)\right\}. \tag{1.98}$$

Lemma 1.5 *The process $\xi(t,\omega)$, $t_0 \leq t \leq t_1$, can be expanded in a series*

$$\xi(t,\omega) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, \varphi_k(t)\xi_k(\omega), \tag{1.99}$$

which is almost surely convergent for every t and satisfies Parseval's identity

$$\int_{t_0}^{t_1} |\xi(t,\omega)|^2\, dt = \sum_{k=1}^{\infty} \lambda_k \xi_k^2. \tag{1.100}$$

---

[7] We denote (see Sect. 1.1)

$$K(s,t) = ((K^{ij}(s,t))) = \operatorname{cov}(\xi(s), \xi(t)).$$

[8] Theorem 1.16 generalizes a result of Shur [259].

Here $\varphi_k(t)$ and λ_k are the normalized eigenfunctions and eigenvalues of the integral equation

$$\int_{t_0}^{t_1} K(t,\tau)\varphi(\tau)\,d\tau = \lambda\varphi(t); \tag{1.101}$$

ξ_k are independent Gaussian random variables with zero expectation and unit variance.

Proof The proof follows easily from the expansion of the process $\xi(t,\omega)$ in the series of eigenfunctions of (1.101). The formulas for the Fourier coefficients

$$\sqrt{\lambda_k}\xi_k(\omega) = \int_{t_0}^{t_1} \xi(t,\omega)\varphi_k(t)\,dt,$$

the orthogonality of the $\varphi_k(t)$ and the fact that the process is Gaussian imply that the random variables ξ_k are independent. Identity (1.100) follows from the completeness of the system of eigenfunctions $\varphi_k(t)$. (For more details see, e.g., [233], [92, Chap. 5, Sect. 2].) $\qquad\square$

Lemma 1.6 *The expectation of the functional*

$$\exp\left\{\alpha\int_{t_0}^{t_1} |\xi(t,\omega)|^2\,dt\right\}$$

exists for all $t_0 < t_1$ and for sufficiently small positive α. Moreover we have the representation

$$\mathbf{E}\exp\left\{\alpha\int_{t_0}^{t_1} |\xi(t,\omega)|^2\,dt\right\} = \prod_{k=1}^{\infty} \frac{1}{\sqrt{1-2\alpha\lambda_k}}. \tag{1.102}$$

If also conditions (1.96), (1.97) hold, then for all $t_0 < t_1$

$$\mathbf{E}\exp\left\{\alpha\int_{t_0}^{t_1} |\xi(t,\omega)|^2\,dt\right\} \le \exp\left\{\frac{\alpha c_1}{1-2\alpha c_2}(t_1-t_0)\right\}.$$

Proof By (1.100), (1.96),

$$\sum_{k=1}^{\infty} \lambda_k = \mathbf{E}\int_{t_0}^{t_1} |\xi(u,\omega)|^2\,du \le c_1(t_1-t_0). \tag{1.103}$$

Consequently $\lambda_{\max} = \max_{1\le k<\infty}\lambda_k$ exists; we may assume without loss of generality that $\lambda_{\max} = \lambda_1$.

The identity (1.102) now follows easily from (1.100) and the following relation, valid for $\alpha < 1/(2\lambda_1)$:

$$\mathbf{E}\exp\{\alpha\lambda_k\xi_k^2(\omega)\} = \frac{1}{\sqrt{1-2\alpha\lambda_k}}.$$

Let us now prove that $\lambda_1 = \lambda_{\max} \leq c_2$. In fact, we easily infer from (1.97) that

$$
\lambda_1 = \int_{t_0}^{t_1} \lambda_1(\varphi_1(s), \varphi_1(s)) \, ds = \int_{t_0}^{t_1} \int_{t_0}^{t_1} (K(s,t)\varphi_1(t), \varphi_1(s)) \, ds \, dt
$$

$$
\leq \int_{t_0}^{t_1} \int_{t_0}^{t_1} \|K(s,t)\| |\varphi_1(t)| |\varphi_1(s)| \, ds \, dt
$$

$$
\leq \int_{t_0}^{t_1} \int_{t_0}^{t_1} \|K(s,t)\| \frac{|\varphi_1(t)|^2 + |\varphi_1(s)|^2}{2} \, ds \, dt \leq c_2. \tag{1.104}
$$

Next, using the elementary inequality $1 + \gamma < e^\gamma$ $(\gamma > 0)$, we get the estimate

$$
\prod_{k=1}^{\infty} \frac{1}{\sqrt{1 - 2\alpha\lambda_k}} = \prod_{k=1}^{\infty} \left(1 + 2\alpha\lambda_k + \frac{4\alpha^2 \lambda_k^2}{1 - 2\alpha\lambda_k} \right)^{1/2}
$$

$$
\leq \exp\left\{ \alpha \left(1 + \frac{2\alpha\lambda_1}{1 - 2\alpha\lambda_1} \right) \sum_{k=1}^{\infty} \lambda_k \right\} = \exp\left\{ \frac{\alpha}{1 - 2\alpha\lambda_1} \sum_{k=1}^{\infty} \lambda_k \right\}.
$$

Hence, in view of (1.103) and (1.104), we get the second assertion of the lemma. $\square$

Proof of Theorem 1.16   Applying the inequality

$$
a < \frac{\alpha}{2} a^2 + \frac{1}{2\alpha} \quad (\alpha > 0), \tag{1.105}
$$

we get

$$
\mathbf{E} \exp\left\{ k_1 \int_{t_0}^{t_1} |\xi(s,\omega)| \, ds \right\} \leq \exp\left\{ \frac{k_1}{2\alpha}(t_1 - t_0) \right\} \mathbf{E} \exp\left\{ \frac{k_1 \alpha}{2} \int_{t_0}^{t_1} |\xi(s,\omega)|^2 \, ds \right\}.
$$

Hence, it follows by Lemma 1.6 that for all $\alpha < 1/k_1 c_2$

$$
\mathbf{E} \exp\left\{ k_1 \int_{t_0}^{t_1} |\xi(s,\omega)| \, ds \right\} \leq \exp\left\{ \left[\frac{k_1}{2\alpha} + \frac{k_1 \alpha c_1}{2(1 - \alpha k_1 c_2)} \right] (t_1 - t_0) \right\}. \tag{1.106}
$$

Setting $\alpha = \alpha^* = 1/(k_1 c_2 + \sqrt{c_1})$, we get (1.98).[9]
     This completes the proof of Theorem 1.16. $\square$

Corollary 1.2 *One can easily eliminate the condition* $\mathbf{E}\xi(t,\omega) = 0$ *in the Theorem,
if one knows that*

$$
|\mathbf{E}\xi(t,\omega)| \leq c_0. \tag{1.107}
$$

---

[9]It is clear that when $\alpha = \alpha^*$ the argument of the exponential function in (1.106) attains its minimum.

Indeed, it is easy to see that estimates similar to (1.94) and (1.95) hold if condition $\mathbf{E}\xi(t, \omega) = 0$ is replaced by (1.107). More precisely, the following analogue of Theorem 1.16 is valid.

If $\xi(t, \omega)$ is a Gaussian process satisfying conditions (1.96), (1.97) and (1.107), then the following estimate is valid for all $k_1 > 0$ and $-\infty < t_0 < t_1 < \infty$:

$$\mathbf{E}\exp\left\{k_1 \int_{t_0}^{t_1} |\xi(s, \omega)| \, ds\right\} \leq \exp\left\{k_1\left(c_0 + \sqrt{c_1} + \frac{k_1 c_2}{2}\right)(t_1 - t_0)\right\}. \quad (1.108)$$

We shall prove one more relation for Gaussian processes which will be needed for the examples in the next section:

Lemma 1.7 *The Gaussian process $\xi(t, \omega) \in \mathbb{R}^l$ satisfies*

$$\mathbf{E}\exp\left\{\int_{t_0}^{t_1} \xi(t, \omega) \, dt\right\} = \exp\left\{\frac{1}{2} \int_{t_0}^{t_1} \int_{t_0}^{t_1} \widehat{K}(s, t) \, ds \, dt + \int_{t_0}^{t_1} \mathbf{E}\xi(t, \omega) \, dt\right\},$$
$$(1.109)$$

where $\widehat{K}(s, t)$ is the vector with components $K^{(1,1)}(s, t), \ldots, K^{(\ell,\ell)}(s, t)$.

Proof It is easily seen that if η is a Gaussian random variable with zero expectation and variance σ^2, then $\mathbf{E}\exp\eta$ exists; in fact

$$\mathbf{E}\exp\eta = \exp\{\sigma^2/2\}. \quad (1.110)$$

The assertion of the lemma follows immediately from (1.110), if we use the fact that each component of the vector $\int_{t_0}^{t_1} \xi(s, \omega) \, ds$ has a Gaussian distribution. $\qquad\square$

Remark 1.13 It follows from Lemma 1.7 that if $\xi(t, \omega)$ is a Gaussian process in $\mathbb{R}^l$ with zero expectation whose correlation function is nonnegative, then condition (1.97) is necessary for Theorem 1.16 to hold. Indeed, if

$$\int_0^\infty K(s, t) \, dt = \infty,$$

then it is clear from (1.109) that the estimate (1.94) does not hold for any k_1 and k_2.

Remark 1.14 Assume that the process $\xi(t, \omega)$ is stationary. Then condition (1.96) is automatically satisfied and $c_1 = \mathbf{E}|\xi(t, \omega)|^2$. Condition (1.97) can then be replaced by the following: The spectral density $f(\lambda)$ of the process $\xi(t, \omega)$ is bounded in the norm. In fact, we used condition (1.97) only to prove that the greatest eigenvalue of (1.101) is bounded. But for stationary stochastic processes it is known that $\max \lambda_k \leq \sup \|f(\lambda)\|$.

Remark 1.15 The estimates (1.94) and (1.95) remain valid when the process $\xi(t, \omega)$ has finite memory (i.e., there is a τ such that the evolution process up to time t and after time $t + \tau$ are independent for all t), and

$$\mathbf{E}\exp\{u|\xi(t, \omega)|\} < \varphi(u) < \infty.$$

It seems likely that also in the general case one could give conditions in terms of the rate of growth of the function $\varphi(u)$ as $u \to \infty$ and certain "mixing" conditions for the process, which imply (1.94) and (1.102). Questions essentially similar to these arise when one investigates existence conditions for the limit of the configuration integral in statistical physics.

1.8 Linear Systems[10]

We now apply the results of this chapter to linear systems of the type

$$\frac{dx}{dt} = (A(t) + \eta(t, \omega))x + B(t, \omega). \tag{1.111}$$

Without loss of generality, we may assume that the elements of the square matrix $\eta(t, \omega)$ have zero expectation.

We consider first the homogeneous system

$$\frac{dx}{dt} = (A(t) + \eta(t, \omega))x. \tag{1.112}$$

Assume that the deterministic system

$$\frac{dx}{dt} = A(t)x \tag{1.113}$$

is exponentially stable, i.e., every solution $x(t, x_0, t_0)$ of the system admits the estimate

$$|x(t, x_0, t_0)| \leq B|x_0|e^{-\alpha(t-t_0)} \quad (\alpha > 0) \tag{1.114}$$

for $t > t_0$, where the constants B, α are independent of x_0, t_0.

By a well-known theorem of Malkin [191, Sect. 75], this implies that there exists a positive definite quadratic form $(C(t)x, x) = W(t, x)$ such that

$$\frac{d^0 W}{dt} \leq -\lambda|x|^2 \quad (\lambda > 0). \tag{1.115}$$

In the sequel we shall frequently use the fact that $d^0 W/dt$ is a quadratic form, being the derivative of a quadratic form along the trajectory of the system (1.113). Moreover

$$\frac{d^0 W}{dt} = \left(\left(CA + A^*C + \frac{\partial C}{\partial t} \right)x, x \right).$$

---

[10]See [43], [125]. One-dimensional linear systems were studied previously in detail in [240], [260] and others.

To be able to apply Theorem 1.12 to the system (1.112), we must write it in the form (1.64). This is easily done if we set $\sigma(x, t)$ equal to the $l \times l^2$ matrix

$$
\sigma(x, t) = \begin{pmatrix} x_1\ x_2\ \dots\ x_l & 0\ 0\ \dots\ 0 & 0\ \dots\ 0 & 0\ 0\ \dots\ 0 \\ 0\ 0\ \dots\ 0 & x_1\ x_2\ \dots\ x_l & 0\ \dots\ 0 & 0\ 0\ \dots\ 0 \\ 0\ 0\ \dots\ 0 & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ 0\ 0\ \dots\ 0 & 0\ 0\ \dots\ 0 & 0\ \dots\ 0 & x_1\ x_2\ \dots\ x_l \end{pmatrix},
$$

and write $\eta_{ik}(t, \omega) = \xi_{(i-1)l+k}(t, \omega)$, where $\xi(t, \omega)$ is a vector in $\mathbb{R}^{l^2}$. Considering the Lyapunov function $V(x, t) = (W(x, t))^{1/2}$ and applying Theorem 1.12, we get the following result.

Theorem 1.17 *Suppose that the solution of the system* (1.113) *satisfies condition* (1.114). *Then there exists a constant $c > 0$ such that the system* (1.112) *is asymptotically stable in probability for every* (*matrix-valued*) *stochastic process $\eta(t, \omega)$ such that the process $\|\eta(t, \omega)\|$ satisfies the law of large numbers and $\mathbf{E}\|\eta(t, \omega)\| < c$. If the process $\|\eta(t, \omega)\|$ also satisfies the strong law of large numbers, the other assumptions remaining unchanged, then the system* (1.112) *is almost surely asymptotically stable in the large.*

Theorem 1.12 also enables us to estimate the constant c. Using the same Lyapunov function and Theorem 1.13, we easily derive sufficient conditions for the system (1.112) to be p-stable, provided the process $\|\eta(t, \omega)\|$ satisfies condition (1.74). Rather than going into further details let us consider an example.

Example 1.10 Consider the following equation in $\mathbb{R}^1$:

$$
\frac{dx}{dt} = (a(t) + \xi(t, \omega))x. \tag{1.116}
$$

It has a solution

$$
x(t) = x_0 \exp\left\{ \int_0^t (a(s) + \xi(s, \omega))\, ds \right\}. \tag{1.117}
$$

Using (1.117) and slightly modifying the proof of Theorem 1.12, we get the following result.

The solution $x(t) \equiv 0$ of (1.116) is asymptotically stable in probability for $a(t) \leq a_0 < 0$ and unstable for $a(t) \geq a_0 > 0$ for every stochastic process $\xi(t, \omega)$ with zero expectation and satisfying the law of large numbers.

Applying Lemma 1.7 to a stationary Gaussian process $\xi(t, \omega)$ with $\mathbf{E}\xi(t, \omega) = 0$ and correlation function $K(t - s)$, we get

$$
\mathbf{E}|x(t, \omega)|^p = |x_0|^p \exp\left\{ p \int_0^t a(s)\, ds + \frac{p^2}{2} \int_0^t \int_0^t K(u - s)\, du\, ds \right\}. \tag{1.118}
$$

Assume that the function $K(u)$ is absolutely integrable and

$$\int_{-\infty}^{+\infty} K(u)\,du = f(0) > 0.$$

Then it is obvious that

$$\int_0^t \int_0^t K(u-s)\,du\,ds = f(0)t + o(t) \quad (t \to \infty). \tag{1.119}$$

Let $a(t) \le a_0 < 0$. It then follows from (1.118) and (1.119) that the solution $x(t) \equiv 0$ of (1.116) is asymptotically p-stable for a Gaussian process $\xi(t, \omega)$ if

$$p < -2a_0/f(0) = p_0$$

and not p-stable if $p > p_0$ and $a(t) = a_0 = \text{const}$.

We now consider the case $a(t) = 0$. It is clear from (1.117) that the solution $x(t) \equiv 0$ is unstable in this case if there exists a function $\alpha(t)$ such that $\alpha(t) \to \infty$ as $t \to \infty$ and the probability

$$\mathbf{P}\left\{\int_0^t \xi(s, \omega)\,ds > \alpha(t)\right\}$$

does not tend to zero. For example this is so when the central limit theorem is applicable to the integral of the process $\xi(t, \omega)$. Fairly broad conditions under which the central limit theorem is applicable in this situation may be found, e.g., in [241].

In the next theorem we adopt the simplifying assumption that A is a constant stable matrix, i.e., a matrix with eigenvalues λ_i such that $\operatorname{Re} \lambda_i < 0$.

Lyapunov showed that, given a stable matrix, one can determine a positive definite matrix C such that the matrix $CA + A^*C$ is negative definite. Let λ denote the greatest positive number such that

$$((CA + A^*C)x, x) \le -\lambda(Cx, x) \tag{1.120}$$

for all $x \in \mathbb{R}^l$. It is not difficult to estimate the number λ from below in terms of the eigenvalues of the matrices C and $CA + A^*C = D$. Let $\lambda_{\max}^C$ and $\lambda_{\max}^D$ denote the greatest eigenvalues of the matrices C and D. Then it is evident that $\lambda > -\lambda_{\max}^D/\lambda_{\max}^C > 0$.

Theorem 1.18 *Let A be a stable $l \times l$ matrix, C a positive definite matrix satisfying condition (1.120), and $\eta(t, \omega) = ((\eta_{ij}(t, \omega)))$, $i, j = 1, \ldots, l$, a Gaussian process. Suppose that the following conditions hold for the process $\tilde{\eta}(t, \omega) = C^{1/2}\eta(t, \omega)C^{-1/2}$:*

$$\|\mathbf{E}\tilde{\eta}(t, \omega)\| \le c_0, \qquad \mathbf{E}\|\tilde{\eta}(t, \omega) - \mathbf{E}\tilde{\eta}(t, \omega)\|^2 \le c_1,$$

$$\int_{-\infty}^{+\infty} \|K(u, s)\|\,du \le c_2$$

(here $K(s, t) = \operatorname{cov}(\tilde{\eta}(s, \omega), \tilde{\eta}(t, \omega))$ is an $l^2 \times l^2$ matrix).

Then the trivial solution of the system

$$\frac{dx}{dt} = (A + \eta(t, \omega))x \qquad (1.121)$$

is asymptotically p-stable for

$$p < \frac{\lambda - 2(c_0 + \sqrt{c_1})}{2c_2},$$

provided $\lambda > 2(c_0 + \sqrt{c_1})$.

Proof We consider the Lyapunov function $V(x) = (Cx, x)$. Using (1.120) and the estimate

$$(C\eta x, x) = (C^{1/2}\eta C^{-1/2}C^{1/2}x, C^{1/2}x) < \|C^{-1/2}\eta C^{1/2}\| \|C^{1/2}x\|^2 = \|\tilde{\eta}\|(Cx, x),$$

we get

$$\frac{dV(x(t, \omega))}{dt} \leq -\lambda V + ((C\eta + \eta^* C)x, x) \leq V(-\lambda + 2\|\tilde{\eta}\|). \qquad (1.122)$$

Therefore

$$[V(x(t, \omega))]^p \leq [V(x_0)]^p \exp\left\{-p\lambda t + 2p \int_0^t \|\eta(s, \omega)\| \, ds\right\}.$$

Hence, calculating expectations and applying the estimate (1.108), we find that

$$\mathbf{E}[V(x(t, \omega))]^p \leq [V(x_0)]^p \exp\{pt(-\lambda + 2c_0 + 2\sqrt{c_1} + 2pc_2)\}. \qquad (1.123)$$

Inequality (1.123) implies directly the assertion of the theorem.   □

Remark 1.16 It is clear from Example 1.10 that in the one-dimensional case condition (1.120) is sufficient for the system to be stable for every value of $c_1 = \sup \mathbf{E}\|\eta(t, \omega)\|^2$, if $c_0 = 0$. It is easy to find examples showing that in the multi-dimensional case this is generally not true: noise of sufficiently high intensity may "overcome" the stability. The only general conclusion from Theorem 1.18 is that when the coefficients of an exponentially stable linear system are perturbed by Gaussian processes satisfying (1.96), (1.97) and (1.107) with sufficiently small c_0 and c_1, the resulting system is asymptotically p-stable.

Remark 1.17 In the one-dimensional case an unstable system driven by Gaussian noise with zero expectation remains unstable. Again, this property does not carry over to multi-dimensional systems. In Chap. 6 we shall present examples of unstable deterministic systems which can be "stabilized" by specially selected Gaussian noise with zero expectation.

Applying Theorem 1.14, we see that an exponentially stable linear system is stable under small random perturbations. It follows from Theorem 1.8 that the system

(1.111) with $\eta(t, \omega) \equiv 0$ is dissipative for every vector $B(t, \omega)$ with finite expectation, provided condition (1.114) (or (1.120)) is satisfied. Let us devote some attention to the case $\eta(t, \omega) \not\equiv 0$, again assuming for simplicity's sake that the process $\eta(t, \omega)$ is Gaussian and the matrix A constant.

Theorem 1.19 *Let A and $\eta(t, \omega)$ satisfy the assumptions of Theorem 1.18 and let $B(t, \omega)$ be a stochastic process with values in $\mathbb{R}^l$, independent of $\eta(t, \omega)$ and with bounded second moment. Then the system*

$$\frac{dx}{dt} = (A + \eta(t, \omega))x + B(t, \omega) \tag{1.124}$$

is dissipative if

$$2c_0 + 2\sqrt{c_1} + 2c_2 < \lambda. \tag{1.125}$$

Proof Setting $V(x) = (Cx, x)$ and using arguments similar to those which yielded (1.122), we infer by means of (1.105) that

$$\frac{dV(x(t, \omega))}{dt} \leq (-\lambda + 2\|\tilde{\eta}\|)V + 2(Cx, B(t, \omega))$$

$$\leq (-\lambda + 2\|\tilde{\eta}\| + \alpha)V + \frac{\|C^{1/2}\|^2 |B(t, \omega)|^2}{\alpha}$$

(where α is any positive number).

Hence, using Lemma 1.1 and the independence of the processes η and B, we can show by standard arguments that

$$\mathbf{E}V(x(t, \omega))$$

$$\leq \mathbf{E}\left\{ V(x_0(\omega)) \exp\left\{ \int_0^t (-\lambda + a + 2\|\tilde{\eta}(s, \omega)\|) \, ds \right\} \right.$$

$$\left. + \frac{\|C^{1/2}\|^2}{\alpha} \int_0^t \mathbf{E} \exp \int_s^t (-\lambda + \alpha + 2\|\tilde{\eta}(u, \omega)\|) \, du \right\} \mathbf{E}|B(s, \omega)|^2 \, ds.$$

If $x_0(\omega)$ satisfies condition (1.42) and the constant α is smaller than $\lambda - 2c_0 - 2\sqrt{c_1} - 2c_2$, then, again using the estimate (1.108), we readily see that for some constant $c_3 > 0$

$$\mathbf{E}V(x(t, \omega)) < c_3.$$

Hence it follows that the process $x(t, \omega)$ is dissipative and its second moment is bounded. $\square$

Remark 1.18 The assertion of Theorem 1.19 may also hold for certain dependent processes $\eta(t, \omega)$ and $B(t, \omega)$. Indeed, before calculating the expectation in the proof one can estimate the expression $\exp\{2 \int_s^t \|\eta(u, \omega)\| \, du\} |B(s, \omega)|^2$, using Young's inequality (1.39). Of course, to continue the proof it is necessary to assume

that $B(t, \omega)$ has bounded moments of order higher than 2. Also, it will be necessary to replace condition (1.125) by a somewhat more stringent condition. We omit the details.

We conclude this chapter by considering one more example.

Example 1.11 Let $\eta_0, \ldots, \eta_n, \ldots$ be a sequence of mutually independent and identically distributed $l \times l$ random matrices, $\mathbf{E}\eta_k = ((0))$, and let A be a constant matrix. Assume that $x(0) = x_0$ and let the evolution of the system on the interval $t \in [k, k+1]$ be governed by the equation

$$\frac{dx(t, \omega)}{dt} = (A + \eta_k(\omega))x(t, \omega). \tag{1.126}$$

Denote $x_k(\omega) = x(k, \omega)$. It is obvious that

$$x_k(\omega) = e^{A+\eta_{k-1}} e^{A+\eta_{k-2}} \cdots e^{A+\eta_0} x_0. \tag{1.127}$$

By virtue of (1.127) the question as to whether the system (1.126) is stable or unstable reduces to the question as to whether the norm of a product of random matrices does or does not tend to zero.

But although various authors have investigated the limiting behavior of a product of random matrices, at this stage it does not appear possible to derive convenient conditions for the stability of the above system.

Assume that the matrix A is stable, and let C be a positive definite matrix satisfying condition (1.120). Applying the method used to prove Theorem 1.18, we get the results below.

If

$$\mathbf{E}\|C^{1/2}\eta_k C^{-1/2}\| < \frac{\lambda}{2}, \tag{1.128}$$

then the stochastic process $x(t, \omega)$ determined by the system (1.126) is almost surely asymptotically stable.

We shall now show that if condition (1.128) is valid, the process $x(t, \omega)$ is asymptotically p-stable for sufficiently small p, provided

$$f(\alpha) = \mathbf{E}\exp\{\alpha\|\eta_k\|\} < \infty \tag{1.129}$$

for sufficiently small $\alpha > 0$.

Lemma 1.8 *Let ξ be a positive random variable such that $\mathbf{E}\exp\{\alpha_0\xi\}$ exists for some $\alpha_0 > 0$. Then*

$$\mathbf{E}\exp\{\alpha\xi\} < \exp\left\{\alpha\mathbf{E}\xi + \frac{\alpha^2}{2}(\mathbf{E}\xi^2 + \varphi(\alpha))\right\} \tag{1.130}$$

holds for sufficiently small α with $\varphi(\alpha) = O(\alpha)$ for $\alpha \to 0$.

The proof of this lemma follows from the inequalities

$$\mathbf{E}\exp\{\alpha\xi\} = \sum_{n=0}^{\infty}\frac{\alpha^n\mathbf{E}\xi^n}{n!} = 1 + \alpha\mathbf{E}\xi + \frac{\alpha^2}{2}\mathbf{E}\xi^2 + \sum_{n=3}^{\infty}\left(\frac{\alpha}{\alpha_0}\right)^n\frac{\alpha_0^n\mathbf{E}\xi^n}{n!}$$

$$\leq 1 + \alpha\mathbf{E}\xi + \frac{\alpha^2}{2}\mathbf{E}\xi^2 + \frac{c\alpha^3}{\alpha_0^3},$$

where $\alpha < \alpha_0$, and the inequality $1 + \gamma < \exp\gamma$ $(\gamma > 0)$.

Suppose that condition (1.129) is satisfied. Then it is obvious that for sufficiently small α the matrices $\tilde{\eta}_k = C^{1/2}\eta_k C^{-1/2}$ also satisfy this condition. On the other hand, from (1.122) we infer that the following estimate is valid when $k - 1 < t < k$ (where k is an integer):

$$(Cx(t,\omega), x(t,\omega))^p \leq (Cx_0, x_0)^p \exp\{-p\lambda(k-1) + 2p(\|\tilde{\eta}_1\| + \cdots + \|\tilde{\eta}_k\|)\}.$$

Since the random variables $\|\tilde{\eta}_i\|$ are independent, it follows from (1.129) and (1.130) that for sufficiently small p

$$\mathbf{E}(Cx(t,\omega), x(t,\omega))^p \leq (Cx_0, x_0)^p \exp\{-p\lambda(k-1) + 2pk\mathbf{E}\|\tilde{\eta}_1\| + kp^2c_1\}.$$

This inequality and (1.128) imply that if conditions (1.128) and (1.129) are satisfied, then the process $x(t,\omega)$ determined by the system (1.126) is asymptotically p-stable for sufficiently small p.

Other conditions for the stability of the system (1.126), based on the fact that the process $x_0, x_1(\omega), \ldots, x_k(\omega), \ldots$ is a Markov chain, will be given in Sects. 6.6 and 6.7.

Chapter 2
Stationary and Periodic Solutions of Differential Equations

2.1 Stationary and Periodic Stochastic Processes. Convergence of Stochastic Processes

A stochastic process $\xi(t) = \xi(t, \omega)$ $(-\infty < t < \infty)$ with values in $\mathbb{R}^l$ is said to be stationary (in the strict sense) if for every finite sequence of numbers $t_1, \ldots, t_n$ the joint distribution of the random variables $\xi(t_1 + h), \ldots, \xi(t_n + h)$ is independent of h. If we replace the arbitrary number h by a multiple of a fixed number θ, $h = k\theta$ ($k = \pm 1, \pm 2, \ldots$), we get the definition of a periodic stochastic process with period θ, or a θ-periodic stochastic process.[1] Stationary and periodic stochastic processes constitute a mathematical idealization of physical noise acting on linear and nonlinear devices functioning in a medium with unvarying or periodically varying properties.

Let $\xi(t)$ be a stationary stochastic process with finite variance. By the definition of stationarity,

$$\mathbf{E}\xi(t) = m = \text{const}, \qquad \text{var}\,\xi(t) = \mathbf{D} = \text{const},$$
$$K(s, t) = \text{cov}(\xi(s), \xi(t)) = K(t - s). \tag{2.1}$$

As already mentioned in Chap. 1, a process satisfying conditions (2.1) is said to be stationary in the wide sense. An important characteristic of stationary processes is their spectral density (see Sect. 1.1).

If $\xi(t)$ is a θ-periodic stochastic process, then $\mathbf{E}\xi(t) = m(t)$ and $\text{var}\,\xi(t) = D(t)$ are evidently periodic with the same period, i.e.

$$m(t + \theta) = m(t), \qquad \mathbf{D}(t + \theta) = \mathbf{D}(t). \tag{2.2}$$

---

[1] There is an enormous literature on the properties of stationary stochastic processes. Among others, we might mention the paper [283] and the books [56], [241], [99]. The properties of periodic processes to be discussed below may be found, e.g., in a paper [57] and in [254].

The matrix-valued function $K(s, t)$ satisfies then the condition

$$K(s + \theta, t + \theta) = K(s, t) \qquad (2.3)$$

for all s, t. A process whose moments satisfy (2.2) and (2.3) is said to be periodic in the wide sense.

It is obvious that a stationary process is periodic with arbitrary period. Conversely, a periodic process can be made stationary by a simple transformation (shift of the argument). Indeed, if τ is a random variable uniformly distributed on the interval $[0, \theta]$ and independent of the θ-periodic stochastic process $\xi(t)$, then the process $\eta(t) = \xi(t + \tau)$ is stationary. To prove this it suffices to observe that for every $t_1, \ldots, t_n, A_1, \ldots, A_n$ the function $\mathbf{P}\{\xi(t_1 + h) \in A_1, \ldots, \xi(t_n + h) \in A_n\}$ is θ-periodic with respect to h, and therefore, for every h,

$$\mathbf{P}\{\eta(t_1 + h) \in A_1, \ldots, \eta(t_n + h) \in A_n\}$$

$$= \frac{1}{\theta} \int_0^\theta \mathbf{P}\{\xi(t_1 + s + h) \in A_1, \ldots, \xi(t_n + s + h) \in A_n\} \, ds$$

$$= \frac{1}{\theta} \int_0^\theta \mathbf{P}\{\xi(t_1 + s) \in A_1, \ldots, \xi(t_n + s) \in A_n\} \, ds$$

$$= \mathbf{P}\{\eta(t_1) \in A_1, \ldots, \eta(t_n) \in A_n\}.$$

It is easily verified that by averaging the moments of the process $\xi(t)$ over the period we obtain the corresponding moments of the process $\eta(t)$. For example,

$$\mathbf{E}\eta(t) = \frac{1}{\theta} \int_0^\theta \mathbf{E}\xi(s) \, ds,$$

$$\mathrm{cov}(\eta(s), \eta(t)) = \frac{1}{\theta} \int_0^\theta \mathrm{cov}(\xi(s + h), \xi(t + h)) \, dh.$$

It is evident that a deterministic periodic function can be regarded as a periodic stochastic process. After a suitable shift of the argument we get a stationary process.

Let $f(t, x)$ be a Borel-measurable function, θ-periodic in t, and $\xi(t)$ a θ-periodic stochastic process. It is then readily seen that the process $f(t, \xi(t))$ is also θ-periodic. For example, if τ is a random variable uniformly distributed on the interval $[0, 2\pi]$, then the process $\xi \sin(t + \tau)$ is stationary for every random variable ξ independent of τ, while the process $\xi \cos t \sin(t + \tau)$ is 2π-periodic. The sample functions of the processes in these examples are periodic. It is easy to construct also examples of periodic processes which almost surely have no periodic sample functions (paths).

In this chapter we shall frequently have to deal with sequences of random variables and with stochastic processes converging in various senses. Therefore let us first recall various definitions of convergence and some results connected with them.[2]

---

[2] See [232], [251], [92].

A sequence of measures $\{\mu_n\}$ defined in $(\mathbb{R}^l, \mathfrak{B})$ is said to be *weakly convergent* to a measure μ if

$$\lim_{n \to \infty} \int_{\mathbb{R}^l} f(x)\mu_n(dx) = \int_{\mathbb{R}^l} f(x)\mu(dx)$$

for every continuous and bounded function $f(x)$ on $\mathbb{R}^l$.

A sequence of random variables ξ_n is *weakly convergent* to ξ if the sequence of measures $\mathbf{P}_n(A) = \mathbf{P}\{\xi_n \in A\}$ converges weakly to the measure $\mathbf{P}(A) = \mathbf{P}\{\xi \in A\}$.

A sequence of random variables ξ_n is said to be *weakly compact* if it contains a weakly convergent subsequence. A sufficient condition for a sequence ξ_n to be weakly compact is that the random variables ξ_n be uniformly bounded in probability, i.e.,

$$\sup_n \mathbf{P}\{|\xi_n| > R\} \to 0 \quad \text{as } R \to \infty.$$

A sequence ξ_n is said to *converge in probability* to ξ if $\mathbf{P}\{|\xi_n - \xi| > \delta\} \to 0$ as $n \to \infty$ for each $\delta > 0$.

Given a sequence ξ_n which converges weakly to ξ_0, one can construct a sequence $\tilde{\xi}_n$ ($n = 0, 1, 2, \dots$) in another probability space $(\widetilde{\Omega}, \widetilde{\mathfrak{A}}, \widetilde{\mathbf{P}})$ such that $\tilde{\xi}_n \to \tilde{\xi}_0$ in probability and the variables ξ_n and $\tilde{\xi}_n$ have the same distribution function for every $n \geq 0$. Skorokhod [251] has generalized these results to stochastic processes as follows.

Theorem 2.1 *Let $\xi_n(t, \omega)$ ($n = 1, 2, \dots$) be a sequence of stochastic processes in $\mathbb{R}^l$ such that for every $t_1, \dots, t_k$ the joint distribution of $\xi_n(t_1), \dots, \xi_n(t_k)$ is weakly convergent to some limit, and the sequence $\xi_n(t)$ is uniformly stochastically continuous, i.e.,*

$$\sup_{n, |s_1 - s_2| < h} \mathbf{P}\{|\xi_n(s_1) - \xi_n(s_2)| > \varepsilon\} \xrightarrow[h \to 0]{} 0. \tag{2.4}$$

Then one can construct a sequence of stochastic processes $\tilde{\xi}_n(s)$ ($n = 0, 1, 2, \dots$) in another probability space $(\widetilde{\Omega}, \widetilde{\mathfrak{A}}, \widetilde{\mathbf{P}})$ such that the process $\tilde{\xi}_0(s)$ is stochastically continuous, $\tilde{\xi}_n(s) \to \tilde{\xi}_0(s)$ in probability for all s and the finite-dimensional distributions of the processes $\xi_n(s)$ and $\tilde{\xi}_n(s)$ coincide for $n > 0$.

Theorem 2.2 *A sufficient condition for a sequence of stochastic processes $\xi_n(t)$ to contain a subsequence of processes with weakly convergent finite-dimensional distributions is that the sequence satisfies condition (2.4) and is uniformly bounded in probability:*

$$\sup_{t, n} \mathbf{P}\{|\xi_n(t)| > R\} \to 0 \quad \text{as } R \to \infty. \tag{2.5}$$

Let the processes $\xi_n(t)$, $\xi(t)$ be continuous on the interval $[a, b]$. Let $\mathcal{C}[a, b]$ denote the space of all continuous functions on $[a, b]$; all the sample functions of the processes $\xi_n(t)$, $\xi(t)$ are almost surely in this class.

A sequence $\xi_n(t)$ is said to be *weakly convergent* to $\xi(t)$ as $n \to \infty$ if for every functional f continuous on $\mathcal{C}[a, b]$

$$\mathbf{E}f(\xi_n(t)) \underset{n \to \infty}{\longrightarrow} \mathbf{E}f(\xi(t)).$$

Prokhorov [232] has proved the following theorem.

Theorem 2.3 *If the finite-dimensional distributions of the processes $\xi_n(t)$ are weakly convergent to some limit and there exist $\alpha > 1$, $k > 0$ and $a > 0$ such that for all t_1, t_2 and n*

$$\mathbf{E}|\xi_n(t_2) - \xi_n(t_1)|^a < k|t_2 - t_1|^\alpha,$$

then the sequence of processes $\xi_n(t)$ is weakly convergent to a process $\xi(t)$ whose finite-dimensional distributions coincide with the above-mentioned limit distributions.

2.2 Existence Conditions for Stationary and Periodic Solutions[3]

An important part of the qualitative theory of differential equations is the study of periodic solutions of systems with periodic right-hand sides.

In a more general setting, this corresponds to the study of existence conditions and properties of periodic and stationary solutions of differential equations whose right-hand side is a periodic or stationary process in t for fixed values of the space variable.

In this section we shall present a general, but not sufficiently effective for applications, solution of this problem. In the next section we shall use this result to derive effective sufficient conditions for the existence of stationary and periodic solutions in terms of auxiliary functions.

Theorem 2.4 *Let $G(x, z)$ $(x \in \mathbb{R}^l$, $z \in \mathbb{R}^k)$ be a function and $\xi(t)$ a stationary stochastically continuous process in $\mathbb{R}^k$, satisfying conditions* (1.23), (1.24). *Then there exists a stationary solution of the equation*

$$\frac{dx}{dt} = G(x, \xi(t)) \tag{2.6}$$

which is stationarily related to $\xi(t)$ if and only if this equation has at least one solution $y(t, \omega)$ satisfying the condition

---

[3]Existence conditions for stationary and periodic solutions of differential equations with random right-hand side have been investigated under different assumptions and by other methods by Vorovich [269] and Dorogovtsev [57].

$$\frac{1}{T} \int_0^T \mathbf{P}\{|y(t, \omega)| > R\} \, dt \to 0 \quad as \ R \to \infty \tag{2.7}$$

uniformly in $T > T_0$ (or $T < -T_0$).

Proof Necessity is obvious, since every stationary solution $y(t, \omega)$ satisfies condition (2.7). To prove sufficiency, we first make the following observation. Solving (2.6) with initial condition $x(0) = x_0(\omega)$ by successive approximations, one may readily verify that the random variable $x(t, \omega)$ is measurable with respect to the minimal σ-algebra containing all possible events $\{\xi(s) \in A_1\}$ ($s \in [0, t]$) and $\{x_0(\omega) \in A_2\}$. Here and below, $A_i \in \mathfrak{B}$, where $\mathfrak{B}$ is the σ-algebra of Borel sets in Euclidean space. Therefore, in order to prove the existence of a stationary process $(X(t), \xi(t))$ satisfying (2.6) it will suffice to show that there exists a random variable $\eta(\omega)$ such that for all $t > 0$, $A_0, A_1, \ldots, A_m, s_1, \ldots, s_m$,

$$\mathbf{P}\{\eta(\omega) \in A_0, \xi(s_1) \in A_1, \ldots, \xi(s_m) \in A_m\}$$
$$= \mathbf{P}\{X(t) \in A_0, \xi(s_1 + t) \in A_1, \ldots, \xi(s_m + t) \in A_m\}, \tag{2.8}$$

where $X(t)$ is the solution of (2.6) with initial condition $x(0) = \eta(\omega)$.

Assume for definiteness that condition (2.7) holds with $T > 0$. Let $\tau_k(\omega)$ be a random variable, uniformly distributed on $[0, k]$ and independent of $\xi(t)$ and $y(0, \omega)$. We set $x_0^{(k)}(\omega) = y(\tau_k(\omega), \omega)$ and

$$x_k(t, \omega) = y(t + \tau_k(\omega), \omega), \qquad \xi_k(t, \omega) = \xi(t + \tau_k(\omega), \omega).$$

It is obvious that

$$\mathbf{P}\{x_k(t) \in A_0, \xi_k(s_1) \in A_1, \ldots, \xi_k(s_m) \in A_m\}$$
$$= \frac{1}{k} \int_0^k \mathbf{P}\{y(t + s) \in A_0, \xi(s_1 + s) \in A_1, \ldots, \xi(s_m + s) \in A_m\} \, ds. \tag{2.9}$$

It follows from (2.9) that for every k the distribution of the process $\xi_k(t)$ is the same as that of the process $\xi(t)$. It also follows from (2.7) that uniformly in $k > 0$,

$$\mathbf{P}\{|x_0^{(k)}(\omega)| > R\} = \frac{1}{k} \int_0^k \mathbf{P}\{|y(t)| > R\} \, dt \xrightarrow[R \to \infty]{} 0. \tag{2.10}$$

By the stochastic continuity of the process $\xi(t)$ and by (2.10), the family $(x_0^{(k)}(\omega), \xi^{(k)}(t, \omega))$ satisfies conditions (2.4) and (2.5). Applying Theorems 2.1 and 2.2, we see that in some probability space $(\widetilde{\Omega}, \widetilde{\mathfrak{A}}, \widetilde{\mathbf{P}})$ there is a sequence $(\tilde{x}_0^{(k)}(\widetilde{\omega}), \tilde{\xi}^{(k)}(t, \widetilde{\omega}))$ with the same distribution as $(x_0^{(k)}(\omega), \xi^{(k)}(t, \omega))$, such that some subsequence $(\tilde{x}_0^{(n_k)}(\widetilde{\omega}), \tilde{\xi}^{(n_k)}(t, \widetilde{\omega}))$ converges in probability to $(\tilde{x}(\widetilde{\omega}), \tilde{\xi}(t, \widetilde{\omega}))$. Obviously, the finite-dimensional distributions of the processes $\tilde{\xi}(t, \widetilde{\omega})$ and $\xi(t, \omega)$ are the same.

We can now construct on the original probability space random variables $x(\omega)$ and $x^{(n_k)}(\omega)$ whose joint distribution with $\xi(t, \omega)$ is the same as the joint distribution of

$$\tilde{x}(\widetilde{\omega}), \quad \tilde{x}_0^{(n_k)}(\widetilde{\omega}), \quad \tilde{\xi}(t, \widetilde{\omega}).$$

We shall prove that (2.8) holds for $\eta(\omega) = x(\omega)$. Let $X_{n_k}(t)$ $(k = 1, 2, \dots)$ denote the solution of (2.6) with initial condition $X_{n_k}(0) = x^{(n_k)}(\omega)$. Now conditions (1.23), (1.24) and the Gronwall–Bellman lemma imply the inequality

$$|X_{n_k}(t) - X(t)| < |x^{(n_k)}(\omega) - x(\omega)| \exp\left\{\int_0^t B(u, \omega) \, du\right\},$$

and so $X_{n_k}(t) \to X(t)$ in probability for every t. Let f be a continuous bounded function. Then, by what we have proved it follows from (2.9) that for each t and $s_1, \dots, s_m$,

$$
\begin{aligned}
\mathbf{E} f(\xi(s_1 + t), &\dots, \xi(s_m + t), X(t)) \\
&= \lim_{k \to \infty} \mathbf{E} f(\xi(s_1 + t), \dots, \xi(s_m + t), X_{n_k}(t)) \\
&= \lim_{k \to \infty} \mathbf{E} f(\xi_{n_k}(s_1 + t), \dots, \xi_{n_k}(s_m + t), x_{n_k}(t)) \\
&= \lim_{k \to \infty} \frac{1}{n_k} \int_0^{n_k} \mathbf{E} f(\xi(s_1 + t + u), \dots, \xi(s_m + t + u), y(t + u)) \, du \\
&= \lim_{k \to \infty} \frac{1}{n_k} \int_0^{n_k} \mathbf{E} f(\xi(s_1 + s), \dots, \xi(s_m + s), y(s)) \, ds \\
&= \mathbf{E} f(\xi(s_1), \dots, \xi(s_m), x(\omega)).
\end{aligned}
\tag{2.11}
$$

This implies (2.8), and hence the assertion of the theorem. $\qquad \square$

The analogous result for a periodic process $\xi(t)$ is given by the following theorem.

Theorem 2.5 *Let $G(x, z)$ $(x \in \mathbb{R}^\ell, z \in \mathbb{R}^k)$ be a given function and $\xi(t)$ a θ-periodic stochastically continuous process in $\mathbb{R}^k$ satisfying conditions (1.23), (1.24). Then there exists a periodic solution of (2.6) which is periodically related to $\xi(t)$ if and only if the equation has at least one solution $y(t, \omega)$ satisfying the condition*

$$\frac{1}{|k| + 1} \sum_{n=0}^{k} \mathbf{P}\{|y(n\theta)| > R\} \to 0 \quad as \ R \to \infty \tag{2.12}$$

uniformly in $k = 1, 2, \dots$ (or $k = -1, -2, \dots$).

Proof The proof is entirely analogous to that of Theorem 2.4. The only difference is that instead of the processes $x_k(t) = y(t + \tau_k)$ one must consider a sequence

$Y_k(t) = y(t + \chi_k)$, where χ_k is a random variable independent of $\xi(t)$ and $y(0, \omega)$ such that $\mathbf{P}\{\chi_k = n\theta\} = 1/(k + 1)$ $(n = 0, 1, \ldots, k)$. $\qquad\square$

As we shall see in the next section, the advantage of condition (2.7) over (2.12) is that it is easier to verify whether (2.7) holds even if no solutions of (2.6) are known. Thus, the following lemma may be sometimes useful.

Lemma 2.1 *Condition* (2.12) *of Theorem* 2.5 *can be replaced by condition* (2.7).

Proof The necessity of condition (2.7) is obvious. Let us prove the sufficiency. Let $y(t) = y(t, \omega)$ be a solution of (2.6) satisfying condition (2.7). Then for each τ, $z(t) = y(t + \tau)$ is a solution of the equation

$$\frac{dz}{dt} = G(z, \xi(t + \tau)). \qquad (2.13)$$

Now let τ be a random variable uniformly distributed on $[0, \theta]$ and independent of the process $\xi(t)$. Then, as shown in Sect. 2.1, $\xi(t + \tau)$ is a stationary process. Moreover, the solution $z(t)$ of (2.13) satisfies condition (2.7), since

$$\frac{1}{T} \int_0^T \mathbf{P}\{|z(t)| > R\} \, dt = \frac{1}{\theta} \int_0^\theta ds \, \frac{1}{T} \int_0^T \mathbf{P}\{|y(t + s)| > R\} \, dt$$

$$\leq \frac{T + \theta}{T} \frac{1}{\theta} \int_0^\theta ds \, \frac{1}{T + s} \int_0^{T+s} \mathbf{P}\{|y(u)| > R\} \, du.$$

Applying Theorem 2.4, we see that (2.13) has a solution $Z_1(t, \omega)$ which is a stationary process. It follows from Theorem 1.5 that

$$\sup_{0 \leq t \leq \theta} |Z_1(t)| \leq |Z_1(0)| + \int_0^\theta |G(Z_1(0), \xi(s + \tau))| \, ds \exp\left\{ \int_0^\theta B(s + \tau, \omega) \, ds \right\}.$$

By conditions (1.23), (1.24) and the stationarity of the process Z_1, the probability of the event

$$\left\{ \sup_{s \leq t \leq s+\theta} |Z_1(t)| > R \right\}$$

is independent of s and

$$\mathbf{P}\left\{ \sup_{s \leq t \leq s+\theta} |Z_1(t)| > R \right\} \to 0 \quad \text{as } R \to \infty. \qquad (2.14)$$

It is clear now that the function $y_1(t, \omega) = Z_1(t - \tau(\omega), \omega)$ is a solution of (2.6). By (2.14), this solution satisfies condition (2.12). Hence, by Theorem 2.5, it follows that (2.6) has a periodic solution. This completes the proof of the lemma. $\qquad\square$

Remark 2.1 The global Lipschitz condition (1.23) is sometimes too restrictive. It can be seen from the proofs of Theorem 2.4 and Lemma 2.1 that this condition is

used only to verify (2.14) and the relation

$$X_{n_k}(t) \to X(t) \quad \text{in probability as } k \to \infty. \tag{2.15}$$

These relations hold if the solutions of (2.6) are uniformly unboundedly continuable in the sense of Remark 1.4 and conditions (1.24), (1.28) are satisfied.

In fact, by conditions (1.24), (1.28) and the Gronwall–Bellman lemma, for every fixed t_0 and all sample functions $X_{n_k}(t, \omega)$, $X(t, \omega)$ satisfying the conditions

$$\sup_{0 \le t \le t_0} |X_{n_k}(t)| \le R, \qquad \sup_{0 \le t \le t_0} |X(t)| \le R, \tag{2.16}$$

we get the inequality

$$|X_{n_k}(t_0) - X(t_0)| \le |x^{(n_k)}(\omega) - x(\omega)| \exp\left\{ \int_0^{t_0} B_R(t, \omega) \, dt \right\}. \tag{2.17}$$

Let $\varepsilon > 0$ be arbitrary. Since the solutions of (2.6) are uniformly unboundedly continuable, the probability of the events (2.16) can be made greater than $1 - \varepsilon/2$ by choosing R sufficiently large. Hence and by considering (2.17) for sufficiently large k we get the inequalities

$$\mathbf{P}\{|X_{n_k}(t_0) - X(t_0)| > \varepsilon\}$$
$$\le \frac{\varepsilon}{2} + \mathbf{P}\left\{ |x^{(n_k)}(\omega) - x(\omega)| \exp\left(\int_0^{t_0} B_R(t, \omega) \, dt \right) > \frac{\varepsilon}{2} \right\} \le \varepsilon.$$

This proves (2.15). The proof of (2.14) is analogous.

This remark, together with Theorem 1.7, implies the following corollaries.

Corollary 2.1 *Let the function $F(x, t), \sigma(x, t)$ and the stochastic process $\xi(t)$ be θ-periodic and satisfy the assumptions of Theorem 1.7. Assume also that the equation $dx/dt = F(x, t) + \sigma(x, t)\xi(t)$ has a solution satisfying condition (2.7). Then this equation also has a θ-periodic solution. Similarly, if F and σ are independent of t and $\xi(t)$ is a stationary process, then the above conditions imply the existence of a stationary solution.*

Corollary 2.2 *Conditions (2.7) and (2.12) are valid if the system (2.6) is dissipative. Therefore, if the system (2.6) is dissipative, $\xi(t, \omega)$ is a stationary (periodic) stochastically continuous process and conditions (1.23), (1.24) are satisfied, then the system (2.6) has a stationary (periodic) solution.*

Example 2.1 Let $G(x, t)$ be a deterministic function which is θ-periodic in t and such that conditions (1.23), (1.24) are satisfied and the equation ($x, G \in \mathbb{R}^l$)

$$\frac{dx}{dt} = G(x, t) \tag{2.18}$$

has at least one bounded solution. It follows from Theorem 2.5 that for some (generally random) initial condition the solution of (2.18) is a periodic stochastic process. For $l \leq 2$ this follows also from a well-known theorem of Massera (see [228, p. 186]). Of course, this result does not guarantee the existence of a deterministic periodic solution of (2.18), since a periodic stochastic process need not have periodic sample functions.

2.3 Special Existence Conditions for Stationary and Periodic Solutions

For systems of the special form

$$\frac{dx}{dt} = F(x,t) + \sigma(x,t)\xi(t) \tag{2.19}$$

one can derive effective conditions which are sufficient for the existence of periodic and stationary solutions.

Theorem 2.6 *Suppose that the vector $F(x,t)$ and the matrix $\sigma(x,t)$ are θ-periodic in t and that they satisfy a local Lipschitz condition; let further $F(0,t) \in \mathbf{L}$ and*

$$\sup_{x,t} \|\sigma(x,t)\| < \infty. \tag{2.20}$$

Assume moreover that the truncated system

$$\frac{dx}{dt} = F(x,t)$$

has a Lyapunov function $V(x,t) \in \mathbf{C}_0$ satisfying the following conditions:

1. *$V(x,t)$ is nonnegative, and*

$$\inf_{t>0} V(x,t) \to \infty \quad as \ |x| \to \infty.$$

2. *d^0V/dt is bounded above, and $\sup_{t>0} d^0V/dt \to -\infty$ as $|x| \to \infty$.*

   Then (2.19) has a θ-periodic solution for each θ-periodic stochastically continuous process $\xi(t)$ with finite expectation. If F and σ independent of t and $\xi(t)$ is a stationary process, then the same conditions imply the existence of a stationary solution.

Proof Let $x(t) = x(t, \omega)$ be a solution of (2.19) satisfying the condition $x(t_0) = x_0$. Using Condition 1 of the theorem, inequality (2.20) and Lemma 1.3, we see that almost surely, for $t > t_0$ and some constant $k > 0$,

$$-V(x_0, t_0) \leq V(x(t), t) - V(x_0, t_0)$$

$$\leq \int_{t_0}^t \frac{d^0 V(x(s), s)}{ds}\, ds + k \int_{t_0}^t |\xi(s)|\, ds. \qquad (2.21)$$

We set $k_1 = \sup_E d^0 V/dt$, $-c_r = \sup_{|x|>r} d^0 V/dt$. It follows from the assumptions of the theorem that

$$k_1 < \infty, \qquad c_r \to \infty \quad \text{as } r \to \infty. \qquad (2.22)$$

Replacing for $|x(s)| > r$ the function $d^0 V/ds$ in (2.21) by the bound $-c_r$ and for $|x(s)| \leq r$ by the bound k_1 and then taking expectations, we get

$$-V(x_0, t_0) \leq -c_r \int_{t_0}^t \mathbf{P}\{|x(s)| > r\}\, ds + k_1(t - t_0) + k \int_{t_0}^t \mathbf{E}|\xi(s)|\, ds.$$

Hence it follows by (2.22) that for some constant k_2

$$\frac{\int_{t_0}^t \mathbf{P}\{|x(s)| > r\}\, ds}{t - t_0} < \frac{k_2}{c_r} \to 0 \quad \text{as } r \to \infty. \qquad (2.23)$$

Condition (2.23) is equivalent to (2.7). Applying Lemma 2.1 and Corollary 2.1, we get the first assertion of the theorem. The second assertion can be proved in the same way. $\qquad\qquad\qquad\square$

Remark 2.2 The assertion of the theorem is valid when the assumption that $\inf_{t>0} V(x, t) \to \infty$ as $|x| \to \infty$ is replaced by the assumption that the solutions of (2.19) are uniformly unboundedly continuable for $t > 0$. It is also sufficient to require that the solutions be unboundedly continuable for $t < 0$ and that the following condition holds: The function $d^0 V/dt$ is bounded below and $d^0 V/dt \to \infty$ as $|x| \to \infty$. (This case reduces to the preceding one if we set $s = -t$.)

Example 2.2 If the system (2.19) is one-dimensional ($x \in \mathbb{R}^1$), then, considering the Lyapunov function $V = |x|$, we have $d^0 V/dt = F(x, t)\, \text{sign}\, x$. Hence Theorem 2.6 and Remark 2.2 yield the following result.

If F and σ are periodic functions of t such that

$$F \in \mathbf{C}_0, \qquad \sigma \in \mathbf{C}_0, \qquad \sup |\sigma| < \infty$$

and either $F(x, t)\, \text{sign}\, x \to -\infty$ or $F(x, t)\, \text{sign}\, x \to \infty$ as $|x| \to \infty$, then (2.19) has a periodic solution in $\mathbb{R}^1$ for every periodic process $\xi(t)$ with bounded expectation. An analogous conclusion holds for stationary solutions as well.

For example, if $f(t)$ is a θ-periodic continuous function and $\xi(t)$ a θ-periodic process, then the equation $dx/dt = xf(t) + \xi(t)$ always has a periodic solution, provided $\mathbf{E}|\xi(t)| < \infty$ and $f(t)$ does not vanish. On the other hand it is obvious that if $F(x, t) > k > -\infty$ (or $F(x, t) < k < \infty$), then (2.19) need not have periodic solutions, since for a suitable choice of σ and ξ the right-hand side of (2.19) will have fixed sign. A more general result is given by the following

Lemma 2.2 *Let $F(x) \in \mathbf{C}$ be a function for which none of the conditions*

$$F(x)\,\text{sign}\,x \xrightarrow[|x|\to\infty]{} \pm\infty \tag{2.24}$$

is valid. Then there exists a stationary stochastic process $\xi(t)$ with finite expectation such that the equation

$$\frac{dx}{dt} = F(x) + \xi(t) \tag{2.25}$$

has no stationary solution.

Proof As already mentioned, the assertion is obvious if the function $F(x)$ is bounded above or below. If it is neither and conditions (2.24) do not hold, then there exist an infinite sequence of points α_k $(k = 1, 2, \ldots)$ and a number c such that $\alpha_k \to \infty$ or $\alpha_k \to -\infty$ as $k \to \infty$, $F(\alpha_k) = c$ and each α_k is a stable equilibrium point of the equation

$$\frac{dx}{dt} = F(x) - c. \tag{2.26}$$

To be more specific, suppose that $\alpha_k \to \infty$. Then the following three cases are possible:

(a) $F(x) \geq c$ for $x < x'$,
(b) $F(x) \leq c$ for $x < x'$,
(c) there exists a sequence $x'_k \to -\infty$ such that $F(x'_k) = c$ and the x'_k are stable equilibrium points of (2.26).

Case (a). We may assume without loss of generality that $x' = \alpha_1$. We claim that in this case (2.25) has no stationary solutions if $\xi(t) = -c + |\eta(t)|$, where $\eta(t)$ is a stationary stochastic process such that for every constant A

$$\mathbf{P}\left\{ \sup_{0 \leq u \leq t} \int_u^{u+1} |\eta(s)|\,ds > A \right\} \to 1 \quad \text{as } t \to \infty. \tag{2.27}$$

(Condition (2.27) holds for instance if $\eta(t)$ is a Gaussian stationary Markov process governed by the generator $d^2/dx - x\,d/dx$.)

Suppose that there exists a stationary process $x(t)$ satisfying (2.25). Since $F(x) \geq c$ for $x < \alpha_1$, the function $x(t)$ is monotone increasing for $x(t) < \alpha_1$, and therefore

$$\mathbf{P}\{x(0, \omega) < \alpha_1\} = 0. \tag{2.28}$$

We shall prove that $\mathbf{P}\{\alpha_1 \leq x(0, \omega) < \alpha_2\} = 0$. To this end, we first observe that, by construction, the points α_k have the following property: once the sample function $x(t_0) = \alpha_k$ at some time t_0, it "cannot" go to the left of α_k for $t > t_0$. Hence in this case it follows from (2.25) that either $X(t + 1) > \alpha_2$ or

$$X(t + 1) - X(t) \geq \int_t^{t+1} |\eta(s)|\,ds + \min_{x \in [\alpha_1, \alpha_2]} (F(x) - c).$$

Hence follows the relation

$$\{x(0,\omega) \geq \alpha_1\} \cap \left\{ \sup_{0 \leq u \leq t-1} \int_u^{u+1} |\eta(s)| \, ds + \min_{x \in [\alpha_1, \alpha_2]} (F(x) - c) > \alpha_2 - \alpha_1 \right\}$$

$$\subset \{x(t,\omega) \geq \alpha_2\}. \tag{2.29}$$

By (2.27), (2.28) and (2.29),

$$\mathbf{P}\{x(0,\omega) \geq \alpha_2\} = \lim_{t \to \infty} \mathbf{P}\{x(t,\omega) \geq \alpha_2\} = 1.$$

Similarly, we show that $\mathbf{P}\{x(0,\omega) \geq a_k\} = 1$ for every k. This contradiction shows that a stationary solution does not exist. The proof for cases (b) and (c) is similar. (In case (b) one sets $\xi(t) = -c - |\eta(t)|$.)   □

Example 2.3 [4] Suppose that for $|x| > x_0$ and some positive integers n and k the coefficients of the equation

$$x'' + f(x)x' + g(x) = \sigma(x, x')\xi(t, \omega) \tag{2.30}$$

satisfy the conditions

$$0 < g(x)/x^{2n+1} < c, \qquad 0 < f(x)/x^{2k} < c,$$

and also the conditions

$$|\sigma| < c; \qquad g(x)F(x) \to \infty \quad \text{as } |x| \to \infty \quad \left(F(x) = \int_0^x f(t) \, dt \right),$$

$$F(x) \operatorname{sign} x > \delta > 0 \quad \text{for } |x| > x_0.$$

Let $\xi(t)$ be a periodic (stationary) stochastic process with finite expectation. Then (2.30) has a periodic (stationary) solution. The proof utilizes Theorem 2.6 applied to the system of equations derived from (2.30), where we set $x' = y$, and take the Lyapunov function

$$V(x, y) = \left[\frac{y^2}{2} + (F(x) - p(x))y + G(x) + \int_0^x f(t)(F(t) - p(t)) \, dt + 1 \right]^\alpha - c_1,$$

with $G(x) = \int_0^x g(t) \, dt$, $p(x) = \gamma \arctan x$, and the positive constants γ, c_1, α so chosen that

$$\min V(x, y) = 0, \qquad V \in \mathbf{C}_0,$$

$$d^0 V/dt \to -\infty \quad \text{as } x^2 + y^2 \to \infty.$$

Note that the conditions of this example hold for a Van der Pol equation in which $f(x) = x^2 - 1$, $g(x) = x$.

---

[4] This example is due to Nevelson.

2.4 Conditions for Convergence to a Periodic Solution

Hitherto we have dealt only with conditions concerning the existence of a periodic (stationary) solution of a differential equation whose right-hand side is a periodic (stationary) process for fixed x. However only those periodic solutions are of practical interest which are stable, in the sense that if the initial conditions lie in a certain class, then the solutions ultimately converge to periodic solutions. In most cases it is sufficient to consider stability for initial conditions which are independent of the right-hand side of the system.

In some cases a periodic solution of a differential equation turns out to be stable in the large, i.e., every solution ultimately converges to a periodic solution. It is clear that if a periodic solution is stable in the large it is unique. These definitions are rather vague, for it is not clear in what sense one should understand the concepts "ultimately" and "convergence to a periodic process". The first of these concepts can be made rigorous as follows.

Definition 2.1 A periodic (stationary) solution $x^0(t, \omega)$ of (2.6) is stable in a certain sense for initial conditions belonging to a class **K** if for all random variables $x_0(t_0, \omega) \in \mathbf{K}$, a.s. the solution $x(t, x_0(t_0, \omega), t_0, \omega)$ of (2.6) with initial condition $x(t_0) = x_0(t_0, \omega)$ converges to $x^0(t, \omega)$ in that same given sense as $t_0 \to -\infty$.

In accordance with the various types of convergence (see Sect. 2.1), we can consider almost sure stability, stability in probability, weak stability, and so on. In this section we shall establish some sufficient conditions for almost sure stability.

The following theorem indicates the connection between the asymptotically stable compact invariant set of a deterministic equation and the periodic (stationary) solutions of the perturbed system obtained when a small stochastic process is superimposed on the deterministic system. To simplify the exposition, we shall confine ourselves to the case in which the invariant set is an equilibrium point, the system of equations is autonomous and the random perturbation stationary.

Theorem 2.7 *Let y_0 be an asymptotically stable singular point of the system*

$$\frac{dx}{dt} = F(x), \tag{2.31}$$

where $F(x) \in \mathbf{C}$. Let $g(x, z)$ ($x \in \mathbb{R}^l$, $z \in \mathbb{R}^k$) be a bounded Borel-measurable function such that $\|\partial g(x, z)/\partial x\|$ is bounded in a neighborhood of the point y_0, and $\xi(t, \omega)$ a stochastically continuous stationary stochastic process in $\mathbb{R}^k$. Then for all sufficiently small $|\varepsilon|$ the equation

$$\frac{dx}{dt} = F(x) + \varepsilon g(x, \xi(t, \omega)) \tag{2.32}$$

has a stationary solution which almost surely satisfies the condition

$$\sup_{-\infty < t < \infty} |x(t, \omega) - y_0| < \delta(\varepsilon) \quad (\delta(\varepsilon) \to 0 \text{ as } \varepsilon \to 0).$$

If moreover the point y_0 is asymptotically stable for the system (2.31) *in the linear approximation, then a sufficiently small neighborhood of the point y_0 contains a unique stationary solution of* (2.32) *which is almost surely stable for every initial condition $x_0(t_0, \omega)$ such that for some $\delta_1(\varepsilon)$*

$$\mathbf{P}\{|x_0(t_0, \omega) - y_0| < \delta_1(\varepsilon)\} = 1. \tag{2.33}$$

Proof Suppose y_0 is asymptotically stable for the system (2.31) and consider a fixed neighborhood of y_0. If $|\varepsilon|$ and $|x(t_0) - y_0|$ are sufficiently small, then no solution of the system (2.32) can leave this neighborhood for $t > t_0$. This follows directly from the stability of the solution $x(t) \equiv y_0$ of (2.31) with respect to continually acting perturbations (see [191]). This together with Theorem 2.4 implies the first part of the theorem.

Since the linear system

$$\frac{dz}{dt} = \frac{\partial F}{\partial x}(y_0)z$$

is asymptotically stable and the matrix $((\partial g/\partial x))$ is bounded in a neighborhood of y_0, there exists a constant $\delta_1(\varepsilon)$ such that for $|x_i - y_0| < \delta_1(\varepsilon)$, all $t > t_0$ and certain positive constants c and λ,

$$|x^{(2)}(t) - x^{(1)}(t)| < ce^{-\lambda(t-t_0)}, \tag{2.34}$$

where $x^{(i)}(t)$ is a solution of (2.32) with initial condition $x^{(i)}(t_0) = x_i$, $i = 1, 2$.

Let $X(t, \omega)$ be some stationary solution of (2.32) in the $\delta_1(\varepsilon)$-neighborhood of the point y_0, and $Y^{(t_0)}(t, \omega)$ a solution of (2.32) satisfying the initial condition $Y^{(t_0)}(t_0, \omega) = x_0(t_0, \omega)$, where $x_0(t_0, \omega)$ satisfies condition (2.33). Setting $x^{(1)} = X(t_0, \omega)$, $x^{(2)} = x_0(t_0, \omega)$ in (2.34), we see that

$$\mathbf{P}\left\{ \lim_{t_0 \to -\infty} Y^{(t_0)}(t, \omega) = X(t, \omega) \right\} = 1$$

as $t_0 \to -\infty$ which implies the required assertions. $\qquad\square$

Note that if we set $x^{(1)} = X(t_0, \omega)$ in (2.34) and let $t_0 \to -\infty$, the evolution of the process $X(t, \omega)$ for $t \in (-\infty, s)$ is determined by that of the process $\xi(t, \omega)$ on the same interval. If moreover $g(x, z)$ is invertible as a function of z, the converse is also true. Thus the process $X(t, \omega)$ has the same regularity and mixing properties (see [241]) as the process $\xi(t, \omega)$.

Theorem 2.8 *Let G be a function which is θ-periodic in t (independent of t) and satisfies the assumptions of Theorem* 1.11, *and $\xi(t, \omega)$ a θ-periodic (stationary) stochastic process. Then the equation*

$$\frac{dx}{dt} = G(x, t, \xi(t, \omega))$$

has a unique periodic (stationary) solution which is almost surely stable for any initial conditions such that $\mathbf{P}\{|x_0(t_0, \omega)| < c\} = 1$ for some c.

The reader should have no difficulty in proving this theorem, employing the arguments used in the proofs of Theorems 2.7 and 1.11.

We conclude this chapter with the following comments.

Remark 2.3 1. Theorem 2.4, which is the fundamental theorem of this chapter, admits various generalizations. For example, it is not hard to prove a corresponding result for equations with delayed argument (see [142]) and for Itô (stochastic) equations (see Chap. 3).

In [106] similar methods were used to prove an analogous theorem for Itô equations with delay.

2. The problem of the existence and stability of stationary (periodic) solutions is also of interest for partial differential equations. For example, let us consider a simple model problem in the strip $0 < x < 1, -\infty < t < \infty$:

$$\left.\begin{aligned}
&\frac{\partial u}{\partial t} = a(x)\frac{\partial^2 u}{\partial x^2} + b(x)\frac{\partial u}{\partial x} + c(x)u + f(x, \xi(t, \omega)) = Lu + f(x, \xi(t, \omega)), \\
&u(0, t) = u(1, t) = 0.
\end{aligned}\right\}$$
(2.35)

It is readily shown that if $\xi(t)$ is a continuous stochastic process and $\mathbf{E}f(x, \xi(t, \omega))$ is bounded uniformly in $x \in [0, 1]$, then problem (2.35) has a stationary solution in the following sense: There exists a function $u(x, t, \omega)$ satisfying the equation and the boundary conditions of (2.35) for almost all ω, which for each fixed x is a stationary stochastic process stationarily related to $\xi(t, \omega)$.

Let $p(x, t, y)$ denote the Green function of the problem

$$\frac{\partial u}{\partial t} = Lu, \qquad u(0, t) = u(1, t) = 0.$$

Then the above-mentioned stationary solution can be determined from the formula

$$u(x, t, \omega) = \int_{-\infty}^{t} ds \int_{0}^{1} p(x, t - s, y) f(y, \xi(s, \omega)) \, dy.$$

It is easy to show that this stationary solution is stable in the sense that every solution of problem (2.35) satisfying the initial condition $u(x, t_0) = \varphi(x, t_0)$ converges almost surely to $u(x, t, \omega)$ as $t_0 \to -\infty$, for every bounded function $\varphi(x, t_0)$.

This model can be readily generalized; for example, instead of homogeneous boundary conditions one can consider conditions of the form

$$u(0, t) = \xi_1(t, \omega), \qquad u_1(1, t) = \xi_2(t, \omega),$$

where $\xi_1(t, \omega), \xi_2(t, \omega)$ are stationary and stationarily related stochastic processes.

It is also easy to prove the existence of a stationary solution in case of an unbounded domain, provided the coefficient $c(x)$ in the operator L satisfies the condition $c(x) \leq c_0 < 0$. There is an analogous result for periodic solutions.

Apparently far more interesting but not so well investigated is the existence problem for stationary solutions of nonlinear partial differential equations. A few papers have been devoted to the solutions of the equations of hydrodynamics with stochastic coefficients (see the survey article [114] which includes a detailed bibliography).

3. We have established above certain results concerning the almost sure stability of stationary and periodic solutions. Although it seems that weak stability is rather more frequently encountered, no general conditions for weak stability are presently known. In particular, the following well-known problem seems to be yet unsolved. Let $F(x,t)$ ($x \in \mathbb{R}^1$) be a periodic function such that $F(x,t) \operatorname{sign} x \to -\infty$ as $|x| \to \infty$. Consider the equation $dx/dt = F(x,t) + \xi(t,\omega)$. What restrictions do we have to impose on the periodic process $\xi(t)$ in order to ensure that every solution of this equation defined by an initial condition independent of $\xi(t)$ converges to some periodic solution? It seems probable that this property is shared by quite a broad class of processes $\xi(t)$. For example, it is known that even in the relatively "unfavorable" case of a deterministic process $\xi(t)$ the property always holds (see [228, Theorem 9.2]).

4. The question of stability of stationary and periodic solutions is intimately connected with the investigation of the properties of a stationary (periodic) solution of (2.19). Suppose that (2.19) has a stationary solution $x(t)$. To simplify matters, assume that F and σ are independent of t and $\xi(t)$ is a stationary process which is ergodic, regular, satisfies a mixing condition, etc. Under what restrictions on F, σ will the process $x(t)$ possess the analogous properties?

In the proof of Theorem 2.7 above we have answered this question only in the simplest case.

Chapter 3
Markov Processes and Stochastic Differential Equations

3.1 Definition of Markov Processes[1]

Consider the following equation in $\mathbb{R}^l$:

$$\frac{dx}{dt} = F(x(s), t), \tag{3.1}$$

where for each t, $F(x(s), t)$ is a functional of the segment of the sample function $x(s)$ for $t - \tau(t) \leq s \leq t$. It is well-known that the specification of $x(t_0)$ does not determine the solution of this equation for $t > t_0$ if $\tau(t) > 0$. Moreover, the solution depends on the "past" of the sample function. Therefore (3.1) is known as an equation with after-effect. As opposed to this, an ODE might be called an equation without after-effect.

Markov processes occupy roughly the same position among stochastic processes as do ODEs among equations with after-effect.

To clarify the meaning of this statement, we first consider the case of discrete time. In what follows it will be assumed that the reader is familiar with the concept of conditional expectation; see, e.g., [185]. Let $x_0, \xi_1, \xi_2, \ldots, \xi_n, \ldots$ be random variables in $\mathbb{R}^l$. We define new random variables by the formulae

$$x_{n+1} - x_n = a_{n+1}(x_n) + \sigma_{n+1}(x_n)\xi_{n+1} \quad (n = 0, 1, 2, \ldots). \tag{3.2}$$

Here the vectors $a_n(x)$ and the matrices $\sigma_n(x)$ are $\mathfrak{B}$-measurable functions in $\mathbb{R}^l$. (Recall that $\mathfrak{B}$ denotes the σ-algebra of Borel sets in $\mathbb{R}^l$.)

In the general case, when the joint distribution of $x_0, \xi_1, \ldots, \xi_n, \ldots$ is arbitrary, the conditional distribution of x_{n+1}, given x_n, is not the same as the conditional distribution of x_{n+1}, given $x_0, x_1, \ldots, x_n$, and in this sense the sequence $x_0, x_1, \ldots$ may be termed a process with after-effect. But if $x_0, \xi_1, \ldots, \xi_n$ are independent, then it is easily shown that for each $A \in \mathfrak{B}$ and $k > 0$,

$$\mathbf{P}\{x_{n+k} \in A \mid x_0, x_1, \ldots, x_n\} = \mathbf{P}\{x_{n+k} \in A \mid x_n\} \quad \text{(a.s.)}. \tag{3.3}$$

---

[1]A more detailed definition of Markov processes may be found in [63, 64].

(Here and below "a.s." and "P-a.s." will serve as abbreviations for almost surely, i.e., the relation in question is valid with probability 1 relative to the measure **P**.)

Intuitively speaking, (3.3) means that the prediction of the behavior of the sequence $x_{n+1}, x_{n+2}, \ldots$ when x_n is known remains unchanged if the entire "history" of the process is known for $k < n$; in other words, the past has no effect on the future when the present is fixed (the past has no after-effects).

Let $P_n(x, A)$ denote the distribution of the random variable $x + a_n(x) + \sigma_n(x)\xi_n$. Then, for each $A \in \mathfrak{B}$,

$$\mathbf{P}\{x_n \in A \mid x_{n-1}\} = P_n(x_{n-1}, A) \quad \text{(a.s.)}. \tag{3.4}$$

It is clear that $P_n(x, A)$ is a Borel-measurable function of x.

A sequence $x_0, x_1, \ldots, x_n, \ldots$ satisfying condition (3.3) for all $A \in \mathfrak{B}$ is called a *Markov chain*.

It can be shown (see [54]) that for every Markov chain there exists a function $P_n(x, A)$, which is a $\mathfrak{B}$-measurable function of x, and which is for every fixed n, x a measure satisfying condition (3.4). This function is the one-step transition probability from x to A at time n. This function immediately generalizes to the transition probability from x at time k to A at time n. This satisfies the relation

$$P(k, x, n, A) = \mathbf{P}\{x_n \in A \mid x_k = x\} \quad \text{(a.s.)}. \tag{3.5}$$

It is clear that $P_n(x, A) = P(n, x, n + 1, A)$.

It is possible to construct a Markov chain for any a priori given family of transition probabilities $P_n(x, A)$. This Markov chain satisfies the Chapman–Kolmogorov equation:

$$P(k, x, m, A) = \int_{\mathbb{R}^\ell} P(k, x, n, dy) P(n, y, m, A) \quad (k < n < m). \tag{3.6}$$

The sequence (3.2) considered in the example above is a very special case of a Markov chain. Nevertheless, every Markov chain can be represented as a system whose evolution at time $n + 1$ is entirely determined by n, its state at time n, and certain random factors which are independent of the entire past history of the system.

In the continuous-time case we introduce the following definition which is analogous to (3.3).

A stochastic process $X(t, \omega)$ with values in $\mathbb{R}^l$, defined for $t \geq 0$ on a probability space $(\Omega, \mathfrak{A}, \mathbf{P})$, is called a *Markov process* if, for all $A \in \mathfrak{B}$, $0 \leq s < t$,

$$\mathbf{P}\{X(t, \omega) \in A \mid \mathcal{N}_s\} = \mathbf{P}\{X(t, \omega) \in A \mid X(s, \omega)\} \quad \text{(a.s.)}, \tag{3.7}$$

where $\mathcal{N}_s$ is the σ-algebra of events generated by all events of the form

$$\{X(u, \omega) \in A\} \quad (u \leq s, \ A \in \mathfrak{B}).$$

It can be proved that there exists a function $P(s, x, t, A)$, defined for $0 \leq s \leq t$, $x \in \mathbb{R}^\ell$, $A \in \mathfrak{B}$, which is $\mathfrak{B}$-measurable in x for every fixed s, t, A, and which

constitutes a measure as a function of the set A, satisfying the condition

$$\mathbf{P}\{X(t, \omega) \in A \mid X(s, \omega)\} = P\{s, X(s, \omega), t, A\} \quad \text{(a.s.).} \tag{3.8}$$

One can also prove that for all x, except possibly those from a set B such that $P\{x(s, \omega) \in B\} = 0$, the Chapman–Kolmogorov equation holds:

$$P\{s, x, t, A\} = \int_{\mathbb{R}^{\ell}} P(s, x, u, dy) P(u, y, t, A). \tag{3.9}$$

The function $P\{s, x, t, A\}$ is called the *transition probability function* of the Markov process. It is usually assumed at this point (and we shall indeed assume this in the sequel) that (3.9) is valid for all $x \in \mathbb{R}^l$, and $P\{s, x, s, \mathbb{R}^l \setminus x\} = 0$.

Conversely, given a transition probability function $P(s, x, t, A)$, one can construct a Markov process with an arbitrary initial distribution. In particular, for $t > s$ one can associate with the function $P(s, x, t, A)$ a family $X^{(s,x)}(t, \omega)$ of Markov processes such that the process $X^{(s,x)}(t, \omega)$ "exits" from the point x at time s, i.e.,

$$\mathbf{P}\{X^{(s,x)}(s, \omega) = x\} = 1. \tag{3.10}$$

Later on we shall often deal with families of Markov processes $X^{(s,x)}(t, \omega)$ of this kind and with the measures generated by them (Markov families, in the terminology of Dynkin [64]).

The transition probability function $P(s, x, t, A)$ is said to be *time-homogeneous* (and the corresponding Markov process is called *time-homogeneous*) if the function $P(s, x, t + s, A)$ is independent of s. It is called *periodic* if $P(s, x, t + s, A)$ is periodic in s.

A time-homogeneous transition probability function is effectively a function of a single time variable, and we shall therefore use the notation $P(s, x, t, A) = P(x, t - s, A)$.

With each time-homogeneous transition probability function $P(x, t, A)$ we can associate two families of operators; the first is defined on functions and the second on measures:

$$T_t V(x) = \int P\{x, t, dy\} V(y) = \mathbf{E} V(X^{(x)}(t, \omega)),$$

$$S_t \mu(A) = \int \mu(dx) P(x, t, A).$$

As usual, we denote by $\mathbf{C}(\mathbb{R}^l)$ ($\mathbf{C}(E)$) the space of continuous function on $\mathbb{R}^l$ (on E). A transition probability function for which the operator T_t maps the space $\mathbf{C}(\mathbb{R}^l)$ into itself is known as a *Feller transition probability function*. If $P(x, t, A)$ is also stochastically continuous, i.e., $P(x, t, U_\varepsilon(x)) \to 1$ as $t \to +0$ for each $\varepsilon > 0$ (where $U_\varepsilon(x)$ is the ε-neighborhood of x), then it is readily seen that $T_t f(x) \to f(x)$ as $t \to +0$ for $f(x) \in \mathbf{C}(\mathbb{R}^l)$. From (3.9) we get then the relation $T_{t+s} = T_t T_s$

$(s > 0, t > 0)$, and hence the family T_t is a homogeneous semigroup on $\mathbf{C}(\mathbb{R}^l)$. Its $\mathbf{C}$-infinitesimal operator (generator) A is defined by the standard formula (see [64])

$$AV(x) = \lim_{t \to +0} \frac{T_t V(x) - V(x)}{t}. \tag{3.11}$$

The domain of definition D_A of the operator A is the set of functions for which the limit in (3.11) exists uniformly for $x \in \mathbb{R}^l$.

A stochastically continuous transition probability function is uniquely determined by its generator (see [64, Chap. II]). Let us show that the case of a non-homogeneous transition probability function can be reduced to the homogeneous case by extension of the phase space.

To do this, we consider the σ-algebra $\mathfrak{B}_0$ of Borel sets on the real axis and we define the function $Q((s, x), t, \Gamma)$ for $t \geq 0$, $s \geq 0$, $x \in \mathbb{R}^l$, $\Gamma = A \times \Delta$ ($A \in \mathfrak{B}$, $\Delta \in \mathfrak{B}_0$) by

$$Q((s, x), t, \Gamma) = P(s, x, s + t, A)\mathbb{1}_\Delta(s + t).$$

(Here and below $\mathbb{1}_\Delta(t)$ denotes the indicator function of the set Δ, i.e., the function equal to 1 for $t \in \Delta$ and 0 for $t \notin \Delta$.) The measure Q can be extended by standard methods (see [97]) to the σ-algebra $\mathfrak{B} \times \mathfrak{B}_0$ generated by the sets Γ of the above form. It is readily seen that the resulting function Q is a homogeneous transition probability function in the phase space $E = \mathbb{R}^l \times I$. If $P(s, x, t, A)$ is stochastically continuous, then Q is also stochastically continuous. Q is the transition probability function of the process

$$Y(t, \omega) = \{X^{(s,x)}(s + t, \omega), s + t\}$$

in $\mathbb{R}^{l+1}$. The generator $\widetilde{A}$ of the semigroup $\widetilde{T}_t$ defined by Q is obviously given by

$$\widetilde{A}V(s, x) = \lim_{h \to +0} \frac{\int P(s, x, s + h, dy)V(s + h, y) - V(s, x)}{h}$$

$$= \lim_{h \to +0} \frac{\mathbf{E}V(s + h, X^{(s,x)}(s + h, \omega)) - V(s, x)}{h}. \tag{3.12}$$

$P(s, x, t, A)$ is known as a *Feller transition probability function* if the operator

$$\widetilde{T}_t V(s, x) = \int P(s, x, s + t, dy)V(s + t, y)$$

transforms functions $V(s, x) \in \mathbf{C}(E)$ into continuous functions.

It follows from the above-mentioned result concerning the time-homogeneous case that the operator $\widetilde{A}$ uniquely determines the function $P(s, x, t, A)$.

An important example of a stochastically continuous Markov process with a Feller transition probability function is the process $\xi(t, \omega)$ of Brownian motion (Wiener process), whose transition probability has a density with respect to the

Lebesgue measure on the real line and moreover

$$P(x, t, A) = \frac{1}{\sqrt{2\pi t}} \int_A \exp\left\{-\frac{(x-y)^2}{2t}\right\} dy. \tag{3.13}$$

It follows from (3.13) that the probability

$$\mathbf{P}\{\xi(t+h) - \xi(t) \in A \mid \xi(t) = x\} = \frac{1}{\sqrt{2\pi h}} \int_A \exp\left(-\frac{z^2}{2h}\right) dz$$

is independent of x. This implies that the increments of the process $\xi(t, \omega)$ on non-overlapping time intervals are independent. In addition, it also follows from (3.13) that

$$\mathbf{E}(\xi(t+h) - \xi(t)) = 0; \qquad \mathrm{var}(\xi(t+h) - \xi(t)) = h. \tag{3.14}$$

It is also not hard to see that

$$\mathbf{E}(\xi(t+h) - \xi(t))^{2n} = \frac{1}{\sqrt{2\pi h}} \int_{-\infty}^{+\infty} e^{-z^2/2h} z^{2n} \, dz = (2n-1)!! \, h^n. \tag{3.15}$$

It follows from (3.15) and Theorem 1.1, that the process $\xi(t)$ has continuous sample functions a.s.

For definiteness, we assume that $\xi(0) = 0$. Given the process $\xi(t)$, one can construct other continuous processes by means of the transformation

$$x(t) = s(0) + \int_0^t b(s) \, ds + \xi\left(\int_0^t \sigma^2(s) \, ds\right).$$

The reader should have no difficulty in verifying that each of the resulting processes also has independent increments and a Gaussian transition probability function. Moreover,

$$\mathbf{E}(x(t+h) - x(t)) = \int_t^{t+h} b(s) \, ds;$$

$$\mathrm{var}(x(t+h) - x(t)) = \int_t^{t+h} \sigma^2(s) \, ds.$$

As we shall see from the sequel, the Wiener process can be used to construct a far more extensive class of Markov processes with continuous sample functions.

It is known (see [252]) that every almost surely continuous process with independent increments in $\mathbb{R}^l$ is Gaussian.

3.2 Stationary and Periodic Markov Processes

We shall now investigate conditions under which a Markov process $X(t, \omega)$ is stationary. One necessary condition, at any rate (see the definition of Sect. 2.1), is

that for every $A, B \in \mathfrak{B}$ the probabilities of the events $\{X(t) \in A\}$ and $\{X(t) \in A, \; X(t+h) \in B\}$ are independent of t. Hence, expressing these probabilities in terms of the transition probability function, we see that the transition probability function of a stationary process is time-homogeneous, and for every $h > 0$ the initial distribution $P_0(A) = P\{X(0) \in A\}$ satisfies the equation

$$P_0(A) = \int P_0(dx) P(x, h, A). \tag{3.16}$$

These two conditions are also sufficient for a Markov process to be stationary. Indeed, for $0 < h_1 < \cdots < h_n$ the probability of the event

$$\{x(t) \in A_0; \, x(t+h_1) \in A_1; \ldots; \, x(t+h_n) \in A_n\}$$

is[2]

$$\int P_0(dx) \int_{A_0} P(0, x, t, dx_1) \ldots \int_{A_n} P(t + h_{n-1}, x_n, t + h_n, dx_{n+1}).$$

It follows from the homogeneity of the transition probability function and from (3.16) that this probability is independent of t.

   In exactly the same way one can show that a Markov process $x(t)$ is θ-periodic if and only if its transition probability function is θ-periodic and the function $P_0(t, A) = P\{X(t) \in A\}$ satisfies the equation

$$P_0(s, A) = \int_{\mathbb{R}^l} P_0(s, dx) P(s, x, s + \theta, A) \equiv P_0(s + \theta, A) \tag{3.17}$$

for every $A \in \mathfrak{B}$.

   However, it is not true that for every time-homogeneous (periodic) transition probability function there exists a corresponding stationary (periodic) Markov process. For example, there is no stationary process for the Wiener transition probability function. To prove this, suppose that such a process exists. Then, it follows by (3.13) and (3.16) that for every set $A \in \mathfrak{B}$ on the real line whose Lebesgue measure $m(A)$ is finite

$$P_0(A) = \int P_0(dx) \int_A \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(x-y)^2}{2t}\right\} dy \leq \frac{1}{\sqrt{2\pi t}} m(A).$$

Letting $t \to \infty$, we get $P_0(A) = 0$. Hence it follows that $P_0(\mathbb{R}^l) = 0$, contradicting $P_0(\mathbb{R}^l) = 1$.

   We shall now determine some further conditions that must hold for a homogeneous (periodic) transition probability function corresponding to a stationary (periodic) Markov process.

---

[2]Henceforth we shall omit the limits of integration when the integration is performed over the entire space $\mathbb{R}^l$.

Theorem 3.1 *A necessary and sufficient condition for the existence of a stationary Markov process with the given time-homogeneous stochastically continuous Feller transition probability function $P(x, t, A)$ is that for some point $x \in \mathbb{R}^l$[3]*

$$\lim_{R \to \infty} \lim_{T \to \infty} \frac{1}{T} \int_0^T P(x, t, U_R^c) \, dt = 0. \tag{3.18}$$

For the proof we need the following lemma.

Lemma 3.1 *A stochastically continuous Feller transition probability function $P(x, t, A)$ is a $\mathfrak{B} \times \mathfrak{B}_0$-measurable function of (x, t) for every $A \in \mathfrak{B}$.*

Proof It will suffice to prove the lemma for closed sets A. If A is closed, we can construct a sequence $f_n(x)$ of monotone decreasing continuous functions converging to the indicator function $\mathbb{1}_A(x)$. It is clear that for each $t > 0$

$$T_t f_n(x) \to P(x, t, A) \quad \text{as } n \to \infty.$$

The function $T_t f_n(x)$ is continuous in x and right-continuous in t, as follows from the Feller property and the stochastic continuity of the transition probability function. Therefore it is Borel-measurable as a function of x, t. Hence the function $P(x, t, A)$ is also measurable. $\square$

Proof of Theorem 3.1 (1) Necessity. Let $P_0(A)$ be the stationary initial distribution. Then, integrating (3.16) with respect to t from 0 to T and applying Fubini's theorem (see [97]), we get

$$P_0(U_R^c) = \int P_0(dx) \frac{1}{T} \int_0^T P(x, t, U_R^c) \, dt.$$

Now suppose that condition (3.18) does not hold. Then

$$\lim_{R \to \infty} \lim_{T \to \infty} \frac{1}{T} \int_0^T P(x, t, U_R^c) \, dt = q(x) > 0.$$

Therefore

$$0 = \lim_{R \to \infty} P_0(U_R^c) \geq \int P_0(dx) q(x) > 0.$$

This is a contradiction and therefore (3.18) must be satisfied.

(2) Sufficiency. It follows from condition (3.18) that for some x_0 there exists a sequence $T_n \to \infty$ such that

$$\frac{1}{T_n} \int_0^{T_n} P(x_0, t, U_R^c) \, dt \to 0 \quad \text{uniformly in } n \text{ as } R \to \infty. \tag{3.19}$$

---

[3]Recall that $U_R^c = \{|x| > R\}$.

Consider the sequence of measures on $\mathbb{R}^l$ defined by

$$P_n(A) = \frac{1}{T_n} \int_0^{T_n} P(x_0, t, A)\, dt.$$

By condition (3.19), this sequence is weakly compact (see Sect. 2.1). Let P_{n_k} be a subsequence converging weakly to some measure P_0. We claim now that the measure P_0 satisfies condition (3.16) and consequently defines the initial distribution of a stationary Markov process.

Let $f(x) \in \mathbf{C}(\mathbb{R}^l)$. Since P_{n_k} is weakly convergent to P_0, and the Feller property holds for the transition probability, we have

$$\int P_0(dx) \int P(x, t, dy) f(y)$$

$$= \lim_{n_k \to \infty} \frac{1}{T_{n_k}} \int_0^{T_{n_k}} ds \int P(x_0, s, dx) \int P(x, t, dy) f(y)$$

$$= \lim_{n_k \to \infty} \frac{1}{T_{n_k}} \int_0^{T_{n_k}} ds\, P(x_0, s + t, dy) f(y)$$

$$= \lim_{n_k \to \infty} \frac{1}{T_{n_k}} \left[\int_0^{T_{n_k}} du \int P(x_0, u, dy) f(y) \right.$$

$$\left. + \int_{T_{n_k}}^{T_{n_k}+t} du \int P(x_0, u, dy) f(y) - \int_0^t du \int P(x_0, u, dy) f(y) \right]$$

$$= \int P_0(dy) f(y).$$

The resulting relation

$$\int P_0(dx) T_t f(x) = \int P_0(dx) f(x)$$

is equivalent to (3.16), i.e. (3.16) holds. This completes the proof.[4] $\square$

Theorem 3.2 *A necessary and sufficient condition for the existence of a θ-periodic Markov process with a given θ-periodic transition probability function $P(s, x, t, A)$ is that for some x_0, s_0*

$$\lim_{R \to \infty} \varlimsup_{n \to \infty} \frac{1}{n} \sum_{k=1}^n P(s_0, x_0, s_0 + k\xi, U_R^c) = 0. \tag{3.20}$$

---

[4]The method of showing the existence of a stationary Markov process employed in the proof of Theorem 3.1 is well-known. It was first used by Krylov and Bogolyubov [163] to prove the existence of an invariant measure for a dynamic system. The method is systematically used for Markov processes, e.g., in [56].

The proof of this theorem is entirely analogous to that of Theorem 3.1. The only difference lies in the definition of $P_n(A)$ which is now given by

$$P_n(A) = \frac{1}{k_n} \sum_{i=1}^{k_n} P(s_0, x_0, s_0 + i\theta, A),$$

where k_n is an increasing sequence of integers such that the sequence of measures $P_n(A)$ is weakly compact.

Remark 3.1 Condition (3.20) of Theorem 3.2 can be replaced by the more easily tested condition

$$\lim_{R \to \infty} \lim_{T \to \infty} \frac{1}{T} \int_0^T P(s, x, s + u, U_R^c) \, du = 0, \tag{3.21}$$

provided the transition probability function $P(s, x, t, A)$ satisfies the following not very restrictive assumption that

$$\alpha(R) = \sup_{x \in U_{\beta(R)}, \, 0 < s, \, t < \theta} P(s, x, s + t, U_R^c) \to 0 \quad \text{as } R \to \infty \tag{3.22}$$

for some function $\beta(R)$ which tends to infinity as $R \to \infty$.

Indeed, for every

$$u \in ((k-1)\theta, k\theta)$$

it is obvious that

$$P(s, x, s + k\theta, U_R^c)$$
$$= \left(\int_{U_{\beta(R)}} + \int_{U_{\beta(R)}^c} \right) P(s, x, s + u, dy) P(s + u, y, s + k\theta, U_R^c)$$
$$\leq P(s, x, s + u, U_{\beta(R)}^c) + \sup_{u \in ((k-1)\theta, k\theta), \, y \in U_{\beta(R)}} P(s + u, y, s + k\theta, U_R^c)$$

and hence, integrating both sides with respect to u from $(k-1)\theta$ to $k\theta$, we see that

$$P(s, x, s + k\theta, U_R^c) \leq \frac{1}{\theta} \int_{(k-1)\theta}^{k\theta} P(s, x, s + u, U_{\beta(R)}^c) \, du + \alpha(R).$$

Summation over k from 1 to n in combination with (3.22) shows that (3.21) implies (3.20). A similar argument shows that if (3.22) is satisfied, then (3.20) implies (3.21).

3.3 Stochastic Differential Equations (SDE)

In Sect. 3.1 we have used the finite-difference equation (3.2) to determine a Markov chain X_n from a sequence of independent random variables ξ_n. It is natural to try to

use the differential analog of (3.16) to construct continuous-time Markov processes. The formal analog of (3.2) for the one-dimensional case is

$$\frac{dX}{dt} = b(t, X) + \sigma(t, X)\dot{\xi}(t). \tag{3.23}$$

The random process $\dot{\xi}(t)$, by analogy with the sequence ξ_n in (3.2), must be a process with independent values for different t. Such a process does not exist, but (3.23) can nonetheless be given a rigorous meaning.

To this end, it is convenient to rewrite the equation in terms of differentials:

$$dX(t) = b(t, X)\, dt + \sigma(t, X)\, d\xi(t). \tag{3.24}$$

What properties must the process $\xi(t)$ possess? First, it must have independent increments, as the integral of a "process" $\dot{\xi}(t)$ with independent values. Second, it must have continuous sample functions, if we wish the solution $X(t)$ of (3.24) to be a continuous stochastic process. As mentioned at the end of Sect. 3.1, such a process is always Gaussian. We may also assume that its mean and variance satisfy conditions (3.14), since this may always be achieved by modifying the coefficients b and σ.

Equations of type (3.23) or (3.24) were first considered by Langevin [176] as far back as 1908, shortly after Einstein and Smoluchowski had published their first papers on the theory of Brownian motion. More systematic investigations of stochastic equations began only in the thirties. The simplest and most convenient construction of the solution of the SDE (3.24) was given by Itô [103, 104]. This construction has been presented in detail in many books (e.g., [56], [251], [64], [92]). We shall therefore present without proof some of Itô's theorems on the existence and properties of solutions of (3.24).

The basic tool for the constructions is the stochastic integral. Let $\xi(t, \omega)$ be a Wiener process on the interval $[a, b]$, defined on a probability space $(\Omega, \mathfrak{A}, \mathbf{P})$. Let $\widetilde{\mathcal{N}}_t$ $(t \geq 0)$ be a family of σ-algebras of sets in $\mathfrak{A}$, related to the Wiener process $\xi(t)$ as follows:

1. $\widetilde{\mathcal{N}}_{t_1} \subset \widetilde{\mathcal{N}}_{t_2}$, if $t_1 < t_2$;
2. $\xi(t)$ is an $\widetilde{\mathcal{N}}_t$-measurable random variable for each $t \geq 0$;
3. the increment $\xi(t + h) - \xi(t)$ of the process $\xi(t)$ is independent of every event $A \in \widetilde{\mathcal{N}}_t$.

For every bounded step function $f(t, \omega) = f(t)$ with jumps at points $t_1, \ldots, t_n$, such that $f(t)$ is $\widetilde{\mathcal{N}}_t$-measurable for each $t \in [a, b]$, the Itô stochastic integral is defined as the sum

$$\int_a^b f(t)\, d\xi(t) = \sum_{i=0}^{n-1} f(t_i)[\xi(t_{i+1}) - \xi(t_i)].$$

Using the independence of $\xi(t_{i+1}) - \xi(t_i)$ and $f(t_i)$, one easily verifies the following properties of integrals of step functions:

$$\mathbf{E}\left(\int_a^b f(t)\,d\xi(t)\Big|\widetilde{\mathcal{N}}_a\right) = 0 \quad \text{(a.s.)}, \tag{3.25}$$

$$\mathbf{E}\left(\left[\int_a^b f(t)d\xi(t)\right]^2\Big|\widetilde{\mathcal{N}}_a\right) = \int_a^b \mathbf{E}(f^2(t)|\widetilde{\mathcal{N}}_a)\,dt \quad \text{(a.s.)}. \tag{3.26}$$

Next, the integral of an $\widetilde{\mathcal{N}}_t$-measurable function $f(t)$ such that $f^2(t)$ is in class $\mathbf{L}$ is defined by means of a passage to the limit (see [92, Chap. 8, Sect. 2]). One then proves that relations (3.25) and (3.26) are valid for every $\widetilde{\mathcal{N}}_t$-measurable function $f(t)$, provided

$$\int_a^b \mathbf{E}(f^2(t)|\widetilde{\mathcal{N}}_a)\,dt < \infty \quad \text{(a.s.)}. \tag{3.27}$$

It can be proved that the stochastic integral

$$\zeta(t) = \int_a^t f(s)\,d\xi(s) = \int_a^b \mathbb{1}_t(s)f(s)\,d\xi(s)$$

(where $\mathbb{1}_t(s)$ is the indicator function of the set $\{s < t\}$) can be so defined that it becomes a separable almost surely continuous stochastic process. This process satisfies the Kolmogorov's inequality

$$\mathbf{P}\left\{\sup_{t\in[a,b]} |\zeta(t)| > c \,\Big|\, \widetilde{\mathcal{N}}_a\right\} \leq \frac{1}{c^2}\int_a^b \mathbf{E}\{f^2(s)|\widetilde{\mathcal{N}}_a\}\,ds. \tag{3.28}$$

Hitherto we have considered stochastic integrals in $\mathbb{R}^1$. It is not difficult to extend the construction to the multi-dimensional case. Let $\sigma_1(t), \ldots, \sigma_k(t)$ be vectors in $\mathbb{R}^l$ whose components $\sigma_{ij}(t)$, $i = 1, \ldots, l$, $j = 1, \ldots, k$, are $\widetilde{\mathcal{N}}_t$-measurable for each fixed t. Let $\xi_1(t), \ldots, \xi_k(t)$ be mutually independent $\widetilde{\mathcal{N}}_t$-measurable Wiener processes such that the random variables $\xi_i(t + h) - \xi_i(t)$ are all independent of every event in $\widetilde{\mathcal{N}}_t$ for $h > 0$. Then stochastic integrals with values in $\mathbb{R}^l$ are defined in the following natural (coordinatewise) manner:

$$\zeta_r(t) = \int_a^t \sigma_r(s)\,d\xi_r(s).$$

Let $b(t) \in \mathbb{R}^l$ be $\widetilde{\mathcal{N}}_t$-measurable vector stochastic process.

The Itô (stochastic) differential $dX(t)$ of the $\widetilde{\mathcal{N}}_t$-measurable process $X(t)$ is defined as

$$dX(t) = b(t)\,dt + \sum_{r=1}^k \sigma_r(t)d\xi_r(t),$$

provided $b(t)$ and $|\sigma_r|^2$ are in class **L**, are $\widetilde{\mathcal{N}}_t$-measurable for each t, and for all $a < t_1 < t_2 < b$

$$X(t_2) - X(t_1) = \int_{t_1}^{t_2} b(t)\,dt + \sum_{r=1}^{k} \int_{t_1}^{t_2} \sigma_r(t)\,d\xi_r(t). \tag{3.29}$$

Denote $\sigma^*(t)$ matrix conjugated to $\sigma(t)$, and $A(t) = \sigma(t)\sigma^*(t)$. Itô [104] established the following analog of the chain rule for stochastic differentials (Itô's formula).

Theorem 3.3 *If the function $u(t, x)$ ($t \in [a, b]$, $x \in \mathbb{R}^l$) has continuous partial derivatives up to second order in x and to first order in t, and the process $X(t)$ with values in $\mathbb{R}^l$ has an Itô differential*

$$dX(t) = b(t)\,dt + \sum_{r=1}^{k} \sigma_r(t)\,d\xi_r(t),$$

then the process $\eta(t) = u(t, X(t))$ also has an Itô differential, and

$$d\eta(t) = \left[\frac{\partial u(t, X(t))}{\partial t} + \sum_{i=1}^{l} b_i(t) \frac{\partial u(t, X(t))}{\partial x_i} \right.$$

$$\left. + \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} a_{ij}(t) \frac{\partial^2 u(t, X(t))}{\partial x_i \partial x_j} \right] dt$$

$$+ \sum_{i=1}^{l} \sum_{r=1}^{k} \sigma_{r_i}(t) \frac{\partial u(t, X(t))}{\partial x_i}\,d\xi_r(t). \tag{3.30}$$

It is clear that the only difference between (3.30) and the usual chain rule is the presence of the term

$$\frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} a_{ij}(t) \frac{\partial^2 u(t, X(t))}{\partial x_i \partial x_j}\,dt.$$

We now return to the SDE (3.24), which will be interpreted as an equation relating the stochastic differentials of a process $X(t, \omega)$ in $\mathbb{R}^l$ and the Wiener processes $\xi_r(t)$; written by means of integrals this becomes

$$X(t) = X(t_0) + \int_{t_0}^{t} b(s, X(s))\,ds + \sum_{r=1}^{k} \int_{t_0}^{t} \sigma_r(s, X(s))\,d\xi_r(s). \tag{3.31}$$

In the sequel we shall mean by a solution of (3.31) on the interval $[t_0, T]$ a stochastic process $X(t)$ such that the random variable $X(t)$ is $\widetilde{\mathcal{N}}_t$-measurable for each t, the integrals in (3.31) exist and equality (3.31) holds almost surely for each $t \in [t_0, T]$.

The following theorem concerns the existence, uniqueness and certain other properties of the solution of (3.31).

Theorem 3.4 *Let the vectors* $b(s, x), \sigma_1(s, x), \ldots, \sigma_k(s, x)$ $(s \in [t_0, T], x \in \mathbb{R}^l)$ *be continuous functions of* (s, x)*, such that for some constant B the following conditions hold in the entire domain of definition*:

$$|b(s, x) - b(s, y)| + \sum_{r=1}^{k} |\sigma_r(s, x) - \sigma_r(s, y)| \leq B|x - y|,$$

$$|b(s, x)| + \sum_{r=1}^{k} |\sigma_r(s, x)| \leq B(1 + |x|). \tag{3.32}$$

Then:

1. *For every random variable* $X(t_0)$ *independent of the processes* $\xi_r(t) - \xi_r(t_0)$ *there exists a solution* $X(t)$ *of* (3.31) *which is an almost surely continuous stochastic process and is unique up to equivalence.*[5]
2. *This solution is a Markov process whose Feller transition probability function* $P(s, x, t, A)$ *is defined for* $t > s$ *by the relation* $P(s, x, t, A) = \mathbf{P}\{X^{s,x}(t) \in A\}$, *where* $X^{s,x}(t)$ *is a solution of the equation*

$$X^{s,x}(t) = x + \int_s^t b(u, X^{s,x}(u)) \, du + \sum_{r=1}^{k} \int_s^t \sigma_r(u, X^{s,x}(u)) \, d\xi_r(u). \tag{3.33}$$

3. *The transition probability function* $P(s, x, t, A)$ *satisfies for* $h \to 0$:

$$\left.\begin{array}{l} \mathbf{E}[X^{s,x}(s+h) - x] = \int (y - x) P(s, x, s+h, dy) = b(s, x)h + O(h^{3/2}), \\ \mathbf{E}[(X_i^{s,x}(s+h) - x_i)(X_j^{s,x}(s+h) - x_j)] = a_{ij}(s, x)h + O(h^{3/2}), \\ \mathbf{P}(s, x, s+h, U_\varepsilon^c(x)) = O(h^{3/2}), \end{array}\right\} \tag{3.34}$$

where all estimates $O(\cdot)$ *are uniform in* s, x *in each bounded domain, and* $a_{ij}(s, x)$ *are elements of a matrix* $A(s, x) = \sigma(s, x)\sigma(s, x)^*$.
4. *There exists a constant* k, *depending only on the dimension of the space* $\mathbb{R}^l$, *on the constant* B *of condition* (3.32) *and on the length* $T - t_0$ *of the interval, such that for all* $s, t \in [t_0, T]$

$$\mathbf{E}|X^{s,x}(t) - x|^4 \leq k(t - s)^2(1 + |x|^4).$$

5. *If the coefficients of* (3.31) *are independent of* s, *then the transition probability function of the corresponding Markov process is time-homogeneous; and if*

---

[5]Two solutions $X_1(t)$ and $X_2(t)$ are said to be *equivalent* if $\mathbf{P}\{X_1(t) = X_2(t)$ for all $t \in [t_0, T]\} = 1$.

the coefficients are θ-periodic in s, then the transition probability function is θ-periodic.

The proof of this theorem, except for the second part of (5), may be found, e.g., in [64], [92]. The proof of the second part of (5) is analogous to that of the first.

Here the matrix $A(s, x)$ is known as the diffusion matrix, and the vector $b(s, x)$ as the drift vector. Their probabilistic meaning is clear from formulas (3.34).

Let us consider the Markov process $X^{s,x}(t, \omega)$ determined by (3.33). We shall often have to calculate the expectation of various random variables which are measurable with respect to the evolution of this process (i.e., with respect to the σ-algebra generated by the events $\{X^{s,x}(t) \in A\}$, $A \in \mathfrak{B}$, $s < t < \infty$). Instead of writing the indices s, x in the symbols for each of these random variables, we shall sometimes attach them to the symbols **E** and **P**. For example, $\mathbf{P}^{s,x}\{X(t, \omega) \in A\} = \mathbf{P}\{X^{s,x}(t, \omega) \in A\}$. If the coefficients of (3.31) are independent of s, we need only consider the process $X^{0,x}(t)$ which will be denoted by $X^x(t)$. Accordingly, the index x for random variables involved in the process $X^x(t)$ will sometimes be attached to the symbols **E** and **P**, so we shall write $\mathbf{P}^x$ and $\mathbf{E}^x$ then.

Let $\mathbf{C}_2$ denote the class of functions on E which are twice continuously differentiable with respect to $x_1, \ldots, x_l$ and continuously differentiable with respect to t. Let $V \in \mathbf{C}_2$. Then it follows from Theorems 3.3 and 3.4 that

$$V(t, X(t)) - V(s, X(s)) = \int_s^t LV(u, X(u))\, du + \sum_{r=1}^{k} \sum_{i=1}^{l} \int_s^t \sigma_{ri} \frac{\partial V}{\partial x_i}\, d\xi_r(u),$$

$$\tag{3.35}$$

where

$$LV(s, x) = \frac{\partial V(s, x)}{\partial s} + \sum_{i=1}^{l} b_i(s, x)\frac{\partial V(s, x)}{\partial x} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(s, x)\frac{\partial^2 V(s, x)}{\partial x_i \partial x_j}.$$

$$\tag{3.36}$$

If moreover the function V and its derivatives are bounded (or increasing no faster than a linear function of x), then, calculating the expectation in (3.35) and using the properties of stochastic integrals and Fubini's theorem, we get

$$\mathbf{E}[V(t, X(t)) - V(s, X(s))] = \int_s^t \mathbf{E}LV(u, X(u))\, du. \tag{3.37}$$

Substituting $X(t) = X^{s,x}(t)$ into this equality, dividing both sides by $t - s$ and letting $t \to s + 0$, we readily find that

$$\lim_{h \to +0} \frac{1}{h}[\mathbf{E}^{s,x} V(s + h, X(s + h)) - V(s, x)] = LV(s, x). \tag{3.38}$$

It follows (see (3.11), (3.12)) that $\widetilde{A}V(s, x) = LV(s, x)$ for all $V \in \mathbf{C}_2$ with compact support, and that for homogeneous processes $AV(x) = LV(x)$.

The operator L defined by (3.36) will be called the generator of the Markov process. It is clear from the definition that this is a local concept, i.e., the value of the operator at a point (s, x) is determined by the values of V in an arbitrarily small neighborhood of the point (s, x).

The left-hand side of formula (3.38) may be undefined for rapidly growing processes (for example, for the Wiener process, when $t > 0$, the function $V = \exp(x^3)$ is not in the domain of the operator T_t). The probabilistic meaning of the operator L for any function $V \in \mathbf{C}_2$ is given by the following lemma:

Lemma 3.2 [6] *Let $X(u)$ be a process satisfying (3.31) on the time interval $[s, T]$, $V \in \mathbf{C}_2$, τ_U the random variable equal to the time at which the sample function of the process $X(u)$ first leaves the bounded neighborhood U, and let $\tau_U(t) = \min(\tau_U, t)$. Suppose moreover that $\mathbf{P}\{X(s) \in U\} = 1$. Then*

$$\mathbf{E}[V(\tau_U(t), X(\tau_U(t))) - V(s, X(s))] = \mathbf{E} \int_s^{\tau_U(t)} LV(u, X(u))\, du. \qquad (3.39)$$

Proof It is known (see [64, Chap. II, Sect. 3]) that the process $Y(t) = X(\tau_U(t))$, obtained by stopping the process $X(t)$ at the instant it reaches the boundary of the domain U, has an Itô differential:

$$dY(t) = \mathbb{1}_{\{\tau_U > t\}}(\omega) b(t, Y(t))\, dt + \sum_{r=1}^{k} \mathbb{1}_{\{\tau_U > t\}}(\omega) \sigma_r(t, Y(t))\, d\xi_r(t).$$

(Since $\{\tau_U > t\} \in \widetilde{\mathcal{N}}_t$, the Itô differential in the formula is defined.) Applying Theorem 3.3 to the process $Y(t)$ and the function V, we get

$$V(\tau_U(t), X(\tau_U(t))) - V(s, X(s))$$

$$= \int_s^{\tau_U(t)} LV\, du + \sum_{r=1}^{k} \sum_{i=1}^{l} \int_s^{\tau_U(t)} \sigma_{ri} \frac{\partial V}{\partial x_i}\, d\xi_r(u). \qquad (3.40)$$

(Here and below we denote

$$\int_s^{\tau_U(t)} \Phi\, d\xi(u) = \int_s^{t} \mathbb{1}_{\{\tau_U > t\}}(\omega) \Phi\, d\xi(u).)$$

This and (3.25) immediately imply (3.39).   □

Remark 3.2 Note that under the assumptions of the lemma the expectation of the random variable $V(t, X(t))$ need not exist and thus formula (3.37) may be false.

---

[6]Lemma 3.2 is a special case of a formula of Dynkin [64, Sect. 5.1].

Remark 3.3 Setting $X(s) = x$ in formula (3.39) and letting $t \to s + 0$, we get

$$\mathfrak{A}V = \lim_{h \to +0} \frac{\mathbf{E}^{s,x} V(\tau_U(s+h), X(\tau_U(s+h))) - V(s,x)}{\mathbf{E}^{s,x}[\tau_U(s+h) - s]} = LV(s,x). \quad (3.41)$$

The operator on the left of (3.41) may be regarded as an extension of the differential operator L. It was first introduced in a more general setting by Dynkin in [62].

3.4 Conditions for Regularity of the Solution[7]

It follows from Theorem 3.4 that if condition (3.32) holds for all $t > t_0$ then the solution $X(t)$ of (3.31) is defined and continuous for all $t > t_0$. Condition (3.32) is rather restrictive. For example, it is intuitively clear that the problem $dX(t) = -X^3(t) \, dt + d\xi(t)$, $X(0) = x_0$, has a unique solution for all $t > 0$ (since the drift coefficient "directs" the motion to the origin), but conditions (3.32) hold for this equation only in a compact domain of the x-space. The same applies to Example 3.5 at the end of this section and also to an important class of stochastic equations which arise in the statistical analysis of partially observable Markov processes (see [256], [250]). It is therefore of paramount importance to find other, broader conditions for the existence and uniqueness of the solution of (3.31) for all $t > t_0$. In this section we shall prove analogs of Theorems 1.3 and 1.7 for Itô stochastic equations.

If conditions (3.32) are valid in every cylinder $I \times U_R$, one can construct a sequence of functions $b_n(t,x)$ and $\sigma_r^{(n)}(t,x)$ such that for $|x| < n$

$$\sigma_r^{(n)}(t,x) = \sigma_r(t,x); \qquad b_n(t,x) = b(t,x),$$

and for each b_n, $\sigma_r^{(n)}$ satisfy conditions (3.32) everywhere in E. By Theorem 3.4, there exists a sequence of Markov processes $X_n(t)$ corresponding to the functions b_n, $\sigma_r^{(n)}$. To simplify matters, we shall limit the discussion to the case in which the distribution of $X_0(t)$ has compact support in $\mathbb{R}^l$. Then it is intuitively clear (and it can be proved rigorously; see [55, 64]) that the first exit times of the processes $X_m(t)$ from the set $|x| < n$ are identical for $m \geq n$. Let this common value be τ_n. It is also clear that the processes themselves coincide up to time τ_n, i.e.,

$$\mathbf{P}\left\{\sup_{t_0 \leq t \leq \tau_n} |X_n(t) - X_m(t)| > 0\right\} = 0, \quad m > n.$$

Let τ denote the (finite or infinite) limit of the monotone increasing sequence τ_n as $n \to \infty$. We call the random variable τ the first exit time of the sample function

---

[7]Necessary and sufficient conditions for the regularity of homogeneous Markov processes of the diffusion type on $\mathbb{R}^1$ were obtained by Feller [71] for the time-homogeneous case. Multidimensional time-homogeneous processes were studied by the author in [117] where it was shown that the conditions of Theorem 3.5 are also necessary for regularity in the nondegenerate case.

from every bounded domain, or briefly the explosion time. This definition is natural, since one easily shows that the values of τ are not changed if we replace the domains $U_n = \{|x| < n\}$ by any other expanding sequence of bounded domains such that the distance from the origin to the boundary tends to infinity.

We now define a new stochastic process $\widetilde{X}(t)$ by setting $\widetilde{X}(t) = X_n(t)$ for $t < \tau_n$. It can be shown that this is always a Markov process for $t < \tau$ (for the definition of a Markov process stopped at a random time τ, see [63]).

We shall say that the process $X(t)$ is *regular* if for any $(s, x) \in E$

$$\mathbf{P}^{s,x}\{\tau = \infty\} = 1. \tag{3.42}$$

If condition (3.42) is satisfied, the process $X(t)$ is almost surely defined for all $t \geq s$. For a process satisfying the assumption of Theorem 3.4, regularity follows from continuity. The following theorem gives a more general sufficient condition for regularity.

Theorem 3.5 *Suppose that conditions* (3.32) *are valid in every cylinder $I \times U_R$ and, moreover, that there exists a nonnegative function $V \in \mathbf{C}_2$ on the domain E such that for some constant $c > 0$*

$$LV \leq cV, \tag{3.43}$$

$$V_R = \inf_{|x|>R} V(t, x) \to \infty \quad as\ R \to \infty. \tag{3.44}$$

Then parts 1, 2 and 5 of Theorem 3.4 hold. Part 3 is valid if the expectations in (3.34) *are replaced by "truncated" expectations (i.e., for example, instead of $\mathbf{E}[X^{s,x}(s + h) - x]$ one considers $\mathbf{E}\{\mathbb{1}(\omega)[X^{s,x}(s + h) - x]\}$, where $\mathbb{1}(\omega)$ is the indicator function of the set $|X^{s,x}(s + h, \omega) - x| < k$). Moreover, this process also satisfies the inequality*

$$\mathbf{E}V(t, X(t)) \leq \mathbf{E}V(t_0, X(t_0))e^{c(t-t_0)}, \tag{3.45}$$

if the expectation on the right exists.

Proof We first prove that under the assumptions (3.43) and (3.44) the process $\widetilde{X}(t)$ constructed at the beginning of the section is regular. From (3.43) it follows that the function

$$W(t, x) = V(t, x)\exp\{-c(t - t_0)\}$$

satisfies $LW \leq 0$. Hence, by Lemma 3.2, for $\tau_n(t) = \min(\tau_n, t)$, we have

$$\mathbf{E}\{V(\tau_n(t), X(\tau_n(t)))\exp[-c(\tau_n(t) - t_0)]\} - \mathbf{E}V(t_0, X(t_0))$$

$$= \mathbf{E}\int_{t_0}^{\tau_n(t)} LW(u, X(u))\,du \leq 0.$$

This, together with the inequalities $\tau_n(t) \leq t$, $V \geq 0$, implies

$$\mathbf{E}V(\tau_n(t), \widetilde{X}(\tau_n(t))) \leq e^{c(t-t_0)}\mathbf{E}V(t_0, X(t_0)). \tag{3.46}$$

From (3.46) we derive the estimate

$$\mathbf{P}\{\tau_n \leq t\} \leq \frac{e^{c(t-t_0)}\mathbf{E}V(t_0, \widetilde{X}(t_0))}{\inf_{|x|\geq n,\, u>t_0} V(u, x)}.$$

Letting $n \to \infty$ and making use of (3.44), we now get (3.42); thus the process $\widetilde{X}(t)$ is a solution of (3.31) for all $t \geq t_0$. This solution is unique up to equivalence. Indeed, it follows from the definition of $\widetilde{X}(t)$ and from the uniqueness of the solution of (3.31) in the domain $|x| < n$ that for every pair of solutions $X(t)$ and $Y(t)$

$$\mathbf{P}\left\{ \sup_{0<t<\tau_n} |X(t) - Y(t)| > 0 \right\} = 0.$$

The desired result now follows by letting $n \to \infty$ and using (3.42).

All other properties of the process just constructed can be proved in a similar manner. For example, we can prove relation (3.45) by letting $n \to \infty$ in (3.46) and using Fatou's lemma (see [97]). When constructing the process $\widetilde{X}(t)$, we assumed that the distribution of $X(t_0)$ has compact support. The case of an arbitrary initial distribution may now be dealt with in the way described in [92, Sect. 8.3]. □

Remark 3.4 It is intuitively clear that whether a process is regular or not depends only on the behavior of the coefficients b and σ_r in a neighborhood of the point at infinity. It is therefore natural to expect the conclusion of Theorem 3.5 to remain valid if a function V satisfying conditions (3.43), (3.44) exists merely in the domain $\{t > 0\} \times U_R^c$ for some $R > 0$. One easily sees from the proof of Theorem 3.5 that in this case the sample function cannot escape to infinity before it exits from the set $|x| > R$. Using the strong Markov property of the process (see Chap. 4), one easily infers that the process is also regular in the sense of definition (3.42).

Remark 3.5 In many cases (see Chap. 4) it is useful to know when the sample function of a diffusion process almost surely does not exit from a given open set D in a finite time. Sufficient conditions for the invariance of the set D in this sense can be derived from Theorem 3.5 by noting that in the proof of this theorem the assumption that the sequence U_n converges to $\mathbb{R}^l$ is not essential. Replacing the sequence U_n by an increasing sequence of open sets D_n whose closures are contained in D, and such that $\bigcup D_n = D$, we get the following result.

Corollary 3.1 *Suppose that in every cylinder $I \times D_n$ the coefficients b and σ_r satisfy conditions (3.32) and there exists a function $V(t, x)$, twice continuously differentiable in x and continuously differentiable in t in the domain $I \times D$, which satisfies condition (3.43) and the condition*

$$\inf_{t>0,\, x\in D\setminus D_n} V(t, x) \to \infty \quad as\ n \to \infty.$$

Then the conclusion of Theorem 3.5 holds provided that also $\mathbf{P}\{X(t_0) \in D\} = 1$. *Moreover the solution satisfies the relation*

$$\mathbf{P}\{X(t) \in D\} = 1 \quad \text{for all } t \geq t_0.$$

As the following theorem shows, conditions (3.43), (3.44) are in a certain sense "the best possible".

Theorem 3.6 *Suppose that conditions* (3.32) *hold in every cylinder* $(t > 0) \times U_R$ *and that, moreover, there exists in the domain* $E = (t > 0) \times \mathbb{R}^l$ *a nonnegative bounded function* $V(t, x) \in \mathbf{C}_2$ *such that for some* $c > 0$

$$LV \geq cV. \tag{3.47}$$

Then for all points s, x *such that* $V(s, x) > 0$ *the process* $X^{s,x}(t)$ *defined by* (3.33) *up to the time* $\tau^{s,x}$ *of first exit from every bounded domain, is not regular. More precisely, for each* $\varepsilon > 0$ *we have*

$$\mathbf{P}\left\{\tau^{s,x} - s < \frac{1}{c} \ln \frac{\sup_E V}{V(s, x)} + \varepsilon\right\} > 0.$$

Proof Exactly as in the proof of Theorem 3.5, we apply Lemma 3.2 and condition (3.47) to get the relation

$$\mathbf{E}V(\tau_n^{s,x}(t), X(\tau_n^{s,x}(t))) \exp\{-c(\tau_n^{s,x}(t) - s)\} \geq V(s, x).$$

Since V is bounded, this implies that

$$\mathbf{E} \exp\{-c(\tau_n^{s,x}(t) - s)\} \sup_E V \geq V(s, x).$$

Letting $n \to \infty$ and putting $\tau^{s,x}(t) = \min(\tau^{s,x}, t)$, we get

$$\mathbf{E} \exp\{-c(\tau^{s,x}(t) - s)\} \sup_E V \geq V(s, x). \tag{3.48}$$

Assume that $\mathbf{P}\{\tau^{s,x} > t\} = 1$ for some $t > s$. Then (3.48) implies the inequality

$$\exp\{-c(t - s)\} \geq V(s, x)/\sup_E V,$$

which is in contradiction with $t - s > c^{-1} \ln(\sup_E V/V(s, x))$. This completes the proof. $\qquad\square$

We append a very brief discussion of the situation arising under the assumptions of Theorem 3.6 when the process $X^{s,x}(t)$ has an explosion. In this case (3.33) determines the process only up to the random time $\tau^{s,x}$. How does the process continue for $t > \tau^{s,x}$? There exists an infinite set of possible continuations. For example, we can set $X^{s,x}(\tau^{s,x} + 0) = y \in \mathbb{R}^l$ (a jump to the point y after explosion), and we

can stipulate other additional conditions determining the evolution of the process after $\tau^{s,x}$. There is an extensive literature on the continuation of Markov processes after the time at which the sample function has hit the boundary. The problem is closely connected with the description of all possible ways in which this process can reach the boundary (assuming that it is a Martin boundary). The case of a one-dimensional time-homogeneous process defined by (3.33) (and a somewhat more general one) has been thoroughly studied by Feller [73] and Ventsel [266]. For the multi-dimensional case extremely interesting results have been obtained by Ventsel [267], Ueno [264, 265] and others.

Example 3.1 Let conditions (3.32) hold everywhere in E. Then the function $V = (|x|^2 + 1)^{n/2}$ satisfies the assumptions of Theorem 3.5 for every $n > 0$. Hence the solution $X(t)$ of (3.31) exists and is almost surely bounded on every finite time interval (this result follows from Theorem 3.4). Moreover, for some constant c_n we have the estimate

$$\mathbf{E}|X(t)|^n < e^{c_n(t-t_0)}\mathbf{E}|X(t_0)|^n.$$

Example 3.2 Suppose that condition (3.32) holds in every bounded x-domain and that for $x \in \mathbb{R}^l$

$$|b| < B(1 + |x| \ln |x|), \tag{3.49}$$

$$\sum_{r=1}^{k} |\sigma_r|^2 < B(1 + |x|^2 \ln |x|). \tag{3.50}$$

Then, using the auxiliary function $V = \ln(|x|^2 + 1)$ and applying Theorem 3.5, we conclude that the solution of (3.31) is regular.

It follows from (3.49) that the inner product $(b(s,x), x)$ increases no faster than $|x|^2 \ln(|x|^2 + 1)$ as $x \to \infty$.

The next example shows that if the rate of increase of $(b(s,x), x)$ is slightly higher, the process fails to be regular.

Example 3.3 Suppose that condition (3.50) holds and that for some $\varepsilon > 0$

$$(b(s,x), x) > |x|^2[\ln(|x|^2 + 1)]^{1+\varepsilon}.$$

Then the process $X(t)$ has an explosion for every initial condition. To prove this it suffices to apply Theorem 3.6 with the auxiliary function

$$V(x) = \exp\{-[\ln(|x|^2 + 1)]^\varepsilon\}.$$

In this example, the sample function explodes because of large drift (the phenomenon also occurs for $\sigma \equiv 0$). It is not hard to find examples in which the lack of regularity is due to large diffusion.

Example 3.4 Consider a process $X(t)$ in $\mathbb{R}^1$ defined by (3.31). Then b and σ are scalar functions and the operator L has the form

$$L = \frac{\partial}{\partial s} + b(s, x) \frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(s, x) \frac{\partial^2}{\partial x^2}.$$

Using the same auxiliary function as in Example 3.3, we see that if

$$\sigma^2 > B(x^2 + 1)[\ln(x^2 + 1)]^{1+\varepsilon},$$

then the process $X(t)$ has an explosion for every function $b(s, x)$ which, with respect to the second variable, increases not faster than linear one. However, when the function $|b(s, x)|$ increases faster than linear, the process may nonetheless be regular. For example, when $b(s, x) = -x^5$ and $\sigma = x^3$, regularity follows from Theorem 3.5 where we should take the auxiliary function $(x^2 + 1)^\alpha$.

Example 3.5 The assumptions of Theorem 3.4 fail to hold even for the well known mechanical system described by the Van der Pol equation driven by white noise of constant intensity σ^2. This system can be described by the Itô equations:

$$dx_1 = x_2\, dt; \qquad dx_2 = [-x_1 + \varepsilon(1 - x_1^2)x_2]\, dt + \sigma\, d\xi(t).$$

The auxiliary function

$$V = \frac{1}{2}\left[x_2 + \varepsilon\left(\frac{x_1^3}{3} - x_1\right)\right]^2 + \frac{x_1^2}{2} + \frac{\sigma^2}{4\varepsilon}$$

satisfies the assumptions of Theorem 3.5 since

$$LV = x_2 \frac{\partial V}{\partial x_1} + [\varepsilon(1 - x_1^2)x_2 - x_1]\frac{\partial V}{\partial x_2} + \frac{\sigma^2}{2}\frac{\partial^2 V}{\partial x_2^2}$$

$$= -\varepsilon\frac{x_1^4}{3} + \varepsilon x_1^2 + \frac{\sigma^2}{2} \leq 2\varepsilon V.$$

Thus the process is regular.

3.5 Stationary and Periodic Solutions of Stochastic Differential Equations

In Sect. 3.2 we gave conditions implying the existence of periodic and stationary Markov processes, stated in terms of the properties of the transition probability functions. These conditions are of little use for SDEs, since the transition probability functions of such processes are usually not expressible in terms of the coefficients of the equation. Fortunately, however, one can state simple conditions in terms of Lyapunov functions for the required properties of the transition probability function. The following theorems are analogous to Theorem 2.6.

Theorem 3.7 *Suppose that the coefficients of* (3.31) *are independent of t and satisfy conditions* (3.32) *in U_R for every $R > 0$, and that there exists a function $V(x) \in \mathbf{C}_2$ in $\mathbb{R}^l$ with the properties*

$$V(x) \geq 0, \tag{3.51}$$

$$\sup_{|x|>R} LV(x) = -A_R \to -\infty \quad \text{as } R \to \infty. \tag{3.52}$$

Suppose moreover that the process $X^x(t)$ is regular for at least one $x \in \mathbb{R}^l$. Then there exists a solution of (3.31) *which is a stationary Markov process.*

Theorem 3.8 *Suppose that the coefficients of* (3.31) *are θ-periodic in t and satisfy conditions* (3.32) *in every cylinder $I \times U$, and suppose further that there exists a function $V(t, x) \in \mathbf{C}_2$ in E which θ-periodic in t, and satisfies condition* (3.52) *and the condition*

$$\inf_{|x|>R} V(t, x) \to \infty \quad \text{as } R \to \infty. \tag{3.53}$$

Then there exists a solution of (3.31) *which is a θ-periodic Markov process.*[8]

Proof of Theorem 3.7 Let $X^x(t)$ be a regular solution of (3.31) and let $V(x)$ satisfy conditions (3.51), (3.52). Lemma 3.2 implies that

$$\mathbf{E}V(X^x(\tau_n(t))) - V(x) = \mathbf{E} \int_0^{\tau_n(t)} LV(X^x(u))\, du.$$

(As before, we denote $\tau_n = \inf\{t : |X^x(t)| > n\}$, $\tau_n(t) = \min(\tau_n, t)$.) Estimating the right-hand side of this equality by means of the obvious inequality

$$LV(X^x(u)) \leq -\mathbb{1}_{\{|X^x(u)|>R\}}(\omega)A_R + \sup_{x \in \mathbb{R}^l} LV(x),$$

we get

$$A_R \mathbf{E} \int_0^{\tau_n(t)} \mathbb{1}_{\{|X^x(u)|>R\}}(\omega)\, du \leq c_1 t + c_2.$$

Since the process $X^x(t)$ is regular, it follows that almost surely $\tau_n(t) \to t$ as $n \to \infty$. Letting $n \to \infty$ and then changing the order of integration in the last inequality, we obtain

$$\frac{1}{t} \int_0^t P(x, u, U_R^c)\, du < \frac{c_3}{A_R}. \tag{3.54}$$

---

[8]Condition (3.52) of Theorems 3.7 and 3.8 may often be replaced by the weaker condition that $LV \leq -1$ outside some compact set (see Chap. 4).

It follows from (3.54), (3.52) and Theorem 3.1 that there exists a stationary initial distribution. The solution of (3.31) with this initial distribution is obviously stationary. □

Proof of Theorem 3.8 Arguments similar to those used in proving (3.54) yield the inequality

$$\frac{1}{T} \int_0^T P(s, x, s + u, U_R^c) \, du < \frac{c_3}{A_R}.$$

Therefore, to prove the theorem we need only to show that condition (3.22) is satisfied and then use the remark following Theorem 3.2.

To prove (3.22), we can again use the method of auxiliary functions. We may assume without loss of generality that the function V satisfying (3.52) and (3.53) also satisfies (3.51) (otherwise we may add to it a constant). Further, it follows from the assumptions of the theorem that $L(V(t, x) - kt) \leq 0$ for a sufficiently large constant k. Using this inequality, the regularity of the process (which follows from Theorem 3.5), and Lemma 3.2, we easily obtain that

$$\mathbf{E}^{s,x} V(t, X(t)) \leq k(t - s) + V(s, x).$$

Together with Chebyshev's inequality, this implies

$$P(s, x, t, U_R^c) \leq \frac{k(t - s) + V(s, x)}{\inf_{|x|>R} V(t, x)}.$$

Thus condition (3.22) will hold if $\beta(R)$ is chosen so that

$$\frac{\sup_{|x|<\beta(R)} V}{\inf_{|x|>R} V} \to 0 \quad \text{as } R \to \infty.$$

This is possible because (3.53) holds. The proof is complete. □

To demonstrate the range of application of Theorems 3.7 and 3.8, we present a few examples.

Example 3.6 Consider the auxiliary function

$$V_1(x) = \sum_{i,j=1}^l c_{ij} x_i x_j = (Cx, x),$$

where C is a positive definite matrix. We have

$$LV_1(x) = 2(Cx, b(t, x)) + \text{tr}(A(t, x)C).$$

Thus a sufficient condition for the existence of a stationary (periodic) solution of (3.31) in the case when the coefficients are independent of t (θ-periodic in t) is that

for some positive definite matrix C

$$2(Cx, b(t, x)) + \text{tr}(A(t, x)C) \to -\infty \quad \text{as } |x| \to \infty,$$

and that the coefficients b and σ_r satisfy conditions (3.32) in every cylinder $I \times U_R$.

Example 3.7 [9] Let us investigate conditions for the existence of a stationary process at the output of a system described by a Lienard equation driven *by* Gaussian white noise $\dot{\xi}(t)$.

The system is described by the equation

$$x'' + f(x)x' + g(x) = \sigma \dot{\xi}(t), \tag{3.55}$$

where σ^2 is the intensity of the white noise at the input (assumed constant for simplicity's sake). Setting $y(t) = x'(t)$, one easily sees that the pair $(x(t), y(t))$ is a Markov process satisfying the following system of Itô equations

$$dx = y\,dt; \qquad dy = [-yf(x) - g(x)]\,dt + \sigma\,d\xi(t)$$

with the generator

$$L = y\frac{\partial}{\partial x} - [yf(x) + g(x)]\frac{\partial}{\partial y} + \frac{\sigma^2}{2}\frac{\partial^2}{\partial y^2} + \frac{\partial}{\partial t}.$$

As in Example 2.3, we set

$$F(x) = \int_0^x f(u)\,du; \qquad G(x) = \int_0^x g(u)\,du; \qquad p(x) = \gamma \arctan x$$

and consider the function

$$V_1(x, y) = \frac{y^2}{2} + [F(x) - p(x)]y + G(x) + \int_0^x f(u)[F(u) - p(u)]\,du + k,$$

which is analogous to the auxiliary function of the above-mentioned example. It is easy to check that conditions (3.51)–(3.53) hold for $V_1(x, y)$ for a suitable choice of γ and k, if the conditions

$$\begin{aligned} \text{sign } g(x) = \text{sign } x \qquad & \text{for } |x| > x_0, \\ g(x)F(x) - \delta|g(x)| \to \infty \qquad & \text{as } |x| \to \infty, \\ G(x) + \delta \int_0^x \frac{F(u)}{1+u^2}\,du \to \infty \qquad & \text{as } |x| \to \infty \end{aligned} \tag{3.56}$$

hold for some $\delta > 0$, $x_0 > 0$.

---

[9]Due to Nevelson.

Thus, if conditions (3.56) are satisfied, the system (3.55) has a stationary output. Conditions (3.56) will obviously hold for the process at the output of a system described by the Van der Pol equation with $f(x) = \varepsilon(x^2 - 1)$, $g = x$. One can give similar conditions which imply the existence of a periodic solution in the case when the driving process contains a periodic component.

Example 3.8 The following fact is well-known in the theory of systems of ODEs: If x_0 is an asymptotically stable equilibrium point of an autonomous system $dx/dt = F(x)$ and $f(t)$ is θ-periodic, then for sufficiently small ε the system $dx/dt = F(x) + \varepsilon f(t)$ has a θ-periodic solution in a neighborhood of the equilibrium point. The problem of extending this result to systems of differential equations describing Markov processes was suggested by N.N. Krasovskii. The method developed in this section can be employed to this end.

Let $X(t)$ be a time-homogeneous stochastic process described by the system (3.31). We consider another process which differs from (3.31) by the presence of a "force" $f(x, t)$ which is θ-periodic in t (the dependence on x means that the value of f may depend on the state of the system):

$$dY(t) = b(Y(t))\, dt + \sum_{r=1}^{k} \sigma_r(Y(t))\, d\xi_r(t) + f(t, Y(t))\, dt. \qquad (3.57)$$

Assume that the unperturbed system ($f(t, y) \equiv 0$) has a stationary solution, and that there exists a function V satisfying conditions (3.51)–(3.53). Then the system (3.57) will have a θ-periodic solution, provided the additional condition $(\mathrm{grad}\, V, f) < c$ holds for some constant c. The proof follows directly from the fact that then the function V satisfies the assumptions of Theorem 3.8 for the process $Y(t)$. It is not difficult to apply this type of reasoning also when the perturbing force itself is a periodic stochastic process. We shall not dwell on this here.

3.6 Stochastic Equations and Partial Differential Equations

The method of investigation of Markov processes based on studying the properties of the solution of some generalized differential equation for the sample functions of the process was, historically speaking, a later development. An earlier method originating in the work of Einstein and Smoluchowski, was subsequently perfected by Kolmogorov, Feller, Dynkin and others. The main idea of the earlier method is to investigate the properties of the solutions of differential equations whose unknowns are expectations of various functions of the processes in question. It turns out that in many cases these expectations are solutions of boundary-value problems for linear parabolic and elliptic equations. The converse is also true. The solution of the first boundary-value problem for the general linear elliptic or parabolic equation admits a probabilistic representation in terms of the expectations of certain functionals of the process $X(t)$.

We shall consider this representation in a few special cases. In so doing we shall confine ourselves to domains in E which are Cartesian products of a certain domain $U \subset \mathbb{R}^l$ having a sufficiently smooth boundary and of a closed interval on the real axis.[10]

Throughout this section we shall consider classical solutions of equations of the types $Lu = -g(s, x)$ or $Lu + q(s, x)u = -g(s, x)$. Here

$$Lu \equiv \frac{\partial u}{\partial s} + \sum_{i=1}^{l} b_i(s, x) \frac{\partial u}{\partial x} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(s, x) \frac{\partial^2 u}{\partial x_i \partial x_j}.$$

By a solution we shall mean a function $u(s, x)$ which is twice continuously differentiable in x and continuously differentiable in s, and such that $Lu = -g$ $(Lu + qu = -g)$ at each point of the given domain.

In case the function u does not depend on s, we shall retain the notation Lu for the "elliptic" operator

$$\sum_{i=1}^{l} b_i(x) \frac{\partial u}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}$$

(this operator is not necessarily elliptic, since the matrix A may be singular).

We shall start from the probabilistic representation for the solution of the equation $Lu = 0$.

Lemma 3.3 *Let $f(s, x)$ be a function, bounded and continuous on the closure of the domain $(t_0, T) \times U$, and satisfying the equation $Lf = 0$ on that domain. Assume that there exists a unique and, if the domain U is unbounded, regular Markov process $X(t)$ associated with the operator L up to the time τ_U at which the process reaches the boundary of U. Then*

$$f(s, x) = \mathbf{E}^{s,x} f(\tau_U(T), X(\tau_U(T))). \tag{3.58}$$

Proof The proof for a bounded domain U follows easily from Lemma 3.2. For $n \to \infty$ we consider an increasing sequence of domains $U^{(n)} \uparrow U$ such that for every n there exists a function $f_n(s, x)$ which is equal to $f(s, x)$ for $x \in U^{(n)}$ and has two continuous derivatives with respect to x and one with respect to s for all $x \in \mathbb{R}^l$, $s \in (t_0, T)$. Applying Lemma 3.2 to the function $f_n(s, x)$ and then letting $n \to \infty$, we get (3.58). The proof for an unbounded domain is similar, except that one considers a sequence of "truncated" domains $U_n = U \cap \{|x| < n\}$ and, using the regularity of the process, one lets $n \to \infty$. $\qquad\square$

Remark 3.6 The stochastic process $X(\tau_U(t))$ is almost surely confined either to the "lateral" surface of the cylinder $[t_0, T] \times \Gamma$ or to its "upper" base $\{s = T\} \times U$.

---

[10]There are numerous papers on the probabilistic representation of solutions of the first boundary-value problem for elliptic and parabolic equations. See e.g., [64, 119, 164, 223].

Hence it follows that under the assumptions of Lemma 3.3 there exists at most one solution of the equation $Lu = 0$ which takes on given values on the set $\Gamma_1 = [t_0, T] \times \Gamma \cup \{s = T\} \times U$. In particular, if $U = \mathbb{R}^l$, we get from (3.58) the following formula for the solution of the Cauchy problem for a parabolic equation:

$$f(s, x) = \mathbf{E}^{s,x} f(T, X(T)) = \int_{\mathbb{R}^l} f(T, y) P(s, x, T, dy). \qquad (3.59)$$

Remark 3.7 If the domain U is bounded, then the assertion of Lemma 3.3 will hold, for instance, if we require that the coefficients b, σ_r satisfy conditions (3.32) in the domain $(t_0, t) \times U$. For an unbounded domain U, it is sufficient to require that conditions (3.32) hold in every compact set and that there exists a function V satisfying conditions (3.43), (3.44) (see Theorem 3.5).

Remark 3.8 It is a well known fact in the theory of differential equations that a sufficient condition for the existence of solutions both to the mixed problem and the Cauchy problem for a parabolic equation is that, in addition to (3.32), the following nondegeneracy condition

$$\sum_{i,j=1}^{l} a_{ij}(s, x)\lambda_i \lambda_j > m(s, x) \sum_{i=1}^{l} \lambda_i^2 \qquad (3.60)$$

is satisfied (here $m(s, x)$ is a positive continuous function on E). Thus, if (3.60) holds, there exists a function f taking on given values on Γ_1 and such that $Lf = 0$. Consequently, formulas (3.58), (3.59) furnish a classical solution of these boundary-value problems for the equation $Lu = 0$. If the function $f(s, x)$ defined by (3.58) or (3.59) is not differentiable, we may regard it as a generalized solution of the equation $Lu = 0$. (It is readily shown that it satisfies the equation $\mathfrak{A}u = 0$; see (3.41).) However, it can be shown that condition (3.60) need not hold in order for the function $f(s, x)$ defined by (3.59) to be differentiable provided that the "initial" function $f(T, x)$ is sufficiently smooth (see, e.g., [92]). The situation is different for the solution of the mixed problem.

Remark 3.9 Comparing formula (3.59) with the formula

$$f(s, x) = \int f(T, y) p(s, x, T, y) \, dy,$$

which expresses the solution of the Cauchy problem for a parabolic equation in terms of a fundamental solution $p(s, x, T, y)$, one can verify that if the parabolic equation $Lu = 0$ has a fundamental solution, then the transition probability function $P(s, x, T, A)$ has a density with respect to Lebesgue measure in $\mathbb{R}^l$ and this density is $p(s, x, T, y)$.

Let us now consider a time-homogeneous process $X(t)$.

Lemma 3.4 *Suppose that the coefficients b and σ_r are independent of t and that they satisfy conditions* (3.32) *in every bounded domain. Let U be a bounded domain in $\mathbb{R}^l$. Let $f(x)$ be twice continuously differentiable with respect to x_i in U, continuous on the closure of U, and such that $Lf = 0$. Assume moreover that for every point $x \in U$*

$$\mathbf{P}^x\{\tau_U < \infty\} = 1. \tag{3.61}$$

Then

$$f(x) = \mathbf{E}^x f(X(\tau_U)) = \int_\Gamma \mathbf{P}^x\{X(\tau_U) \in dy\} f(y) \tag{3.62}$$

holds. Formula (3.62) *remains valid for an unbounded domain U if the process $X(t)$ is regular and the function $f(x)$ is bounded in U.*

Proof It will suffice to prove the lemma for a function $f(x) \in \mathbf{C}_2$; the general case can then be treated by means of a suitable limiting process, as in the proof of Lemma 3.3. If $f(x) \in \mathbf{C}_2$, we infer from Lemma 3.3 that

$$f(x) = \mathbf{E}^x\{f(X(\tau_U))\mathbb{1}_{\{\tau_U \le t\}}(\omega)\} + \mathbf{E}^x\{f(X(t))\mathbb{1}_{\{\tau_U > t\}}(\omega)\}$$

holds for each t. Letting $t \to \infty$ and using (3.61), we get (3.62).   □

Remark 3.10 It follows from Lemma 3.4 that in particular the Dirichlet problem for the "elliptic" equation

$$\sum_{i=1}^l b_i(x)\frac{\partial u}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^l a_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} = 0$$

in the domain U has at most one solution if condition (3.61) holds. It is known from the theory of differential equations that the Dirichlet problem has a solution if the coefficients satisfy conditions (3.32) and (3.60). If condition (3.60) fails to hold, the generalized solution (3.62) may turn out to be discontinuous. In the next section we shall present a sufficient condition for (3.61) to hold; this will be related to the existence of an auxiliary (Lyapunov) function. Comparing formula (3.62) with the formula (see [202])

$$f(x) = \int_\Gamma K(x, y) f(y) \sigma_\Gamma(dy)$$

(where $K(x, y)$ is the normal derivative of the Green function), we see that the measure $\mathbf{P}^x\{X(\tau_U) \in \gamma\}$ has density $K(x, y)$ with respect to the "surface area" $\sigma_\Gamma(dy)$ on Γ, and this density is bounded from above and below uniformly in x on every compact subset of U.

Let us now discuss the probabilistic representation of the solution of the inhomogeneous equation

$$Lf = -g. \tag{3.63}$$

We need only consider zero boundary conditions.

Lemma 3.5 *Let U be a bounded domain in $\mathbb{R}^l$. Let $g(s, x)$ be a continuous bounded function. Let $f(s, x)$ be a function twice continuously differentiable with respect to x_i and continuously differentiable with respect to s in the domain $(t_0, T) \times U$, which is continuous on the closure of this domain, vanishes on the set $\Gamma_1 = [t_0, T] \times \Gamma \cup \{s = T\} \times U$, and satisfies equation* (3.63). *Assume moreover that in every domain which is bounded with respect to x, conditions* (3.32) *are valid. Then*

$$f(s, x) = \mathbf{E}^{s,x} \int_s^{\tau_U(T)} g(u, X(u)) \, du. \tag{3.64}$$

This representation remains valid for an unbounded domain U if the process $X(t)$ is regular and the functions f and g are bounded in $(t_0, T) \times U$.

The proof, analogous to the proofs of Lemmas 3.3 and 3.4, is left to the reader.

Lemma 3.6 *Let the coefficients b and σ_r be independent of t and suppose they satisfy conditions* (3.32) *in every bounded domain. Let $g(x)$ be continuous and bounded in $U \cup \Gamma$ and let $f(x)$ be twice continuously differentiable and bounded in U, vanishing on Γ, and such that*

$$Lf = \sum_{i=1}^{l} b_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} = -g.$$

Assume moreover that for $x \in U$

$$\mathbf{E}^x \tau_U < c. \tag{3.65}$$

Then the function f can be written as

$$f(x) = \mathbf{E}^x \int_0^{\tau_U} g(X(t)) \, dt.$$

The proof is analogous to that of Lemma 3.4.

Remark 3.11 Condition (3.65) is satisfied, e.g., if U is a bounded domain in which the nondegeneracy condition (3.60) holds with a function $m(s, x) = m(x)$ which is positive on the closure of U (see Corollary 3.3 in the next section). For an unbounded domain, condition (3.65) does not hold, as a rule. Then Lemma 3.6 is of little use. In the next section we shall prove a better result for this case.

The following modification of Lemma 3.3 is known as the Feynman–Kac formula.

Lemma 3.7 *Let the conditions of Lemma* 3.3 *be fulfilled with the replacement the equation $Lf = 0$ to the equation*

$$Lf + q(s, x)f = 0; \qquad (3.66)$$

here $q(s, x)$ is the bounded and continuous in $(t_0, T) \times U$ function. Then

$$f(s, x) = \mathbf{E}^{s,x} \left\{ f(\tau_U(T), X(\tau_U(T))) \exp \left[\int_s^{\tau_U(T)} q(t, X(t)) \, dt \right] \right\}.$$

Proof Let $X(t)$ be the solution of SDE with the generator L, and $X(s) = x$. Let the process $Y(t)$ be defined by the equation $dY(t) = q(t, X(t))Y(t) \, dt$; $Y(s) = y$. Then

$$Y(t) = y \exp \left\{ \int_s^t q(s_1, X(s_1)) \, ds_1 \right\}. \qquad (3.67)$$

From the other hand, the pair $(X(t), Y(t))$ is a Markov diffusion process with the generator

$$\widetilde{L}(s, x, y) = L(s, x) + q(s, x) \, y \, \frac{\partial}{\partial y}. \qquad (3.68)$$

We further have due to (3.66), (3.68)

$$\widetilde{L}(f(s, x) \cdot y) = yLf(s, x) + q(s, x)yf(s, x) = y(Lf + qf) = 0.$$

So, applying again Dynkin's formula for the process $(X(t), Y(t))$ to the function $v(s, x, y) = f(s, x)y$ and making use of (3.67), we have

$$y\mathbf{E}^{s,x} \left\{ f(\tau_U(T), X(\tau_U(T))) \exp \left[\int_s^{\tau_U(T)} q(s_1, X(s_1)) \, ds_1 \right] \right\} = yf(s, x).$$

Lemma is proved. $\qquad \qquad \square$

Proof of the following lemma is completely analogous, and it is left to the reader again.

Lemma 3.8 *Let the conditions of Lemma* 3.5 *be valid with the replacement of* (3.63) *to the equation*

$$Lf + q(s, x)f = -g(s, x).$$

Then

$$f(s, x) = \mathbf{E}^{s,x} \left\{ f(\tau_U(T), X(\tau_U(T))) \exp \left(\int_s^{\tau_U(T)} q(t, X(t)) \, dt \right) \right\}.$$

Remark 3.12 Analogs of Lemmas 3.4 and 3.6 for the time-homogeneous case can be proved by similar way, if $q(x) \leq 0$. For the case $q(x) > 0$ see [130].

Remark 3.13 Let the coefficients of L and q do not depend on s. Then with help of the substitution $t = T - s$ we arrive in particular to the following representation of the Cauchy problem

$$\frac{\partial u(t, x)}{\partial t} = L(x)u(t, x) + q(x)u(t, x); \qquad u(0, x) = \varphi(x);$$

$$u(t, x) = \mathbf{E}^x \left\{ \varphi(X(t)) \exp\left[\int_0^t q(X(t_1)) \, dt_1 \right] \right\}. \tag{3.69}$$

3.7 Conditions for Recurrence and Finiteness of Mean Recurrence Time

Let U_1 be some (bounded or unbounded) domain, and denote its complement U_1^{c} by U. A process $X(t)$ is said to be *recurrent relative to the domain* U_1 (or U_1-*recurrent*) if it is regular, and for every $s, x \in U$

$$\mathbf{P}^{s,x}\{\tau_U < \infty\} = 1. \tag{3.70}$$

Recurrence is an extremely important concept for the investigation of the properties of sample functions for large time values. A simple condition for recurrence is given by

Theorem 3.9 *A process $X(t)$ is recurrent relative to the domain U_1 if it is regular and there exists in $I \times U$ a nonnegative function $V(s, x)$, twice continuously differentiable with respect to x and continuously differentiable with respect to s, such that*

$$LV(s, x) \leq -\alpha(s), \tag{3.71}$$

where $\alpha(s) \geq 0$ is a function for which

$$\beta(t) = \int_0^t \alpha(s) \, ds \to \infty \quad \text{as } t \to \infty. \tag{3.72}$$

Moreover, the expectation of the random variable $\beta(\tau_U)$ exists and satisfies the inequality

$$\mathbf{E}^{s,x}\beta(\tau_U) \leq \beta(s) + V(s, x). \tag{3.73}$$

Proof We define a random variable $\tau_U^{(n)}(t)$ by $\tau_U^{(n)}(t) = \min\{\tau_U, t, \tau_n\}$. (As before, $\tau_n = \inf\{t : |X(t)| = n\}$.) By Lemma 3.2,

$$\mathbf{E}^{s,x} V(\tau_U^{(n)}(t), X(\tau_U^{(n)}(t))) - V(s, x) = \mathbf{E}^{s,x} \int_s^{\tau_U^{(n)}(t)} LV(u, X(u)) \, du$$

and therefore, by (3.71),

$$\mathbf{E}^{s,x}\beta(\tau_U^{(n)}(t)) \le \beta(s) + V(s,x).$$

Since $\tau_U^{(n)}(t) \to \tau_U(t)$ almost surely as $n \to \infty$ (since the process is regular), it follows via Fatou's lemma that

$$\mathbf{E}^{s,x}\beta(\tau_U(t)) \le \beta(s) + V(s,x). \tag{3.74}$$

Hence it follows that

$$\mathbf{P}^{s,x}\{\tau_U \ge t\} \le \frac{\beta(s) + V(s,x)}{\beta(t)}.$$

Letting $t \to \infty$ and using (3.72), we see that (3.70) holds. The other part of the theorem follows from (3.74) if we let $t \to \infty$ and again use Fatou's lemma. $\square$

Corollary 3.2 *Suppose that the assumptions of Theorem* 3.9 *hold, with a function* $\alpha(s) = cs^{n-1}$ $(c > 0)$. *Then the random variable* τ_U *has an n-th moment, and*

$$\mathbf{E}^{s,x}\tau_U^n - s^n \le \frac{n}{c}V(s,x).$$

In particular we see that the recurrence time for the bounded domain U_1 *has finite expectation, if* $\inf_{|x|>R} V \to \infty$ *as* $R \to \infty$, $V \ge 0$, *and* $LV \le -c$ *in the domain* $I \times U$ *for some positive constant* c.

Let the domain $U = U_1^c$ be bounded with respect to one of the coordinates, i.e., $U \subset \{x_i^{(0)} < x_i < x_i^{(1)}\}$. Suppose that the diffusion is nonsingular in the same coordinate uniformly in s and that the function $b_i(s,x)$ is bounded from above or below, so that

$$0 < a_0 < a_{ii}(s,x), \qquad b_i(s,x) < b_0 \quad (\text{or } b_i(s,x) > b_0). \tag{3.75}$$

Consider the auxiliary function

$$V(s,x) = e^{\gamma s}[k - (x_i + \beta)^{2n}],$$

where the constants γ, k, β and n will be specified later. Obviously,

$$LV(s,x) = e^{\gamma s}[\gamma k - \gamma(x_i + \beta)^{2n} - 2nb_i(s,x)(x_i + \beta)^{2n-1}$$
$$- 2n(2n-1)a_{ii}(s,x)(x_i + \beta)^{2n-2}].$$

Assume for definiteness that $b_i(s,x) < b_0$. We now choose $\beta < 0$ to be large enough in absolute value so that $x_i + \beta < 1$, we determine the number n by the condition $(2n-1)a_0 > b_0 \sup_U |x_i + \beta| + 1$, and finally we set $k = \sup_U (x_i + \beta)^{2n}$, $\gamma = k^{-1}$.

Then there is a positive constant c such that the inequality $LV(s,x) < -ce^{\gamma s}$ holds.

We may now draw the following conclusion from Theorem 3.9:

Corollary 3.3 *If $U \subset \{x_i^{(0)} < x_i < x_i^{(1)}\}$ and conditions (3.75) hold in U, then the process $X(t)$ is recurrent relative to U_1. Moreover the random variable τ_U then has moments of arbitrary order, actually $\mathbf{E}_{s,x} e^{\gamma \tau_U}$ exists for sufficiently small γ.*

A similar conclusion is valid if instead of (3.75) we assume that

$$b_i(s,x) > b_0 > 0 \quad (\text{or } b_i(s,x) < -b_0 < 0) \tag{3.76}$$

holds in U. In this case, we must set $V(s,x) = e^{\gamma s}(k - x_i)$. Note that conditions (3.75) are satisfied if U is bounded, the coefficients of the operator L are bounded in U, and the matrix A is nonsingular in this domain.

Let us now devote some further attention to time-homogeneous processes, confining the discussion for the sake of simplicity to the case in which the nondegeneracy condition (3.60) holds in every bounded domain.

Theorem 3.10 *Let U_1 be a bounded domain whose boundary Γ is regular[11] relative to $U = U_1^c$ and suppose that the coefficients b, σ_r are independent of t and satisfy conditions (3.32) in every compact set. Assume further (3.60), and that the corresponding Markov process $X(t)$ is regular.*

Then the process $X(t)$ is recurrent relative U_1 if and only if the Dirichlet problem

$$Lu = \sum_{i=1}^{l} b_i(x) \frac{\partial u}{\partial x} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} = 0;$$

$$u|_\Gamma = f(s) \tag{3.77}$$

has a unique bounded solution in U.

Proof Let $\tau_U^{(n)} = \min(\tau_U, \tau_n)$ denote the time at which the sample function of the process first exits from the domain $U \cap U_n$ ($U_n = \{|x| < n\}$). It follows from the assumptions of the theorem and from Corollary 3.3 that $P^x\{\tau_U^{(n)} < \infty\} = 1$. Applying Lemma 3.4, we see that the function

$$\mathbf{P}^x\{X(\tau_U^{(n)}) \in \Gamma\} = u_n(x)$$

is the unique solution in $U \cap U_n$ of the problem

$$Lu = 0; \qquad u|_\Gamma = 1; \qquad u|_{|x|=n} = 0.$$

The sequence $u_n(x)$, is a monotone sequence of bounded solutions of the equation $Lu = 0$. Thus it is obvious that it converges, as $n \to \infty$, to a limit which is also a solution in U of this equation and satisfies the condition $u|_\Gamma = 1$. (This follows

---

[11]The term "regular" is used here in the sense customary in the theory of elliptic equations (see e.g. [225]).

from Harnack's second theorem for the equation $Lu = 0$; see [249].) If the Dirichlet problem in U has a unique bounded solution, then evidently

$$\lim_{n \to \infty} u_n(x) = 1. \tag{3.78}$$

On the other hand, it is obvious that

$$\{\tau_U < \infty\} \subset \bigcup_{n=1}^{\infty} \{X(\tau_U^{(n)}) \in \Gamma\},$$

whence it follows in view of (3.78) that

$$\mathbf{P}^x\{\tau_U < \infty\} = \lim_{n \to \infty} \mathbf{P}^x\{X(\tau_U^{(n)}) \in \Gamma\} = 1. \tag{3.79}$$

This proves the sufficiency.

Now assume that the process $X(t)$ is U_1-recurrent, and suppose that the Dirichlet problem in U has two bounded solutions u_1 and u_2. Then their difference $v(x) = u_2(x) - u_1(x)$ satisfies the conditions $v|_{\Gamma} = 0$ on Γ, and the condition $v(x) < k$ for some constant k. It follows from Lemma 3.4 that for every n

$$v(x) = \mathbf{E}^x v(X(\tau_U^{(n)})) \leq k(1 - \mathbf{P}^x\{X(\tau_U^{(n)}) \in \Gamma\}).$$

Letting $n \to \infty$ and using the fact that $\mathbf{P}^x\{\tau_U^{(n)} \in \Gamma\} \to \mathbf{P}^x\{\tau_U < \infty\}$ by virtue of the regularity of the process, we get $v(x) \equiv 0$. This contradiction concludes the proof. □

Remark 3.14 The regularity of the process $X(t)$ is not needed for the sufficiency part of the proof.

Theorem 3.11 *Under the assumptions of Theorem* 3.10, *a necessary and sufficient condition for the existence of the expectation of the random variable τ_U is that there exists in U a function $V(x)$, twice continuously differentiable and nonnegative, such that*

$$LV(x) = -1.$$

The function $u(x) = \mathbf{E}^x \tau_U$ is then the smallest positive solution in U of the problem

$$Lu = -1, \qquad u|_{\Gamma} = 0. \tag{3.80}$$

Proof We first assume that there exists a function $V(x)$ satisfying the conditions of the theorem. Then $\mathbf{E}^x \tau_U$ exists and by Theorem 3.9, we have $\mathbf{E}^x \tau_U \leq V(x)$. Since $\tau_U^n \uparrow \tau_U$ almost surely as $n \to \infty$, it follows that $\mathbf{E}^x \tau_U^{(n)} \to \mathbf{E}^x \tau_U$. By Lemma 3.6, the function $u_n(x) = \mathbf{E}^x \tau_U^{(n)}$ is a solution in U of the problem

$$Lu_n(x) = -1; \qquad u_n|_{\Gamma} = 0. \tag{3.81}$$

Hence the function $v_n(x) = u_{n+1}(x) - u_n(x)$ is L-harmonic (i.e., satisfies the equation $Lv_n = 0$) in the domain $U \cap \{|x| < n\}$. It is also clear that $v_n(x) > 0$. The function $u(x)$ is the sum of the series

$$u(x) = u_{n_0}(x) + \sum_{k=n_0}^{\infty} v_k(x). \tag{3.82}$$

It is known that the sum of a convergent series of positive L-harmonic functions is also an L-harmonic function. Thus (3.82) implies that the function $u(x)$ is twice continuously differentiable and satisfies the equation $Lu = -1$. Because of the assumed regularity properties of the boundary Γ, we have $u|_{\Gamma} = 0$. We claim that u is the smallest positive solution of the problem (3.80). To prove this we consider another positive solution $W(x)$ of (3.80). Since the function $u_n(x)$ satisfies the boundary conditions $u_n|_{\Gamma} = 0$; $u_n|_{|x|=n} = 0$, it follows from the maximum principle for elliptic equations (see [225]) that $u_n(x) \leq W(x)$ in $U \cap \{|x| < n\}$ for all n. Letting $n \to \infty$, we get $u(x) \leq W(x)$, as required.

The necessity part of the proof is even simpler. If $u(x_0) = \mathbf{E}^{x_0} \tau_U < \infty$ for at least one $x_0 \in U$, then, using the representation (3.82), we verify that the sequence $u_n(x_0)$ converges to $u(x_0)$. Hence it follows from Harnack's second theorem for L-harmonic functions that the series (3.82) converges uniformly on every compact set. Its limit $u(x)$ also satisfies the equation $Lu = -1$. Thus there exists a function $V(x)$ satisfying the required conditions (for example, we can set $V(x) = u(x)$). $\qquad \square$

3.8 Further Conditions for Recurrence and Finiteness of Mean Recurrence Time[12]

In this section we derive a few corollaries from the results of Sect. 3.7. We also consider some examples.

Example 3.9 Considering the auxiliary function $V(x) = \sum_{i,j} b_{ij} x_i x_j = (Bx, x)$, we infer from Theorems 3.5 and 3.9 that the process $X(t)$ is recurrent relative to the domain U if for some function $\alpha(s) \geq 0$ satisfying condition (3.72) and a nonsingular positive definite matrix B the condition

$$(Bx, b(s, x)) + \text{tr}(A(s, x)B) \leq -\alpha(s) \tag{3.83}$$

holds in $\{t > 0\} \times U$. It follows from Corollary 3.2 that the process $X(t)$ has finite mean recurrence time for U_1 if condition (3.83) holds with a function $\alpha(s) \leq -1$.

By imposing on the process an additional, not too stringent restriction, one can derive a more convenient condition for recurrence. In the following lemma the domain $U_1 \in \mathbb{R}^l$ is assumed bounded.

---

[12]For the nondegenerate case, results resembling those of this section have been established by Wonham [277]. His methods were different from ours.

Lemma 3.9 *Suppose that the process $X(t)$ almost surely exits from each bounded domain in a finite time. Then a sufficient condition for U_1-recurrence is that there exist a nonnegative function $V(t, x)$ in the domain $\{t > 0\} \times U$ such that*

$$V_R = \inf_{t > 0, |x| \geq R} V(t, x) \to \infty \quad as \ R \to \infty,$$

$$LV \leq 0.$$

Proof Let τ_U, τ_R, $\tau_{U,R}$ denote the time of first exit of the sample function of the process from the domains U, U_R and $U \cap U_R$ respectively. Denote $\min(\tau_{U,R}, t)$ by $\tau_{U,R}(t)$. For each $s \geq 0$, $x \in U$, it follows from the assumptions of the lemma and from Lemma 3.2 that

$$\mathbf{E}^{s,x} V(\tau_{U,R}(t), X(\tau_{U,R}(t))) \leq V(s, x).$$

Hence, as in the proof of Theorem 3.9, we get

$$\mathbf{E}^{s,x} V(\tau_{U,R}, X(\tau_{U,R})) \leq V(s, x).$$

This inequality and Chebyshev's inequality imply that

$$\mathbf{P}^{s,x}\{\tau_R < \tau_U\} \leq \frac{V(s, x)}{V_R} \to 0 \quad as \ R \to \infty.$$

Since moreover the process is continuous, we have $\mathbf{P}^{s,x}\{\tau_R = \tau_U\} = 0$. These relations yield the assertion of the lemma, since

$$\mathbf{P}^{s,x}\{\tau_U < \infty\} \geq \mathbf{P}^{s,x}\{\tau_U < \tau_R\} \to 1 \quad as \ R \to \infty. \qquad \square$$

Remark 3.15 The first assumption of the lemma (recurrence of the process $X(t)$ relative to the exterior of every bounded domain) holds for instance when either of the conditions (3.75) or (3.76) is satisfied in every bounded domain.

Lemma 3.10 *A sufficient condition for nonrecurrence of the process $X^{s_0, x_0}(t)$ defined by the equation*

$$X^{s_0, x_0}(t) = x_0 + \int_{s_0}^{t} b(u, X^{s_0, x_0}(u)) \, du + \int_{s_0}^{t} \sum_{r=1}^{k} \sigma_r(u, X^{s_0, x_0}(u)) \, d\xi_r(u), \quad (3.84)$$

relative to the domain $U_1 = U^c$ with boundary Γ is that there exists in $\{t > 0\} \times U$ a function $V(t, x)$ such that

$$\sup_{x \in \Gamma, t \geq s_0} V(t, x) \leq 0; \qquad LV(t, x) \geq 0; \qquad \sup_{(t > 0) \times U} V(t, x) < k < \infty,$$

$$(3.85)$$

$$V(s_0, x_0) > 0.$$

Proof It follows easily from $LV \geq 0$ and from Lemma 3.2 that

$$\mathbf{E}^{s_0, x_0} V(\tau_U^{(n)}(t), X(\tau_U^{(n)}(t))) \geq V(s_0, x_0),$$

where $\tau_U^{(n)}(t) = \min(\tau_n, \tau_U, t)$. Hence in view of (3.85), we have

$$k\mathbf{P}^{s_0, x_0}\{\tau_U > \min(\tau_n, t)\} \geq V(s_0, x_0). \tag{3.86}$$

It will now be convenient to distinguish two cases: (a) the process $X^{s_0, x_0}(t)$ is regular in U, so that $\lim_{n \to \infty} \tau_n = \tau = \infty$ almost surely; (b) the process is nonregular, i.e., $\tau < \infty$ with positive probability. In case (a) we let $n \to \infty$ in (3.86) and we conclude that $\mathbf{P}^{s_0, x_0}\{\tau_U = \infty\} > 0$. In case (b) the process is nonrecurrent, since it is nonregular. This proves the lemma. $\qquad \square$

Lemmas 3.9 and 3.10 easily yield necessary and sufficient conditions for recurrence in the one-dimensional case.

Example 3.10 Let the process $X(t)$ be defined by the equation

$$dX(t) = b(X(t)) \, dt + \sigma(X(t)) \, d\xi(t),$$

where b and σ are continuously differentiable in $\mathbb{R}^1$ and $\sigma(x) \neq 0$. The generator for this process is

$$L = b(x) \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2(x) \frac{\partial^2}{\partial x^2}.$$

Consider the functions

$$Q(x) = \exp\left\{-2 \int_0^x \frac{b(z)}{\sigma^2(z)} dz\right\}; \qquad W(x) = \int_0^x Q(y) \, dy.$$

It is easily seen that $LW = 0$. If moreover $W(x) \to \pm\infty$ as $x \to \pm\infty$, then the function $W(x) \operatorname{sign} x$ satisfies the assumptions of Lemma 3.9, and therefore the process $X(t)$ is recurrent relative to every segment containing the origin. But if the function $W(x)$ is bounded in the domain $x > 0$ (or $x < 0$), then the function $W(x) - \varepsilon$ (or $-W(x) - \varepsilon$) satisfies the assumptions of Lemma 3.10.

Thus, the condition

$$\int_0^x \exp\left\{-2 \int_0^y \frac{b(z)}{\sigma^2(z)} dz\right\} dy \to \pm\infty \quad \text{as } x \to \pm\infty$$

is necessary and sufficient for a time-homogeneous one-dimensional process to be recurrent.

Note also that for the recurrent process $X(t)$ the transformation $W(x)$ generates the so-called canonic coordinate $y = W(x)$ (see [64]). In this coordinate the process $Y(t) = W(X(t))$ has the generator

$$\widetilde{L}(y) = \frac{1}{2} \sigma^2(x(y)) Q^2(x(y)) \frac{d^2}{dy^2}.$$

So the study of time-homogeneous one-dimensional recurrent diffusion processes can be reduced to the study diffusion process without drift term.

Lemma 3.11 *Let U_1 be a bounded domain with boundary Γ. Assume that there exist in the domain $\{t > 0\} \times U$ functions $V(t, x)$ and $W(t, x)$ such that*

(1) $V > 0$; $W|_\Gamma \le 0$; $LV \le 1$, $LW \ge 0$;
(2) *for some increasing system of bounded domains $U_n \supset U_1$ with boundaries Γ_n,*

$$\frac{\inf_{\Gamma_n} V}{\sup_{\Gamma_n} W} = A_n \to \infty \quad as\ n \to \infty;$$

(3) *the process $X(t)$, defined in the domain $U \cap U_n$ up to the time $\tau^{(n)}$ at which it first reaches the boundary of this domain, satisfies the condition*

$$\mathbf{P}^{s,x}\{\tau^{(n)} < \infty\} = 1$$

for all n. Then $\mathbf{E}^{s,x}\tau_U = \infty$ for all points s, x such that $W(s, x) > 0$.

Proof By assumption, the function $V - A_n W$ satisfies in $U \cap U_n$ the conditions

$$L(V - A_n W) \le 1; \qquad (V - A_n W)|_{\Gamma_n} \ge 0; \qquad (V - A_n W)|_\Gamma \ge 0.$$

This together with Lemma 3.2 and $X(\tau^{(n)}) \in \Gamma \cup \Gamma_n$, implies that

$$A_n W(s, x) - V(s, x) \le \mathbf{E}^{s,x} V(\tau^{(n)}, X(\tau^{(n)})) - A_n W(\tau^{(n)}, X(\tau^{(n)}))$$

$$+ A_n W(s, x) - V(s, x)$$

$$\le \mathbf{E}^{s,x}(\tau^{(n)} - s).$$

Since $A_n \to \infty$ as $n \to \infty$, we see from the above that

$$\lim_{n \to \infty} \mathbf{E}^{s,x}\tau^{(n)} = \infty$$

for all points s, x such that $W(s, x) > 0$. The assertion of the lemma is now obvious if we observe that $\tau^{(n)} < \tau_U$ almost surely. $\qquad \square$

Remark 3.16 It is not hard to find conditions, in terms of auxiliary functions implying the validity of assumption (3) of Lemma 3.11 (see Sect. 3.7). In particular, this assumption holds if (3.32) and the nondegeneracy condition (3.60) are valid in every domain bounded with respect to x.

Example 3.11 We conclude from Lemma 3.9 that a process $X(t)$ in $\mathbb{R}^l$ satisfying the nondegeneracy condition (3.60) is recurrent relative to every bounded domain U_1 containing the origin when there exists a positive definite symmetric matrix B such that,

$$(Bx, b(t, x)) + \text{tr}(A(t, x)B) \le \frac{2(BA(t, x)Bx, x)}{(Bx, x)} \tag{3.87}$$

holds in $U = U_1^c$. The condition for recurrence given by inequality (3.87) is more general than that given by inequality (3.83), since the right-hand side of (3.87) is always positive. The sufficiency of condition (3.87) can be proved using the auxiliary function $V(x) = \ln(Bx, x) + k$.

We now consider some particular consequences of (3.87). First let $B = A = J$, where J is the $l \times l$ identity matrix. Then (3.87) becomes $(x, b(t, x)) \leq 2 - l$. Hence, in particular, if $l = 1$ or $l = 2$, then the process $X(t)$ associated with the operator

$$L = \frac{\partial}{\partial t} + \sum_{i=1}^{l} b(t, x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{l} \frac{\partial^2}{\partial x_i^2}, \tag{3.88}$$

is recurrent, provided $(x, b(t, x)) \leq 0$ outside some bounded domain. For example, the Wiener process on the line and the Wiener process in the plane ($b \equiv 0$) are recurrent.

Lemma 3.10 yields various sufficient conditions for recurrence relative to U_1 (and hence also relative to every domain contained in U_1).

Example 3.12 Consider the auxiliary function

$$V(x) = 1 - k(Bx, x)^{-\alpha}.$$

It is readily seen that

$$LV(x) = 2k\alpha(Bx, x)^{-\alpha-1} \left[(Bx, b(t, x)) + \text{tr}(A(t, x)B) \right.$$
$$\left. - \frac{2(1 + \alpha)(BA(t, x)Bx, x)}{(Bx, x)} \right].$$

If for some α

$$(Bx, b(t, x)) + \text{tr}(A(t, x)B) \geq \frac{2(1 + \alpha)(BA(t, x)Bx, x)}{(Bx, x)}, \tag{3.89}$$

then for a suitable choice of k and U_1 the function V satisfies the assumptions of Lemma 3.10. Hence we may conclude that the process $X(t)$ is nonrecurrent if there exist a positive definite *symmetric* matrix B and a constant $\alpha > 0$ for which inequality (3.89) holds outside some bounded domain. Setting $B = A = J$, we infer from (3.86) that if

$$(Bx, b(x)) \geq 2 - l + \varepsilon$$

in a neighborhood of the point at infinity, then the process described by the operator (3.88) is nonrecurrent. In particular, it follows that the Wiener process is nonrecurrent in every space of dimension higher than 2.

It is readily seen that the condition that B be positive definite can be considerably weakened; it suffices to assume that B is positive semi-definite and that at least one

of its eigenvalues is nonzero. In this case the neighborhood of the point at infinity must be replaced by a (connected) component of the set $(Bx, x) > y > 0$.

For example, let B be a matrix all of whose elements except b_{ii} are zeros, and $b_{ii} = 1$. It then follows from (3.89) that the process $X(t)$ is nonrecurrent if for some i and $\varepsilon > 0$ the condition

$$x_i b_i(t, x) \geq (1 + \varepsilon) a_{ii}(t, x)$$

holds in the domain $x_i > \gamma$ or $x_i < -\gamma$ (where γ is sufficiently large).

Example 3.13 We set $V = (B_1 x, x)$, $W = (B_2 x, x)^\alpha - k$, where B_1 and B_2 are positive definite symmetric matrices, $0 < \alpha < 1$, and the constant k is sufficiently large. Obviously,

$$LV = 2(B_1 x, b(t, x)) + 2 \operatorname{tr}(A(t, x) B_1),$$

$$LW = 2\alpha V^{\alpha-1} \left[(B_2 x, b(t, x)) + \operatorname{tr}(A(t, x) B_2) - 2(1 - \alpha) \frac{(B_2 A(t, x) B_2 x, x)}{(B_2 x, x)} \right].$$

Applying Lemma 3.11, we see that the mean recurrence time in U_1 is infinite if there are positive definite matrices B_1, B_2 and a number $\varepsilon > 0$ such that in U

$$(B_1 x, b(t, x)) + \operatorname{tr}(A(t, x) B_1) < \text{const},$$

$$(B_2 x, b(t, x)) + \operatorname{tr}(A(t, x) B_2) > \varepsilon \frac{(B_2 A(t, x) B_2 x, x)}{(B_2 x, x)}.$$

We again consider the process associated with the operator (3.88). Setting $B_1 = B_2 = J$, we see that the mean recurrence time in a bounded domain for this process does not exist, if for some $\varepsilon > 0$ the scalar product $(x, b(x))$ satisfies

$$-l + \varepsilon < (x, b(x)) < \text{const}$$

for $|x| > R$. In particular, these inequalities are valid for the Wiener process in $\mathbb{R}^l$ for every l.

Chapter 4
Ergodic Properties of Solutions of Stochastic Equations

The main results of this chapter are an extended and improved version of Sects. 4.2 and 4.3 of the author's paper [117]. Under different assumptions, results similar to those given below in Sects. 4.4 and 4.5 have been derived by Maruyama and Tanaka [194]. The existence proof given in Sects. 4.6 and 4.7 for the limit of the transition probability function is similar to that of Theorem 3 in [100]. For the discrete time case, the existence of the limit of the transition probability function has been proved under similar assumptions by Nagaev [213].

4.1 Kolmogorov Classification of Markov Chains with Countably Many States

In the preceding chapters we found some sufficient conditions for the existence of a stationary Markov process defined by a SDE. The following two questions are also of great interest: (a) When is the stationary Markov process associated with a given stochastic equation unique? (b) When can it be said that a Markov process with arbitrary initial distribution from a given class converges in some sense to a stationary one?

In this chapter we shall consider these and some related questions in terms of the properties of the coefficients of the operator L. Using the results of Chap. 3, one can reformulate the results in terms of Lyapunov functions.

The material in this chapter is organized similarly as in the well-known paper of Kolmogorov [145] on Markov chains with countably many states E_i and discrete time. In that paper Kolmogorov divides the classes of communicating essential[1] states of the chain into the two types of recurrent and nonrecurrent (or transient) ones.

A class of states E_i is said to be if the probability L_{ii} of reaching E_i at least once from E_i is equal to 1. (It can be proved that in a single class either $L_{ii} < 1$ for

---

[1] A state E_i is said to be *unessential* if the transition probability from E_i to some other state E_j is positive, but the reverse transition from E_j to E_i is impossible.

all i or all L_{ij} are equal to 1.) A nonrecurrent class is said to be *transient*. It is easy to prove (see below in Sect. 4.2 the analog of this result for continuous time) that the n-step transition probability $p_{ij}^{(n)}$ from E_i to E_j for a transient class tends to zero as $n \to \infty$. Hence it follows that no stationary distribution exists for a transient class of states.

Kolmogorov also divides recurrent classes into two types. A class of states is said to be *positive recurrent* if the mean recurrence time M_{ii} for a state E_i, starting from E_i, is finite. (It can be proved that in a single class either $M_{ii} = \infty$ for all i or all M_{ij} are finite.) Otherwise the class is said to be a *null recurrent class*. It can be shown that in a null recurrent class $p_{ij}^{(n)} \to 0$ as $n \to \infty$. For a positive recurrent class it can be shown under a certain additional assumption (that the class consists of a single subclass) that

$$\lim_{n \to \infty} p_{ij}^{(n)} = \pi_j \neq 0.$$

Thus within a positive recurrent class of states consisting of a single subclass a limit distribution π_j will ultimately be established, irrespective of the initial distribution. It is easily shown that this limit distribution is stationary.

Kolmogorov has proved that a necessary and sufficient condition for the existence of a stationary distribution within a given class A is that the expectation of referring time from i to i $M_{ii} < \infty$ holds for at least one state $E_i \in A$.

In this chapter we shall derive sufficient conditions for the existence of a stationary distribution and we shall prove theorems on the limit behavior of the transition probability for continuous Markov processes described by SDEs in terms of functions analogous to L_{ij} and M_{ij}. These results will be derived in terms of functions analogous to L_{ij} and M_{ij}. The only significant modification is that recurrence for a single state must be replaced by recurrence for a compact set.

As before, we shall use the symbols $\mathbf{E}^x$ and $\mathbf{P}^x$ to designate the expectation of a random variable and the probability of an event when these are determined by the evolution of a time-homogeneous Markov process $X^{0,x}(t)$. In this and the following chapters we shall make constant use of the strong Markov property, introduced by Dynkin and Yushkevich [65]. In Chap. 3 we considered the random variable τ_U equal to the first exit time of the sample path from the domain U, $\tau_U(t) = \min(\tau_U, t)$, and other variables, all possessing the property that the event "the random variable assumes a value smaller than s" depends only on the evolution of the process up to time s, i.e.

$$\{\tau \leq s\} \in \mathcal{N}_s. \tag{4.1}$$

A random variable τ satisfying condition (4.1) will be called a *Markov time* or a *random variable independent of the future*. Roughly speaking, a Markov process is said to be *strongly Markov* if the future is independent of the past not only for a fixed instant of time but also for any Markov time τ.

A Markov process such that the transition probability function $P(s, x, t, A)$ is Borel-measurable as a function of (s, x) is said to be *strongly Markov* if for any

Markov time τ, $\mathcal{N}_\tau$-measurable random variable $\eta \geq \tau$, $x \in \mathbb{R}^l$ and $A \in \mathfrak{B}$,

$$\mathbf{P}\{X^{s,x}(\eta) \in A \mid \mathcal{N}_\tau\} = P(\tau, X^{s,x}(\tau), \eta, A) \quad \text{(a.s.)}. \tag{4.2}$$

It can be proved that any right-continuous process with Feller transition probability function is strongly Markov [65]. Therefore, the solution of an Itô equation (3.31) is also a strong Markov process. This fact will be repeatedly used in the sequel.

4.2 Recurrence and Transience

In Sects. 3.7 and 3.8 we studied conditions which imply the recurrence solutions of stochastic equations relative to a domain U, i.e., conditions under which the paths issuing from any point $x \in \mathbb{R}^l \setminus U$ almost surely reach the set U. In this section we shall show that for time-homogeneous processes with nonsingular diffusion matrix the recurrence property does not depend on the choice of the bounded open domain U.

Let $X(t)$ be a regular time-homogeneous Markov process in $\mathbb{R}^l$ described by the SDE

$$dX(t) = b(X) \, dt + \sum_{r=1}^{k} \sigma_r(X) \, d\xi_r(t). \tag{4.3}$$

Here and in the next section we shall assume that the diffusion matrix

$$A(x) = ((a_{ij}(x))), \quad a_{ij}(x) = \sum_{r=1}^{k} \sigma_r^i(x) \sigma_r^j(x)$$

of the process $X(t)$ is nonsingular, i.e., the smallest eigenvalue of the matrix $A(x)$ is bounded away from zero in every bounded domain.

Lemma 4.1 *If $X(t)$ is recurrent relative to some bounded open domain U, then it is recurrent relative to any nonempty open domain in $\mathbb{R}^l$.*

Proof It will suffice to prove that the process is recurrent relative to a domain U_0 with regular boundary Γ_0 such that $U_0 \cup \Gamma_0 \subset U$. We may also assume that the boundary Γ of U is regular (otherwise we replace U by a suitable domain containing U).

Let U_1 be a domain with regular boundary Γ_1 such that $U \cup \Gamma \subset U_1$. Let $x \in U_0^c$ be arbitrary. We claim the $\mathbf{P}^x\{\tau_{U_0^c} < \infty\} = 1$.

Assuming for definiteness that $x \in U_1^c$, we consider the following random variables: τ_1', the time at which the path of the process first reaches the set Γ, and τ_1, the first time after τ_1' at which the path reaches Γ_1. We now define two random variables inductively: τ_n', the first time after τ_{n-1} at which the path reaches Γ, and τ_n, the first time after τ_n' at which it reaches Γ_1 (Fig. 4.1). By the assumption that the process is

Fig. 4.1
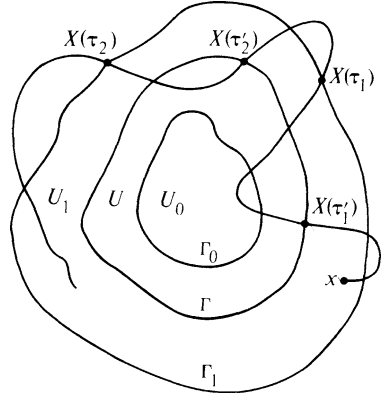
recurrent relative to U and by Corollary 3.3, it follows that $\tau_n < \infty$ almost surely for all $n = 2, 3, \ldots$.

Set $U_3 = U_1 \setminus U_0$. By Lemma 3.4, the function

$$u(x) = \mathbf{P}^x \{ X(\tau_{U_0}) \in \Gamma_0 \}$$

is a solution of the elliptic equation

$$Lu = 0$$

in the domain U_3 satisfying the boundary conditions

$$u|_{\Gamma_0} = 1; \qquad u|_{\Gamma_1} = 0.$$

By the strong maximum principle for solutions of elliptic equations,

$$\max_{x \in \Gamma} \mathbf{P}^x \{ X(\tau_{U_3}) \in \Gamma_1 \} = p_0 < 1. \tag{4.4}$$

Using the fact that the random variables $\tau_1', \tau_1, \tau_2', \tau_2, \ldots$ are Markov times and applying the strong Markov property of the process $X(t)$, and also (4.4), we get

$$\mathbf{P}^x \{ \tau_{U_0^c} = \infty \}$$

$$\leq \sup_{z \in \Gamma_1} \mathbf{P}^z \left\{ \bigcap_{i=1}^{\infty} (X(\tau_{U_3}^i) \in \Gamma_1) \right\}$$

$$= \sup_{z \in \Gamma_1} \int_{\Gamma} \mathbf{P}^z \{ X(\tau_1') \in dy \} \int_{\Gamma_1} \mathbf{P}^y \{ X(\tau_{U_3}) \in dz_1 \} \mathbf{P}^{z_1} \left\{ \bigcap_{i=1}^{\infty} (X(\tau_{U_3}^i) \in \Gamma_1) \right\}$$

$$\leq p_0 \sup_{z \in \Gamma_1} \mathbf{P}^z \left\{ \bigcap_{i=1}^{\infty} (X(\tau_{U_3}^i) \in \Gamma_1) \right\},$$

where $\tau_{U_3}^i$ denotes the first time after τ_i' at which the path exits from the set U_3.

These inequalities imply that $\mathbf{P}^x\{\tau_{U_0^c} = \infty\} = 0$, since $p_0 < 1$. This completes the proof of the lemma. $\qquad\square$

In view of this lemma the following definitions are natural. A regular process $X(t)$ described by (4.3) with nonsingular diffusion matrix is said to be *recurrent* if there exists a bounded domain U such that for all $x \in U^c$,

$$\mathbf{P}^x\{\tau_{U^c} < \infty\} = 1.$$

If there exist a non-empty open domain U and a point $x \in U^c$ such that $\mathbf{P}^x\{\tau_U < \infty\} < 1$, the process is said to be *transient*.

It follows from the definition of recurrence and from Lemma 4.1 that the sample path of a recurrent process is almost surely dense in $\mathbb{R}^l$ and it prevails an infinite time in every fixed neighborhood of any point. The situation is different for transient processes.

Lemma 4.2 *If the process $X(t)$ is transient, then for any compact set K and any $x \in \mathbb{R}^l$ the probability of the event "the sample path of the process $X^x(t)$ never passes through the set K after some random but finite instant of time $t_0(\omega)$" equals 1.*

Proof Let K be a compact set in $\mathbb{R}^l$, and let $U_R \supset K$ ball of radius R containing K. Since the process is transient, the value of the function

$$u(x) = \mathbf{P}^x\{\tau_{U_R^c} < \infty\}$$

is less than 1 at some point $x \in U_R^c$. This function satisfies the equation $Lu = 0$ (as the limit of a monotone sequence of L-harmonic functions; see Sect. 3.7). By the strong maximum principle,

$$\max_{|x|=R+1} \mathbf{P}^x\{\tau_{U_R^c} < \infty\} = q < 1. \tag{4.5}$$

Let $x \in U_R^c$. We now define the random variables $\tau_1', \tau_1, \tau_2', \ldots$ as follows: $\tau_1' = \tau_{U_{R+1}}$, τ_1 is the first time after τ_1' at which the set $|x| = R$ is reached, τ_i' is the first time after τ_1 at which the set $|x| = R+1$ is reached, and so on. For example, let us estimate the probability $\mathbf{P}^x\{\tau_2 < \infty\}$. Using the strong Markov property of the process and also (4.5), we get

$$\mathbf{P}^x\{\tau_2 < \infty\} = \int_{|y_1|=R+1} \mathbf{P}^x\{X(\tau_1') \in dy_1\} \int_{|z|=R} \mathbf{P}^{y_1}\{\tau_{U_R} < \infty; X(\tau_{U_R^c}) \in dz\}$$

$$\times \int_{|y_2|=R+1} \mathbf{P}^z\{X(\tau_2') \in dy_2\} \mathbf{P}^{y_2}\{\tau_{U_R^c} < \infty\}$$

$$\leq q^2.$$

Similarly, one shows that $\mathbf{P}^x\{\tau_n < \infty\} \leq q^n$.

Thus $\{\tau_n < \infty\}$ is a sequence of events whose probabilities decrease in geometric progression. By the Borel–Cantelli lemma, this implies that almost surely only finitely many of these events will occur. This proves the lemma. $\square$

Corollary 4.1 *Since* $\{X(t) \in K\} \subset \{t_0(\omega) > t\}$, *it follows that a transient process* $X(t)$ *satisfies*

$$\mathbf{P}^x\{X(t) \in K\} = P(x, t, K) \to 0 \quad \text{as } t \to \infty \tag{4.6}$$

for every compact set $K \subset \mathbb{R}^l$.

Lemma 4.3 *If the process* $X(t)$ *is transient, then the random variable* ζ_K *equal to the total time which the sample path of the process spends in the compact set* K *satisfies*

$$\mathbf{E}^x |\zeta_K|^n < c_n < \infty \tag{4.7}$$

for any $n < 0$.

Proof It is sufficient to prove this when K is the ball $|x| \leq R$. Let $q < 1$ be the number defined by (4.5). The random variable τ_{R+1} defined as the time to the first exit of the path of $X(t)$ from the set $\{|x| < R + 1\}$ has bounded expectation, due to Corollary 3.3. By Chebyshev's inequality, there is a constant $t > 0$ such that

$$\mathbf{P}^x\{\tau_{R+1} \geq t\} < \frac{1-q}{2}. \tag{4.8}$$

Since for any $x \in K$

$$\mathbf{P}^x\{\zeta_K \geq t\} \leq \mathbf{P}^x\{\tau_{R+1} \geq t\} + \sup_{|y|=R+1} \mathbf{P}^y\{\tau_{U_R^c} < \infty\}, \tag{4.9}$$

it follows from (4.5) and (4.8) that

$$\mathbf{P}^x\{\zeta_K \geq t\} \leq \frac{1+q}{2} = q_1 < 1$$

for $x \in K$. It is clear that (4.9) is valid for all $x \in \mathbb{R}^l$. A necessary condition for the occurrence of the event $\{\zeta_K \geq 2t\}$ is obviously that of the event $\{\zeta_K \geq t\}$ occurs and that beginning from the time at which ζ_K is first equal to t, the path remains in K for a time not less than t. Hence (the rigorous argument involves the strong Markov property)

$$\mathbf{P}^x\{\zeta_K \geq 2t\} \leq q_1^2, \qquad \ldots, \qquad \mathbf{P}^x\{\zeta_K \geq nt\} \leq q_1^n. \tag{4.10}$$

Inequalities (4.10) guarantee the existence of bounded moments for all powers of the random variable ζ_K, since

$$\mathbf{E}^x \zeta_K^n \leq \sum_{m=0}^{\infty} \mathbf{P}^x \{\zeta_K > mt\}[(m+1)t]^n$$

for $x \in K$, and hence also for all $x \in \mathbb{R}^l$.   □

Corollary 4.2 *By virtue of the equality*

$$\mathbf{E}^x \zeta_K = \mathbf{E}^x \int_0^{\infty} \mathbb{1}_K(X(t))\,dt = \int_0^{\infty} P(x, t, K)\,dt$$

it follows from Lemma 4.3 *that the function* $\int_0^{\infty} P(x, t, K)\,dt$ *is bounded in* $\mathbb{R}^l$ *for a transient process* $X(t)$.

Remark 4.1 It is readily seen that the proofs of the lemmas in this section essentially use far weaker properties of the process $X(t)$ than the nonsingularity of the diffusion matrix. For example, the following generalization of Lemma 4.1 can be proved with help of small modification of reasonings, used for the proof of this lemma.

Lemma 4.4 *The solution* $X(t)$ *of the stochastic equation* (4.3) *is recurrent relative to the domain* U_2 *when it is recurrent relative to* U_1 *and*

$$\inf_{x \in U_1} P(x, T, U_2) > 0 \tag{4.11}$$

for some $T > 0$.

It is easy to see that condition (4.11) is satisfied if the domain U_1 is bounded and the transition probability of $X(t)$ has an everywhere positive density.

4.3 Positive and Null Recurrent Processes

Suppose that the conditions formulated at the beginning of Sect. 4.2 are satisfied, and let the process $X(t)$ be recurrent. As in the case of a process with countably many states, the asymptotic behavior of the transition probability depends essentially on whether the mean recurrence time for a bounded domain is finite or infinite.

Lemma 4.5 *Under the above assumptions, if* $\mathbf{E}^{x_0} \tau_{U^c}$ *is finite for some bounded open domain* U *and* $x_0 \in U^c$, *then* $\mathbf{E}^x \tau_{U_0^c}$ *is finite for all nonempty open domains* U_0 *and all* $x \in U_0^c$.

Proof Let $\mathbf{E}^{x_0} \tau_{U^c} < \infty$ for some bounded open domain U and $x_0 \in U^c$. It was shown in the proof of Theorem 3.11 that then $\mathbf{E}^x \tau_{U^c} < \infty$ for all $x \in U^c$.

We must prove that $\mathbf{E}^x \tau_{U_0^c} < \infty$ for any other non-empty open domain U_0. As in the proof of Lemma 4.1, we need only deal with the case in which U and U_0 are open domains with regular boundaries Γ and Γ_0 respectively and $U_0 \cup \Gamma_0 \subset U$.

As in the proof of Lemma 4.1, we construct an auxiliary domain U_1 and consider the corresponding times $\tau_1', \tau_1, \ldots, \tau_n', \tau_n, \ldots$ (see Fig. 4.1). We shall call the portion of the sample path of the process from τ_{n-l} to τ_n the n-th cycle, and we set $\tau_0 = 0$. The event A: "the path of $X(t)$ reaches the set U_0" can first occur either during the first cycle, or during the second, etc. If it occurs during the n-th cycle, then obviously,

$$\tau_{U_0^c} < \tau_n.$$

On the other hand, it follows by (4.4) and the strong Markov property that the probability that A will not occur during the $n-1$ preceding cycles is majorized by p_0^{n-1}. Hence we have the estimate

$$\mathbf{E}^x \tau_{U_0^c} < \mathbf{E}^x \tau_1 + p_0 \mathbf{E}^x \tau_2 + \cdots + p_0^{N-1} \mathbf{E}^x \tau_n + \cdots.$$

Without loss of generality, we may assume that $x \in \Gamma_1$. Obviously,

$$\sup_{x \in \Gamma_1} \mathbf{E}^x \tau_1 \leq \sup_{x \in \Gamma_1} \mathbf{E}^x \tau_{U^c} + \sup_{x \in \Gamma} \mathbf{E}^x \tau_{U_1} = B < \infty.$$

Therefore

$$\mathbf{E}^x \tau_{U_0^c} \leq B + 2Bp_0 + \cdots + nBp_0^{n-1} + \cdots < \infty$$

as required. $\square$

In accordance with Kolmogorov's terminology for Markov chains with countably many states, a recurrent process such that the mean recurrence time for some (hence for each) bounded open set is finite will be called a *positive recurrent process*. 0therwise we have a *null recurrent process*.

Sufficient conditions for a process to be positive recurrent, null recurrent, or transient, in terms of the coefficients b and σ_r, were given above (Sects. 3.7, 3.8).

4.4 Existence of a Stationary Distribution

In Sects. 4.4 through 4.7 we shall study ergodic properties of positive recurrent Markov processes. First, in Sect. 4.4, we shall establish the existence of a stationary distribution for such processes. This will enable us to apply the ergodic theorem for stationary processes and thus, in Sect. 4.5, to establish the strong law of large numbers for functions of diffusion-type Markov processes.

Next, in Sects. 4.6 and 4.7, we shall prove a theorem which states that, under certain restrictions, the transition probability of a time-homogeneous process from a point x to a set A in time t tends to a limit as $t \to \infty$. This limit is independent of the "point of departure" x. It equals the stationary distribution $\mu(A)$. In Sect. 4.8 we discuss certain generalizations of previous results. In Sect. 4.9 we shall prove

corollaries concerning the behavior of the solutions of parabolic differential equations for large time values. Sections 4.10 and 4.11 are devoted to a more detailed consideration of a null recurrent process on the real line.

The main assumptions adopted in Sects. 4.4 through 4.7 may be described as follows. We shall stipulate that the process $X(t)$ has finite mean recurrence time for some bounded open domain U, and within this domain all sample paths "mix sufficiently well" (while outside U the diffusion matrix may be as strongly singular as desired; for example, the process may be deterministic outside U).

More precisely, we shall adopt the following

Assumption (B) *There exists a bounded open domain* $U \subset \mathbb{R}^l$ *with regular boundary* Γ, *having the following properties*:

(B.1) *In the domain U and some neighborhood thereof, the smallest eigenvalue of the diffusion matrix $A(x)$ is bounded away from zero.*

(B.2) *If $x \in \mathbb{R}^l \setminus U$, the mean time τ at which a path issuing from x reaches the set U is finite, and $\sup_{x \in K} \mathbf{E}^x \tau < \infty$ for every compact subset $K \subset \mathbb{R}^l$.*

We consider a domain U_1 with sufficiently smooth boundary such that condition (B.1) holds in U_1 and $U \cup \Gamma \subset U_1$.

The construction used above in the proof of Lemma 4.1 is basic for what follows. This construction divides an arbitrary sample path of the process into "cycles":

$$[\tau_0, \tau_1); \quad [\tau_1, \tau_2); \quad \ldots; \quad [\tau_n, \tau_{n+1}); \quad \ldots.$$

Here $\tau_0 = 0$, and the times $\tau_1', \tau_1, \tau_2', \tau_2, \ldots$ are defined inductively: τ_{n+1}' is the first time after τ_n at which the set Γ is reached, and τ_{n+1} is the first time after τ_{n+1}' at which the path reaches Γ_1 (see Fig. 4.1 on p. 102).

The process $X(t)$ is U-recurrent by condition (B.2) and U_1-recurrent by condition (B.1) and Corollary 3.3. Hence, all the random variables $\tau_1' < \tau_1 < \cdots < \tau_n' < \tau_n < \cdots$ are almost surely finite.

Suppose that $X(0) = x \in \Gamma_1$, and consider the sequence $X(\tau_i) = \widetilde{X}_i$. It follows from the strong Markov property of $X(t)$ that this sequence is a Markov chain on Γ_1. Let $\widetilde{P}(x, \gamma)$ denote the one-step transition probability of this chain, and set

$$\mathbf{E}^x f(\widetilde{X}_1) = \int_{\Gamma_1} \widetilde{P}(x, dy) f(y).$$

Let us first establish some properties of the process $\widetilde{X}_n$. We denote by $\widetilde{P}^{(n)}(x, \gamma)$ the n-step transition probability of this process.

Lemma 4.6 *The Markov chain $\widetilde{X}_1, \ldots, \widetilde{X}_n, \ldots$ has a unique stationary distribution $\mu(\gamma)$, which satisfies*

$$|\widetilde{P}^{(n)}(x, \gamma) - \mu(\gamma)| < k^n \tag{4.12}$$

uniformly in $\gamma \in \Gamma_1$ for some constant $k < 1$.

Proof By Remark 3.10, for all x the measure $\mathbf{P}^x(X(\tau_{U_1}) \in \gamma)$ has a density $\sigma_{\Gamma_1}(\gamma)$ on Γ_1, relative to surface area, which is bounded away from zero. Hence the obvious equality

$$\widetilde{P}(x, \gamma) = \int \mathbf{P}^x\{X(\tau_1') \in dz\}\mathbf{P}^z\{X(\tau_{U_1}) \in \gamma\}$$

implies that $\widetilde{P}(x, \gamma)$ has the same property. Now it is well-known (see [56, Sect. 5.5]) that this condition is sufficient for the existence of a unique stationary distribution $\tilde{\mu}(\gamma)$ of the Markov chain $\widetilde{X}_n$ and for (4.12) to hold.   □

Lemma 4.7 *Let τ be a Markov time, $\mathbf{E}^x\tau < \infty$ and $f(x)$ a Borel-measurable function. Then*

$$\mathbf{E}^x \int_0^\tau f(X(t+s))\,ds = \mathbf{E}^x \int_0^\tau \mathbf{E}^{X(s)} f(X(t))\,ds. \tag{4.13}$$

Proof Since τ is a Markov time, the indicator $\mathbb{1}_{s<\tau}(s)$ of the set $\{s < \tau\}$ is $\widetilde{\mathcal{N}}_s$-measurable. Therefore

$$\mathbf{E}^x \int_0^\tau f(X(t+s))\,ds = \int_0^\infty \mathbf{E}^x\mathbf{E}^x\{\mathbb{1}_{\{s<\tau\}}(s)f(X(t+s))|\widetilde{\mathcal{N}}_s\}\,ds$$

$$= \mathbf{E}^x \int_0^\infty \mathbb{1}_{\{s<\tau\}}(s)\mathbf{E}^x\{f(X(t+s))|\widetilde{\mathcal{N}}_s\}\,ds$$

$$= \mathbf{E}^x \int_0^\tau \mathbf{E}^x\{f(X(t+s))|\widetilde{\mathcal{N}}_s\}\,ds.$$

This, together with (3.7), (3.8), implies the assertion.   □

Theorem 4.1 *If (B) holds, then the Markov process $X(t)$ has a stationary distribution $\mu(A)$.*

Proof Let $A \in \mathfrak{B}$. Let τ^A denote the time spent by the path of $X(t)$ in the set A during the first cycle. We define a measure $\nu(A)$ by

$$\nu(A) = \int_{\Gamma_1} \tilde{\mu}(dx)\mathbf{E}^x\tau^A. \tag{4.14}$$

Then, for any continuous bounded function $f(x)$,

$$\int_{\Gamma_1} f(X)\,\nu(dx) = \int_{\Gamma_1} \tilde{\mu}(dx)\mathbf{E}^x \int_0^{\tau_1} f(X(t))\,dt. \tag{4.15}$$

Recall that $\tilde{\mu}(\gamma)$ is the stationary distribution of the Markov chain $\widetilde{X}_n$, i.e., for any bounded Borel-measurable function $g(x)$ on Γ_1 we have

$$\int_{\Gamma_1} \mathbf{E}^x g(\widetilde{X}_1)\tilde{\mu}(dx) = \int_{\Gamma_1} g(x)\tilde{\mu}(dx). \tag{4.16}$$

From (4.15) and (4.13) we see by performing the substitution $t + s = u$ that

$$\int_{\mathbb{R}^l} \mathbf{E}^x f(X(t)) \nu(dx)$$

$$= \int_{\Gamma_1} \tilde{\mu}(dx) \mathbf{E}^x \int_0^{\tau_1} \mathbf{E}^{X(s)} f(X(t)) \, ds$$

$$= \int_{\Gamma_1} \tilde{\mu}(dx) \mathbf{E}^x \int_0^{\tau_1} f(X(t+s)) \, ds$$

$$= \int_{\Gamma_1} \tilde{\mu}(dx) \mathbf{E}^x \int_t^{t+\tau_1} f(X(u)) \, du$$

$$= \int_{\Gamma_1} \tilde{\mu}(dx) \mathbf{E}^x \int_0^{\tau_1} f(X(u)) \, du + \int_{\Gamma_1} \tilde{\mu}(dx) \mathbf{E}^x \int_{\tau_1}^{\tau_1 + t} f(X(u)) \, du$$

$$- \int_{\Gamma_1} \mu(dx) \mathbf{E}^x \int_0^t f(X(u)) \, du. \tag{4.17}$$

It follows from (4.16) that

$$\int_{\Gamma_1} \tilde{\mu}(dx) \mathbf{E}^x \int_{\tau_1}^{\tau_1 + t} f(X(u)) \, du = \int_{\Gamma_1} \tilde{\mu}(dx) \mathbf{E}^x \mathbf{E}^{\tilde{X}_1} \int_0^t f(X(u)) \, du$$

$$= \int_{\Gamma_1} \tilde{\mu}(dx) \mathbf{E}^x \int_0^t f(X(u)) \, du. \tag{4.18}$$

By (4.17) and (4.18), we have

$$\int_{\mathbb{R}^l} \mathbf{E}^x f(X(t)) \, \nu(dx) = \int_{\mathbb{R}^l} f(x) \, \nu(dx). \tag{4.19}$$

We now see from (4.19) that the measure given by $\mu(A) = \nu(A)/\nu(\mathbb{R}^l)$ defines the required stationary distribution. $\qquad\square$

Remark 4.2 The measure $\nu(A)$ is invariant (i.e. it satisfies condition (4.19)) even if no assumption is made concerning the finiteness of $\mathbf{E}^x \tau_1$. It is sufficient to require that the process $X(t)$ is U-recurrent. In this case the measure ν is merely σ-finite and $\nu(\mathbb{R}^l) = \infty$. For details, see [194], [117].

4.5 Strong Law of Large Numbers

Chung's proof [48] of the law of large numbers for random variables $\zeta_n = \sum_{i=1}^n f(X_i)$, where X_i is a recurrent Markov chain with countably many states, employs the following method. Each random variable ζ_n is split into components. The k-th component contains the terms $\sum f(X_i)$ for $\tau_k \le i < \tau_{k+1}$ (τ_k is the time

at which a fixed state of the chain is reached for the k-th time). According to this approach, ζ_n may be described as a sum of independent random variables and a certain (generally small) remainder term. Therefore the law of large numbers for ζ_n follows from the corresponding law for sums of independent random variables. Maruyama and Tanaka [193] adopt a similar approach to prove the law of large numbers for one-dimensional diffusion processes. This reduction is impossible for multi-dimensional processes, since a multi-dimensional process does not generally return to the initial point. However, the law of large numbers can be used for stationary sequences.

Theorem 4.2 *Suppose that condition* (B) *holds, and let* μ *be the stationary distribution of the process* $X(t)$, *constructed in the proof of Theorem* 4.1. *Let* $f(x)$ *be a function integrable with respect to the measure* μ. *Then*

$$\mathbf{P}\left\{ \frac{1}{T} \int_0^T f(X^x(t))\, dt \xrightarrow[T\to\infty]{} \int_{\mathbb{R}^l} f(y)\mu(dy) \right\} = 1 \qquad (4.20)$$

for all $x \in \mathbb{R}^l$.

Proof We first prove (4.20) for the initial distribution

$$\mathbf{P}\{X(0) \in \gamma\} = \tilde{\mu}(\gamma), \quad \gamma \in \Gamma_1, \qquad (4.21)$$

where $\tilde{\mu}(\gamma)$ is the stationary distribution of the Markov chain $\widetilde{X}_n$ existing according to Lemma 4.1. Under this assumption the sequence of random variables

$$\eta_n = \int_{\tau_n}^{\tau_{n+1}} f(X(t))\, dt$$

is a random sequence which is stationary in the narrow sense, and it follows from (4.14) that

$$\mathbf{E}\eta_n = \int_{\mathbb{R}^l} f(x)\nu(dx).$$

We easily see from (4.12) that the sequence η_n is metrically transitive.

Let $\nu_1(T)$ denote the number of cycles completed up to time T. It is obvious that

$$\int_0^T f(X(t))\, dt = \sum_{n=0}^{\nu_1(T)} \eta_n + \int_{\tau_{\nu_1(T)}}^T f(X(t))\, dt. \qquad (4.22)$$

We may assume without loss of generality that $f(x) > 0$. It then follows from (4.22) that

$$\sum_{n=0}^{\nu_1(T)} \eta_n \leq \int_0^T f(X(t))\, dt \leq \sum_{n=0}^{\nu_1(T)+1} \eta_n. \qquad (4.23)$$

Since the sequence $\eta_1, \ldots, \eta_n, \ldots$ is stationary and metrically transitive, the law of large numbers for such sequences (see [92]) implies the relation

$$\mathbf{P}\left\{\frac{1}{k}\sum_{n=0}^{k}\eta_n \underset{k\to\infty}{\longrightarrow} \int_{\mathbb{R}^l} f(x)\,\nu(dx)\right\} = 1. \tag{4.24}$$

In particular for $f(x) \equiv 1$ we get from the above that

$$\mathbf{P}\left\{\frac{\tau_k}{k} \underset{k\to\infty}{\longrightarrow} \int_{\mathbb{R}^l} \tilde{\mu}(dx)\mathbf{E}^x\tau = \nu(\mathbb{R}^l)\right\} = 1. \tag{4.25}$$

We claim that almost surely,

$$\frac{\tau_{\nu_1(T)}}{T} \to 1 \quad \text{as } T \to \infty. \tag{4.26}$$

Using (4.25) and the fact that $\nu_1(T) \to \infty$ as $T \to \infty$, we see that almost surely,

$$\lim_{T\to\infty}\frac{\tau_{\nu_1(T)+1}}{\tau_{\nu(T)}} = \lim_{T\to\infty}\frac{\tau_{\nu_1(T)+1}}{\nu_1(T)+1}\lim_{T\to\infty}\frac{\nu_1(T)}{\tau_{\nu_1(T)}} = 1.$$

Hence, using the inequalities $\tau_{\nu_1(T)} \le T \le \tau_{\nu_1(T)+1}$ we get (4.26). From (4.23)–(4.26) we readily infer that

$$\mathbf{P}\left\{\frac{1}{T}\int_0^T f(X(t))\,dt = \frac{\int_0^T f(X(t))\,dt}{\nu_1(T)}\frac{\nu_1(T)}{\tau_{\nu_1(T)}}\frac{\tau_{\nu_1(T)}}{T} \underset{T\to\infty}{\longrightarrow} \int_{\mathbb{R}^l} f(x)\mu(dx)\right\} = 1.$$

This proves (4.20) for the initial distribution (4.21). It follows that (4.20) is valid for almost all points $x \in \Gamma_1$ with respect to the measure $\tilde{\mu}(\gamma)$ (hence also with respect to the measure $\sigma_{\Gamma_1}(\gamma)$).

Let x be any point of $\mathbb{R}^l$. It is clear from the proof of Lemma 4.6 that the measure $\mathbf{P}^x\{X(\tau_1) \in \gamma\}$ is absolutely continuous with respect to the measure $\sigma_{\Gamma_1}(\gamma)$ for any $x \in \mathbb{R}^l$. Hence, using the equalities

$$\mathbf{P}^x\left\{\lim_{T\to\infty}\frac{1}{T}\int_0^T f(X(t))\,dt = a\right\}$$

$$= \mathbf{P}^x\left\{\lim_{T\to\infty}\frac{1}{T}\int_{\tau_1}^T f(X(t))\,dt = a\right\}$$

$$= \int_{y\in\Gamma_1}\mathbf{P}^x\{X(\tau_1) \in dy\}\mathbf{P}^y\left\{\lim_{T\to\infty}\frac{1}{T}\int_0^T f(X(t))\,dt = a\right\},$$

we get the assertion of the theorem for all $x \in \mathbb{R}^l$. $\square$

Corollary 4.3 *If $f(x)$ is bounded, then*

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \mathbf{E}^x f(X(t))\,dt = \int_{\mathbb{R}^l} f(x)\mu(dx). \tag{4.27}$$

In particular if $A \in \mathfrak{B}$, then

$$\frac{1}{T} \int_0^T P(x, t, A) \, dt \to \mu(A) \quad as \ T \to \infty. \tag{4.28}$$

These relations follow from (4.20) by Lebesgue's dominated convergence theorem.

Corollary 4.4 *If* (B) *holds, then the stationary distribution of the process $X(t)$ is unique.*

Indeed, let $\mu_1(A)$ be another stationary distribution. Then

$$\int_{\mathbb{R}^l} \mu_1(dx) P(x, t, A) = \mu_1(A). \tag{4.29}$$

Integrating (4.29) with respect to t from 0 to T and using (4.28), we get $\mu(A) = \mu_1(A)$.

The results (4.27) and (4.28) can be considerably strengthened. To be precise, one can prove that the transition probability function itself (not merely its Cesàro sense) tends to $\mu(A)$ as $t \to \infty$, provided certain additional assumptions are made. We present a general outline of one of the proofs of this theorem [54]. A Markov process $X(t)$ with stationary initial distribution (the existence of the latter was proved in Sect. 4.4) generates a dynamical system in the space of sample paths of the process. This system has a finite invariant measure. Employing the ergodicity property of this system (which follows from the results of this section), we conclude that there is no angular variable and that the μ-singular component of the transition probability tends to 0 as $t \to \infty$. (It can be shown that condition (B) is sufficient to guarantee these properties.) Applying the Von Neumann–Koopman mixing theorem (see Hopf [98, pp. 36–37]), one can prove that the transition probabilities tend to a limit. However, this method of proof, for all its generality, is probably difficult to extend to the time-non-homogeneous case. Therefore in Sects. 4.6 and 4.7 we shall employ a different method which does not depend on the general theory of dynamical systems and can be generalized to nonhomogeneous processes.

4.6 Some Auxiliary Results

In this section we shall derive some further consequences from condition (B), concerning the properties of the stationary distribution μ and the transition probability function $P(x, t, A)$ of the process $X(t)$. These properties will be used in our proof (Sect. 4.7) of the theorem on the limiting behavior of the transition probabilities as $t \to \infty$.

Lemma 4.8 *If condition* (B) *holds for* U, *then the stationary distribution* $\mu(A)$ *of the process* $X(t)$, *constructed in Sect.* 4.4, *possesses the property*

$$\inf_{x \in U} \mu(U_\delta(x)) = \beta(\delta) > 0 \quad \text{for } \delta > 0. \tag{4.30}$$

Proof It is clear from the construction of $\mu(A)$ in Sect. 4.4 that this measure is proportional to the mean time spent by the path of the process in the set A during one cycle, provided the initial distribution on Γ_1 coincides with $\tilde{\mu}(\gamma)$. This time is more or equal to the time $\tau^A(\Gamma_1)$ spent in the set A during a halfcycle (from the time the set Γ is reached to the time Γ_1 is reached).

The function $u(x) = \mathbf{E}^x \tau^{U_\delta(x_0)}(\Gamma_1)$ can be expressed as

$$u(x) = \mathbf{E}^x \int_0^{\tau_{U_1}} \mathbb{1}_{U_\delta(x_0)}(X(t)) \, dt.$$

Hence, applying Lemma 3.6,[2] we see that the function $u(x)$ is a solution in U_1 of the problem

$$Lu + \mathbb{1}_{U_\delta(x_0)}(x) = 0, \qquad u|_{\Gamma_1} = 0.$$

Since the Green function of this problem is bounded away from zero in every domain which, together with its boundary, lies in U_1, this implies the inequality

$$\inf_{x \in \Gamma, \, x_0 \in U} \mathbf{E}^x \tau^{U_\delta(x_0)}(\Gamma_1) > 0.$$

This inequality, together with the above-mentioned estimate

$$\mu(U_\delta(x_0)) \geq \inf_{x \in \Gamma, \, x_0 \in U} \mathbf{E}^x \tau^{U_\delta(x_0)}(\Gamma_1)$$

yields the assertion of the lemma.   □

Lemma 4.9 *If condition* (B) *holds, then the function* $P(x, t, A)$ *is a uniformly continuous function of* x *for* $x \in U, t > t_0, A \in \mathcal{B}$ *(where* $t_0 > 0$ *is arbitrary).*

Proof We have the following identity for events:

$$\{X(t) \in A\} = \{X(t) \in A, \ \tau_{U_1} > t\} \cup \{X(t) \in A, \ \tau_{U_1} \leq t\}.$$

Therefore, by the strong Markov property of the process $X(t)$, we get

---

[2]In Lemma 3.6, the integrand is assumed to be continuous. However, by "sandwiching" the function $\mathbb{1}_{U_\delta(x_0)}(x)$ between two continuous functions and using the properties of solutions of elliptic equations (recall that L is a nonsingular elliptic operator in U_1), one easily shows that Lemma 3.6 is valid also for this function. Analogous remarks apply to the other boundary-value problems considered in this and the following section.

$$P(x, t, A) = \mathbf{P}^x \{X(t) \in A, \ \tau_{U_1} > t\}$$

$$+ \int_{\Gamma_1} \int_{u=0}^{t} \mathbf{P}^x \{\tau_{U_1} \in du, \ X(\tau_{U_1}) \in dz\} P(z, t - u, A). \qquad (4.31)$$

By Lemma 3.3, this formula implies that the function $u(t, x) = P(x, t, A)$ satisfies in the cylinder $(t > 0) \times U_1$ the following nondegenerate parabolic equation

$$Lu(t, x) = -\frac{\partial u}{\partial t} + \sum_{i=1}^{l} b_i(x) \frac{\partial u}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} = 0. \qquad (4.32)$$

The boundary conditions for this equation are

$$u(0, x) = \mathbb{1}_A(x); \qquad u(t, x)|_{x \in \Gamma_1} = P(x, t, A)|_{x \in \Gamma_1}. \qquad (4.33)$$

Let us use now certain known estimates for the solutions of (4.32) with bounded measurable boundary conditions (4.33) (see [76, Theorem 3]). According to these, the function $u(t, x)$ has bounded derivatives with respect to x in every cylinder $\{t > t_0\} \times K$, where $K \subset U_1$ is compact and $t_0 > 0$. The proof is complete. $\qquad \square$

Lemma 4.10 *If condition* (B) *holds, then for any* $x \in \mathbb{R}^l$ *and* $\varepsilon > 0$ *there exist numbers* $R > 0$ *and* $t_0(x) > 0$ *such that*

$$P(x, t, U_R^c) = \mathbf{P}^x \{|X(t)| > R\} < \varepsilon \qquad (4.34)$$

for $t > t_0(x)$.

Proof Given an arbitrary $\varepsilon > 0$, it follows from Lemma 4.9 that there exists a $\delta > 0$ such that for $x_0 \in U \cup \Gamma$, $|x - x_0| < \delta$, $A \in \mathfrak{V}$, $t > t_0$,

$$|P(x, t, A) - P(x_0, t, A)| < \frac{\varepsilon}{4}. \qquad (4.35)$$

Let $\beta(\delta)$ be the number defined by (4.30). Choose $R > 0$ large enough so that the stationary measure of the set $U_R^c = \{|x| \geq R\}$ satisfies

$$\mu(U_R^c) < \frac{\varepsilon \beta(\delta)}{4}. \qquad (4.36)$$

We shall now make use of the equality (see (4.19))

$$\mu(U_R^c) = \int \mu(dx) P(x, t, U_R^c). \qquad (4.37)$$

By this equality, we get from (4.35) and (4.36) the estimate

$$\frac{\beta(\delta)\varepsilon}{4} > \mu(U_R^c) = \int \mu(dx) P(x, t, U_R^c)$$

$$\geq \int_{U_\delta(x_0)} \mu(dx) P(x, t, U_R^c) \geq \left[P(x_0, t, U_R^c) - \frac{\varepsilon}{4} \right] \mu(U_\delta(x_0))$$

for $x_0 \in U$. Thus, in view of (4.30) we have the inequality

$$P(x_0, t, U_R^c) \leq \frac{\varepsilon}{2}, \tag{4.38}$$

valid for $x_0 \in U \cup \Gamma$, $t > t_0$. Since $t_0 > 0$ is arbitrary, inequality (4.38) is valid for all $t > 0$.

Thus the lemma is proved for $x_0 \in U \cup \Gamma$. Now let $x_0 \in (U \cup \Gamma)^c$. Since the process $X(t)$ is recurrent relative to the set U, there exists a number $t_0(x_0) > 0$ such that

$$\mathbf{P}^{x_0}\{\tau_{U^c} > t_0(x)\} < \frac{\varepsilon}{2}. \tag{4.39}$$

By arguments similar to those used for showing (4.31), we deduce that

$$P(x_0, t, U_R^c) = \int_\Gamma \int_{s=0}^t \mathbf{P}^{x_0}\{\tau_{U^c} \in ds, X(\tau_{U^c}) \in dy\} P(y, t - s, U_R^c)$$
$$+ \mathbf{P}^{x_0}\{\tau_{U^c} > t, |X(t)| > R\}$$

for $t > t_0(x)$. Inequalities (4.38) and (4.39) imply (4.34). The proof is complete. $\square$

Lemma 4.11 *If condition* (B) *holds, then there exists a constant* $T > 0$ *such that for any* $\delta > 0$

$$\inf_{t>0, x, x_0 \in U \cup \Gamma} \mathbf{E}^x \int_t^{t+T} \mathbb{1}_{U_\delta(x_0)}(X(s)) \, ds > 0 \tag{4.40}$$

(*i.e., the mean time spent by the path, starting from* x, *in the* δ-*neighborhood of* x_0 *during the time interval* $[t, t + T]$ *is bounded away from zero uniformly in* t, $x_0 \in U$, $x \in U$).

Proof Let K be a compact set such that

$$P(x, t, K) > \frac{1}{2} \tag{4.41}$$

for all $x \in U$, $t > 0$. Such a set K exists by (4.38). Now, since $\mathbf{E}^x \tau_{U^c}$ is bounded in every compact set, we can choose a constant $T \geq 0$ such that

$$\mathbf{P}^x\{\tau_{U^c} > T - 1\} < \frac{1}{2} \tag{4.42}$$

holds for all $x \in K$.

We now consider the function

$$\mathbf{E}^x \int_0^{\tau_{U^c}(t)} \mathbb{1}_{U_\delta(x_0)}(X(s)) \, ds = v^{(\delta)}(t, x).$$

By Lemma 3.5, this function is a solution in $\{t > 0\} \times U_1^c$ of the parabolic equation[3]

$$Lv^{(\delta)}(t,x) + \mathbb{1}_{U_\delta(x_0)}(x) = 0,$$

satisfying the boundary conditions

$$v^{(\delta)}(0,x) = 0; \qquad v^{(\delta)}(t,x)|_{x \in \Gamma_1} = 0.$$

Since the operator L is non-singular in the above domain, it follows from the general properties of solutions of parabolic equations that

$$\inf_{x,x_0 \in U \cup \Gamma} v^{(\delta)}(1,x) > 0. \tag{4.43}$$

Next, we see that for $x \in U$

$$\mathbf{E}^x \int_t^{t+T} \mathbb{1}_{U_\delta(x_0)}(X(s))\, ds$$

$$\geq \int_{y \in K} P(x,t,dy) \mathbf{E}^y \int_0^T \mathbb{1}_{U_\delta(x_0)}(X(s))\, ds$$

$$\geq \int_{y \in K} P(x,t,dy) \int_{z \in U \cup \Gamma} \int_{u=0}^{T-1} \mathbf{P}^y \{\tau_{U^c} \in du, X(\tau_{U^c}) \in dz\} v^{(\delta)}(1,z).$$

Using (4.41)–(4.43), we infer from the last inequality that (4.40) holds.   □

Lemma 4.12 *If condition (B) is satisfied, then for any $\delta > 0$ there exists a constant $\alpha_\delta > 0$ such that for all $x \in \mathbb{R}^l$, $x_0 \in U$ and $t > t_0(x)$*

$$P(x,t,U_\delta(x_0)) > \alpha_\delta.$$

Proof (1) Since

$$\mathbf{E}^x \int_t^{t+T} \mathbb{1}_{U_\delta(x_0)}(X(s))\, ds = \int_t^{t+T} P(x,s,U_\delta(x_0))\, ds,$$

it follows from Lemma 4.11 that for any $x \in U$ there exists a sequence $s_1, s_2, \ldots, s_n,$ $\ldots$ $(s_n \in [(n-1)T, nT])$ such that

$$P(x,s_n,U_\delta(x_0)) > \gamma_\delta > 0.$$

Now let $t \in [(n-1)T, nT]$. Then

$$P(x,t,U_\delta(x_0)) \geq \int_{y \in U_{\delta/2}(x_0)} P(x,s_{n-1},dy) P(y,t-s_{n-1},U_\delta(x_0)). \tag{4.44}$$

---

[3] See footnote on p. 113.

It is clear that $0 < t - s_{n-1} < 2T$ and

$$P(y, t - s_{n-1}, U_\delta(x_0)) \geq \mathbf{P}^y\{X(t - s_{n-1}) \in U_\delta(x_0), \tau_{U_1} > t - s_{n-1}\}. \qquad (4.45)$$

By Lemma 4.8, the function $w(t, y) = \mathbf{P}^y\{X(t) \in U_\delta(x_0), \tau_{U_1} > t\}$ is a solution in $\{t > 0\} \times U_1$ of the equation $Lw(t, y) = 0$ satisfying the boundary conditions

$$w(0, y) = \mathbb{1}_{U_\delta(x_0)}(y), \qquad w(t, y)|_{y \in \Gamma_1} = 0.$$

From general properties of solutions of parabolic equations we readily conclude that

$$\inf_{0 \leq t \leq 2T, \ y \in U_{\delta/2}(x_0), \ x_0 \in U} w(t, y) > 0.$$

Using this estimate in combination with (4.44) and (4.45) we deduce that for all $x \in U$, $x_0 \in U$, $t > 0$, $\delta > 0$

$$P(x, t, U_\delta(x_0)) \geq 2\alpha_\delta > 0.$$

(2) Now let $x \in U^c$. Take $t_0(x)$ such that $P^x\{\tau_{U^c} > t_0(x)\} < 1/2$. Then for $t > t_0(x)$

$$P(x, t, U_\delta(x_0)) \geq \int_{z \in \Gamma} \int_{u=0} \mathbf{P}^x\{\tau_{U^c} \in du, X(\tau_{U^c}) \in dz\} P(x, t - u, U_\delta(x_0))$$

$$\geq \frac{1}{2} 2\alpha_\delta = \alpha_\delta.$$

This complete the proof of the lemma. $\square$

4.7 Existence of the Limit of the Transition Probability Function

Lemma 4.10 guarantees that the process $X^x(t)$ is bounded in probability for all $x \in \mathbb{R}^l$. As we observed in Sect. 2.1, this implies that the family of distributions $P(x, t_n, A)$ is weakly compact for any sequence $t_n \to \infty$. In this section we shall show that the limit does not depend on which sequence $t_n \to \infty$ is chosen. The proof uses essentially Lemmas 4.9 and 4.12.

Theorem 4.3 *If condition (B) holds, then for any continuous bounded function $f(x)$ and any Borel subset A with boundary Γ such that $\mu(\Gamma) = 0$, we have*

$$T_t f(x) = \int_{\mathbb{R}^l} P(x, t, dy) f(y) \to \int_{\mathbb{R}^l} f(y) \mu(dy), \qquad (4.46)$$

$$P(x, t, A) \to \mu(A) \qquad (4.47)$$

as $t \to \infty$.

Proof Let $f(x)$ be an arbitrary continuous bounded function in $\mathbb{R}^l$. Then the function $T_t f(x)$ is also bounded. It is also continuous in x uniformly in $t > t_0$ ($t_0 > 0$) on any compact subset K of U_1 (see proof of Lemma 4.9). Hence there exists a sequence t_n tending to infinity as $n \to \infty$ such that $\lim_{n \to \infty} T_{t_n} f(x) = c(x)$, where $c(x)$ is a function continuous on U_1, and the convergence is uniform in U.

We set

$$c_- = \min_{U \cup \Gamma} c(x) = c(x_1); \qquad c_+ = \max_{U \cup \Gamma} c(x) = c(x_2),$$
$$a = \inf_{x \in \mathbb{R}^l} \varliminf_{t \to \infty} T_t f(x); \qquad b = \sup_{x \in \mathbb{R}^l} \varlimsup_{t \to \infty} T_t f(x). \tag{4.48}$$

To prove (4.46), it will suffice to show that $a = b$. For if $a = b$, then the function $T_t f(x)$ has a limit independent of x as $t \to \infty$. This limit is then equal to $\int f(y)\mu(dy)$, by Theorem 4.2 (see Corollary 4.3).

By (4.48), there exist for any $\varepsilon > 0$ a point $x_0 \in \mathbb{R}^l$ and a sequence $t'_n \to \infty$ such that

$$T_{t'_n} f(x_0) < a + \varepsilon. \tag{4.49}$$

Also, for any $x \in \mathbb{R}^l$

$$T_t f(x) > a - \varepsilon \tag{4.50}$$

if t is sufficiently large. Moreover, for any compact subset $K \subset \mathbb{R}^l$ there exists a $t_0 > 0$ such that inequality (4.50) holds for $t > t_0$ and all $x \in K$. (Indeed, if A_{t_0} consists of all points x such that (4.50) is true for $t > t_0$, then the sets A_t obviously possess the property that $A_s \subset A_t$ for $s < t$ and the union $\bigcup_{t>0} A_t$ coincides with $\mathbb{R}^l$.)

Our next task is to prove that $c_+ \leq a$. To do this, we shall assume that

$$c_+ > a \tag{4.51}$$

and we shall derive a contradiction.

Set $\gamma = (c_+ - a)/2$ and choose a number $\delta > 0$ such that $c(x) > c_+ - \gamma$ for $x \in U_\delta(x_2)$. Since the sequence $T_{t_n} f(x)$ converges to $c(x)$ uniformly in U, we can find a number n_0 such that for $x \in U_\delta(x_2)$ and $n > n_0$, we also have

$$T_{t_n} f(x) > c_+ - \gamma. \tag{4.52}$$

Now, using Lemma 4.10, we choose a compact set $K_1 \supset U$ such that

$$P(x_0, t, K_1^c) < \varepsilon \tag{4.53}$$

for all $t > t_1$.

We next take $t_2 > 0$ large enough so that inequality (4.50) holds for all $x \in K_1$ if $t > t_2$. Finally, using Lemma 4.12, we choose t_3 such that

$$P(x_0, t, U_\delta(x_2)) > \alpha_\delta \tag{4.54}$$

for $t > t_3$.

We now consider some fixed number t_n ($n > n_0$), and then a number t'_k such that $t'_k - t_n \geq \max(t_1, t_2, t_3)$; it follows from (4.49), (4.50) and (4.53) that

$$a + \varepsilon > T_{t'_n} f(x_0) = \int T_{t'_n} f(y) P(x_0, t'_k - t_n, dy)$$

$$\geq \int_{U_\delta(x_2)} (c_+ - \gamma) P(x_0, t'_k - t_n, dy)$$

$$+ \int_{K_1 \setminus U_\delta(x_2)} (a - \varepsilon) P(x_0, t'_k - t_n, dy) - \varepsilon \max_{\mathbb{R}^l} |f(x)|.$$

Since $\gamma = (c_+ - a)/2$, this inequality can be rewritten as

$$a + \varepsilon \geq (a - \varepsilon) P(x_0, t'_k - t_n, K_1)$$

$$+ \gamma P(x, t'_k - t_n, U_\delta(x_2)) - \varepsilon \max_{\mathbb{R}^l} |f(x)|$$

$$\geq (a - \varepsilon)(1 - \varepsilon) + \gamma P(x_0, t'_k - t_n, U_\delta(x_2)) - \varepsilon \max_{\mathbb{R}^l} |f(x)|. \qquad (4.55)$$

From (4.54) and (4.55) follows now the estimate

$$\gamma < \frac{\varepsilon(2 + a + \max_{\mathbb{R}^l} |f(x)|)}{a_\delta}.$$

Since ε is arbitrarily small, it follows that $\gamma \leq 0$ and this contradicts our assumption that $c_+ > a$.

Consequently, $c_+ \leq a$. Similarly, we show that $c_- \geq b$. Thus $c_+ \leq a \leq b \leq c_- \leq c_+$, and so $a = b$. As we have mentioned above, this implies (4.46).

The validity of (4.46) means that the measure $P(x, t, A)$ converges weakly to $\mu(A)$ as $t \to \infty$ for every fixed x. It is known (see [232]) that this implies the second assertion of the theorem. $\qquad \square$

4.8 Some Generalizations

In this section we shall show how to generalize the results of Sects. 4.4 through 4.7 to a wider class of Markov processes and to diffusion processes in other phase spaces. At the end of the section we shall generalize the theorems of Sects. 4.5 and 4.7 to the nonergodic case.

1. We assumed above that the diffusion matrix is nonsingular on the domain U_1. An analysis of the proof shows that in fact a weaker condition is sufficient; for example, we need only require that there exist a sufficiently smooth positive Green function for the first boundary-value problem for the parabolic equation $Lu = 0$ in $(t > 0) \times U$.

2. Many of the results of Sects. 4.4 through 4.7 can be proved for Markov processes in a general Banach space. Neither is there any need to confine the discussion

to processes with continuous sample paths. Indeed, the decomposition of the path
into cycles, which was the basic construction in the previous sections, carries over
to processes whose paths may have discontinuities of the first kind (jumps). For
details, see [194] and [117].

All the results of Sects. 4.4 through 4.7, without exception, carry over to diffusion
processes on smooth manifolds. In particular, a process with nonsingular diffusion
matrix on a compact smooth manifold is always positive recurrent. This follows
from the fact that the mean exit time of the sample path of a nonsingular diffusion
process from a bounded domain is bounded (see Corollary 3.2).

3. As was already mentioned above, the construction used in Sect. 4.4 to yield
the stationary distribution carries over to U-recurrent processes with infinite mean
recurrence time for U. This measure is invariant in the sense of (4.19). By a slight
modification of the arguments in Sect. 4.5, one can then prove the following gener-
alization of Theorem 4.2.

Theorem 4.4 *Suppose that condition* (B.1) *holds, and let the process* $X(t)$ *be* U-
recurrent. Let $\nu(A)$ *be the measure defined by* (4.14) *and let* $f(x)$ *and* $g(x)$ *be
functions integrable with respect to this measure such that* $\int g(x)\nu(dx) \neq 0$.
Then for all $x \in \mathbb{R}^l$

$$\mathbf{P}^x\left\{\lim_{T\to\infty}\frac{\int_0^T f(X(t))\,dt}{\int_0^T g(X(t))\,dt}=\frac{\int f(x)\nu(dx)}{\int g(x)\nu(dx)}\right\}=1. \tag{4.56}$$

This result, deduced from somewhat different assumptions, may be found in
[194], [117].

Now, if the expected time at which U is reached is infinite, we have

$$\nu(\mathbb{R}^l)=\infty, \tag{4.57}$$

and thus, using a monotone increasing sequence of functions $g_n(x)$ such that
$g_n(x)\uparrow 1$ as $n\to\infty$, we get from (4.56) that

$$\mathbf{P}^x\left\{\frac{1}{T}\int_0^T f(X(t))\,dt\xrightarrow[T\to\infty]{}0\right\}=1$$

for any ν-integrable function $f(x)$. Taking expectations and applying the Lebesgue
theorem, as in (4.27) and (4.28), we conclude that if $f(x)$ is ν-integrable and K is
a compact set,

$$\frac{1}{T}\int_0^T \mathbf{E}^x f(X(t))\,dt\to 0, \qquad \frac{1}{T}\int_0^T P(x,t,K)\,dt\to 0 \tag{4.58}$$

as $T\to\infty$.

By methods similar to those used in Sects. 4.6, 4.7, one can show that for a null
recurrent process the assumptions of Theorem 4.4 imply

$$\lim_{t\to\infty}\mathbf{E}^x f(X(t))=0, \qquad \lim_{t\to\infty}P(x,t,K)=0. \tag{4.59}$$

We shall not go into the proof of these formulas.

4. It was proved in Sects. 4.4 through 4.7 that if condition (B) holds, then the Markov process $X(t)$ has a unique stationary distribution, whose "domain of attraction" is the entire space. However, one frequently encounters the situation in which there exist domains of attraction with different stationary distributions. To bring such situations into the range of our discussion, we need only modify condition (B.2).

Let us say that a set D is *invariant* for the process $X(t)$ if $\mathbf{P}^x\{X(t) \in D\} = 1$ for all $t > 0$, $x \in D$.

Suppose that, besides condition (B.1) *as formulated in Sect.* 4.4, *the following condition* (B.2$'$) *is satisfied*:

There exists a set $D \supset U$, *invariant for the process* $X(t)$, *such that* $D = \bigcup_{n=1}^{\infty} K_n$ *(K_n are compact subsets) and* $\sup_{K_n} \mathbf{E}^x \tau < \infty$ *for all* n.

Analyzing the previous proofs, we see that under conditions (B.1), (B.2$'$) there exists a unique stationary distribution μ for the process $X(t)$, such that $\mu(\mathbb{R}^l \setminus D) = 0$. Moreover, for all $x \in U$, the strong law of large numbers (4.20) and the theorem on the limiting behavior (4.46) and (4.47) of the transition probability are valid.

Call F a *set of inessential states* of the process $X(t)$ if

$$P(x, t, F) \to 0 \tag{4.60}$$

for all $x \in \mathbb{R}^l$, $t \to \infty$.

Now suppose that the entire phase space $\mathbb{R}^l$ can be decomposed as the union of a finite or countable family of invariant sets D_i and a set of inessential states $F = \mathbb{R}^l \setminus \bigcup D_i$. Assume further that each of the sets D_i contains a "mixing region" U_i, i.e., U_i and D_i satisfy conditions (B.1), (B.2$'$).

Each set D_i is called an *ergodic* set of the process $X(t)$. Suppose a specific stationary distribution $\mu_i(A)$ is established in each ergodic set. Since the sets D_i are mutually disjoint, the measures $\mu_i(A)$ are singular with respect to each other. It is easily shown that for any positive constants $k_1, k_2, \ldots$ such that $\sum_i k_i = 1$, the measure

$$\mu(A) = \sum_i k_i \mu_i(A) \tag{4.61}$$

is also stationary.

The converse is also valid: Any distribution which is stationary for the process $X(t)$ is expressible in the form (4.61).

In fact, let $\mu(A)$ be some stationary distribution. Then

$$\mu(A) = \int \mu(dx) P(x, t, A). \tag{4.62}$$

Setting $A = F$, letting $t \to \infty$ and using (4.60), we conclude that $\mu(F) = 0$. Now let $B \subset D_i$. Then

$$\mu(B) = \sum_j \int_{D_j} \mu(dx) P(x, t, B) = \int_{D_i} \mu(dx) P(x, t, B) \tag{4.63}$$

by the invariance of the set D_i. The invariance of D_i and (4.63) imply that the measure ν defined by $\nu(A) = \mu(A \cap D_i)/\mu(D_i)$, is a stationary distribution for $X(t)$ such that $\nu(\mathbb{R}^l \setminus D_i) = 0$. Since there is exactly one such measure, we get

$$\mu(A \cap D_i) = \mu(D_i)\mu_i(A).$$

Using this and the condition $\mu(F) = 0$, we finally get

$$\mu(A) = \sum_i \mu(D_i)\mu_i(A).$$

Suppose the point $X(0) = x$ belongs to ergodic set D_i. Then, as we already have seen above, the strong law of large numbers (4.20) and the theorem on the limiting behavior of the transition probability (4.46), (4.47) are valid, provided we set $\mu = \mu_i$ in these formulas.

We now consider the case $X(0) = x \in F$. First, since the event $B = \{X(t) \in F$ for all $t > 0\}$ implies the event $\{X(t) \in F\}$, it follows by (4.60) that $\mathbf{P}^x(B) = 0$. Using this fact and the strong law of large numbers (which is valid in each ergodic set), we get the equality

$$\mathbf{P}\left\{ \lim_{T \to \infty} \frac{\int_0^T f(X(t))\,dt}{T} = \xi \right\} = 1.$$

Here ξ is the random variable equal to

$$f_i = \int f(x)\mu_i(dx),$$

when the path $X(t)$ leads from F to the set D_i. Let $\pi_i(x)$ denote the probability that a path issuing from x "settles" in the set D_i. Since $\mathbf{P}^x(B) = 0$, it follows that $\sum_i \pi_i(x) = 1$. By the above, we have $\mathbf{P}^x\{\xi = f_i\} = \pi_i(x)$.

Similar arguments yield the conclusion that

$$\lim_{t \to \infty} P(x, t, A) = \sum_i \pi_i(x)\mu_i(A)$$

for $x \in F$.

4.9 Stabilization of the Solution of the Cauchy Problem for a Parabolic Equation

In previous sections we applied properties of solutions of parabolic and elliptic differential equations in order to study properties of diffusion processes. Recent years have seen the publication of numerous papers in which, conversely, probabilistic methods are employed to study the properties of solutions of second-order parabolic

and elliptic equations. In this section we shall see that the results proved in Sects. 4.4 through 4.8 yield information about the behavior of the solutions of the Cauchy problem for a parabolic equation for large time values.

Let $X(t)$ be a time-homogeneous Markov process, regular in $\mathbb{R}^l$, with generator

$$L = \frac{\partial}{\partial s} + \sum_{i=1}^{l} b_i(x) \frac{\partial}{\partial x} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j}$$

$$\equiv \frac{\partial}{\partial s} + \left(b(x), \frac{\partial}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(x), \frac{\partial}{\partial x} \right)^2.$$

As above, we shall assume that the coefficients $b(x)$, $\sigma_r(x)$ satisfy a Lipschitz condition

$$\sum_{r=1}^{k} |\sigma_r(x) - \sigma_r(y)| + |b(x) - b(y)| < B|x - y|$$

in every compact subset $K \subset \mathbb{R}^l$, where the Lipschitz constant B may depend on K. Let us assume now also that the nondegeneracy condition (B.1) holds in every compact set.

It is well known (see [66]) that the above conditions are sufficient for the existence of a solution of the problem

$$Lu = 0; \qquad u(0, x) = f(x) \tag{4.64}$$

in the domain $(s < 0) \times \mathbb{R}^l$ for any bounded continuous function $f(x)$. It follows from Lemma 3.3 and the remarks following its proof that this solution is unique and that it can be written as

$$u(s, x) = \mathbf{E}^{s,x} f(X(0)).$$

The process $X(t)$ is time-homogeneous. Therefore, making the substitution $s = -t$, we see that the function

$$u(-t, x) = u_1(t, x) = \mathbf{E}^x f(X(t)) \tag{4.65}$$

is a solution of the following initial-value problem, equivalent to problem (4.64):

$$\frac{\partial u_1}{\partial t} = \left(b(x), \frac{\partial}{\partial x} \right) u_1 + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(x), \frac{\partial}{\partial x} \right)^2 u_1, \qquad u_1(0, x) = f(x). \tag{4.66}$$

Lemma 4.13 *If the process $X(t)$ is regular and transient, then for any continuous initial function $f(x)$ with compact support the solution $u_1(t, x)$ of problem (4.66) tends to zero as $t \to \infty$. Moreover, the function*

$$v(x) = \int_0^\infty u_1(t, x)\, dt$$

is bounded in $\mathbb{R}^l$.

Proof Let $f(x) = 0$ for $x \in \mathbb{R}^l \setminus K$. Then (4.65) implies

$$|u_1(t, x)| \leq \max_{\mathbb{R}^l} |f(x)| P(x, t, K). \tag{4.67}$$

Hence the first assertion of the theorem follows by (4.6). The second assertion follows from (4.67) and the Corollary 4.2. $\qquad\square$

Remark 4.3 It can be shown that the function $v(x)$ satisfies the elliptic equation

$$\sum_{i=1} b_i(x) \frac{\partial v}{\partial x_i} + \frac{1}{2} \sum_{i,j=1} a_{ij}(x) \frac{\partial^2 v}{\partial x_i \partial x_j} = -f(x) \tag{4.68}$$

and

$$v(x) \to 0 \quad \text{as } |x| \to \infty. \tag{4.69}$$

Thus the conditions of the theorem are sufficient for the existence of a solution of problem (4.68), (4.69) in $\mathbb{R}^l$.

Consider the sufficient conditions for transience proved in Chap. 3. Using these and Lemma 4.13, one can derive sufficient conditions for the solution of problem (4.66) to tend to zero as $t \to \infty$ and for the existence of a solution of problem (4.68), (4.69). These conditions will then be given in terms of the coefficients of the equation.

Lemma 4.14 *Let $X(t)$ be a null recurrent process. Then, for any continuous initial function $f(x)$ with compact support, the solution $u_1(t, x)$ of problem (4.65) tends to zero as $t \to \infty$. If $f(x) \geq 0$ and $f(x) \not\equiv 0$, then*

$$\int_0^T u_1(t, x) \, dt \to \infty \quad \text{as } T \to \infty.$$

Proof The first assertion follows from (4.67) and (4.59), and the second from the equality

$$\int_0^T u_1(t, x) \, dt = \mathbf{E}^x \int_0^T f(X(t)) \, dt$$

and the fact that a recurrent process spends an infinite time in the neighborhood of any point. $\qquad\square$

Stronger results can be obtained if $X(t)$ is a positive recurrent process.

Lemma 4.15 *Let $X(t)$ be a positive recurrent process. Then, for any continuous bounded function $f(x)$, the solution $u_1(t, x)$ of problem (4.66) tends to a constant as $t \to \infty$. This constant is given by*

$$\int_{\mathbb{R}^l} f(x) \mu(dx),$$

where $\mu(A)$ is the stationary initial distribution of the process $X(t)$.

The proof follows from the representation (4.65) of the solution of problem (4.66) and from Theorem 4.3.

The following Lemma may be useful for the actual computation of the limit of $u_1(t, x)$ as $t \to \infty$.[4]

Lemma 4.16 *The distribution $\mu(A)$ has a density $p_0(x)$ with respect to Lebesgue measure in $\mathbb{R}^l$. This density is the unique bounded solution of the equation*

$$L^* p_0 \equiv \frac{1}{2} \sum_{i,j=1}^{l} \frac{\partial^2}{\partial x_i \partial x_j} (a_{ij}(x) p_0) - \sum_{i=1}^{l} \frac{\partial}{\partial x_i} (b_i(x) p_0) = 0, \qquad (4.70)$$

satisfying the additional condition

$$\int_{\mathbb{R}^l} p_0(x)\, dx = 1. \qquad (4.71)$$

Proof The stationary distribution $\mu(A)$ satisfies the condition

$$\mu(A) = \int_{\mathbb{R}^l} \mu(dx) P(x, t, A). \qquad (4.72)$$

Since the measure $P(x, t, A)$ has a density $p(x, t, y)$ with respect to Lebesgue measure, which is simply Green's function[5] of (4.66), it follows from (4.72) that the measure $\mu(A)$ also has a density $p_0(x)$; i.e.

$$\mu(A) = \int_A p_0(x)\, dx. \qquad (4.73)$$

From (4.72) and (4.73) we get the formula

$$p_0(y) = \int_{\mathbb{R}^l} p_0(x) p(x, t, y)\, dx. \qquad (4.74)$$

It is well known (see [101]) that the solution $v(t, x)$ of the problem

$$\frac{\partial v}{\partial t} = L^* v; \qquad v(0, x) = p_0(x),$$

can be written as

$$v(t, x) = \int_{\mathbb{R}^l} p_0(y) p(y, t, x)\, dy.$$

This, together with (4.74) shows immediately that the function $p_0(x)$ satisfies (4.70). Before proving uniqueness, we need the following lemma.

---

[4]Equation (4.70) for the function $p_0(x)$ is known as Fokker–Planck–Kolmogorov equation, see [144]. Lemmas 4.16 and 4.17 were proved in [100, Sect. 4] under more general assumptions.

[5]The existence of Green's function under these assumptions follows, e.g., from results of Eidelman [66] and Mikhailov [197].

Lemma 4.17 *If $X(t)$ is a positive recurrent process, then the function $p(x, t, y)$ satisfies the relation*

$$\lim_{t \to \infty} p(x, t, y) = p_0(y). \tag{4.75}$$

Proof By the Chapman–Kolmogorov equation

$$p(x, t, y) = \int_{\mathbb{R}^l} p(x, t - 1, z) p(z, 1, y) \, dz,$$

the function $p(x, t, y)$ satisfies

$$p(x, t, y) = \mathbf{E}^x p(X(t - 1), 1, y).$$

Hence, by Theorem 4.3, we have

$$\lim_{t \to \infty} p(x, t, y) = \int_{\mathbb{R}^l} p_0(z) p(z, 1, y) \, dz$$

for every y. This equality and (4.74) imply the desired assertion. $\square$

We can now complete the proof of Lemma 4.16. Let $q_0(x)$ be another solution of problem (4.70), (4.71). Then this is also a solution of the problem

$$\frac{\partial v}{\partial t} = L^* v; \qquad v(0, x) = q_0(x).$$

Therefore

$$q_0(x) = \int q_0(z) p(z, t, x) \, dz$$

and if we now let $t \to \infty$ and recall (4.75), we see that $q_0(x) = p_0(x)$. This completes the proof of Lemma 4.16. $\square$

Remark 4.4 Recall that sufficient conditions for recurrence and positive recurrence were established in Chap. 3. Using the results of Sects. 3.2 and 3.3, we can derive necessary and sufficient conditions for a process $X(t)$ to be transient, recurrent, null or positive. Such conditions will then be given in terms of the properties of the operator

$$L_1 = \sum_{i=1}^{l} b_i(x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j}.$$

For example, a regular process $X(t)$ is recurrent if and only if there exists a unique solution of the exterior Dirichlet problem for the equation $L_1 u = 0$ in some (hence every, see Sect. 3.2) domain D^c such that D is bounded. The process is positive recurrent if and only if the exterior Dirichlet problem for the equation $L_1 u = -1$ has a positive solution. These statements follow from Theorems 3.10, 3.11 and the results of Sects. 3.2 and 3.3.

Remark 4.5 Many results of this chapter can be generalized to the time non-homogeneous case. For example, in [100] and [123] sufficient conditions are established in terms of Lyapunov functions for a nonhomogeneous Markov process to be bounded in probability. It has been shown that, under certain assumptions on the coefficients, a time non-homogeneous process possesses a certain analog of a stationary distribution, i.e., a distribution to which all others ultimately converge (see [100, Theorem 5]). It is not difficult to prove that if the transition probability function is periodic, this distribution is periodic as a function of time and determines a periodic Markov process.

4.10 Limit Relations for Null Recurrent Processes

In Sect. 4.5 we studied the behavior of the functional

$$\zeta(T) = \int_0^T f(X(t))\,dt$$

as $T \to \infty$, assuming condition (B) to hold. We showed that $\zeta(T)/T$ converges to a constant a.s., and $\mathbf{E}\zeta(T)/T$ converges to the same constant. If $X(t)$ is a null recurrent process, the random variable $T^{-1}\zeta(T)$ converges in probability to zero if f is ν-integrable. This leads one to expect the existence of a non-trivial limit distribution of $\zeta(T)$ for normalizing factors other than T^{-1}.

The analogous problem for Markov chains with countably many states was studied by Feller in his well known paper [71]. He proved there that the limit distribution of the number of hits in each state of the chain depends essentially on the distribution of the random variable τ equal to the length of one cycle (see Sect. 4.5). Since in the case at hand we have $\mathbf{E}\tau = \infty$, it is natural to conjecture that τ belongs to the domain of attraction of a stable law with exponent $\alpha < 1$. It turns out that in this case the limit distribution of $T^{-\alpha}\zeta(T)$ for a one-dimensional diffusion process coincides with that established by Feller in [71]. What is new here is that one can establish conditions for convergence to various laws, and also the values of the normalizing factors in terms of the coefficients of the generator of the Markov process $X(t)$.

Our exposition will be based on certain known facts. For the reader's convenience we summarize these in the following two lemmas.

The reader can find the proofs in Sects. 13.5 and 13.6 of [74].

Lemma 4.18 *Let $\alpha > 0$ and let $G(x)$ be a function, monotone on the half-line $[0, \infty)$, whose Laplace transform*

$$g(s) = \int_0^\infty e^{-sx}\,dG(x)$$

exists for $s > 0$. Then either of the relations

$$G(x) \sim \frac{cx^{\alpha}}{\Gamma(1+\alpha)}, \qquad x \to \infty,$$
$$g(s) \sim cs^{-\alpha}, \qquad s \to 0,$$

implies the other. (*As usual,* $\Gamma(\alpha)$ *denotes the gamma-function.*)

Lemma 4.19 *Let* $0 < \alpha < 1$, *and let* $\xi_1, \ldots, \xi_n, \ldots$ *be independent identically distributed random variables,* $F(x) = P\{\xi_1 \geq x\}$, *and* $F(0) = 1$. *Let* $G_{\alpha}(x)$ *be the distribution function of a stable law with exponent* α, *whose Laplace transform is* $\exp\{-s^{\alpha}\}$. *Then any one of the relations*

$$F(x) \sim \frac{cx^{-\alpha}}{\Gamma(1-\alpha)}, \qquad\qquad x \to \infty, \qquad (4.76)$$

$$\varphi(s) = \mathbf{E}e^{-s\xi_i} \sim 1 - cs^{\alpha}, \qquad s \to 0, \qquad (4.77)$$

$$P\left\{\frac{\xi_1 + \cdots + \xi_n}{(nc)^{1/\alpha}} < x\right\} \to G_{\alpha}(x), \qquad n \to \infty, \qquad (4.78)$$

implies the two others.

Let $\nu(T)$ denote the integer-valued random variable defined by

$$\xi_1 + \cdots + \xi_{\nu(T)} \leq T, \qquad \xi_1 + \cdots + \xi_{\nu(T)} + \xi_{\nu(T)+1} > T.$$

The following lemma is a very special case of the above-mentioned results.

Lemma 4.20 *Let the random variables* ξ_i *satisfy the assumptions of Lemma 4.19 and suppose* (4.76) *holds. Then*

$$\mathbf{P}\left\{\frac{\nu(T)}{T^{\alpha}} > \frac{x}{c}\right\} \to G_{\alpha}(x^{-1/\alpha}), \qquad \mathbf{E}\nu(T) \sim \frac{T^{\alpha}}{c\Gamma(1+\alpha)} \qquad (4.79)$$

also as $T \to \infty$.

Proof The definition of $\nu(T)$ implies that

$$\mathbf{P}\{\nu(T) \geq k\} = \mathbf{P}\{\xi_1 + \cdots + \xi_k < T\}.$$

Using (4.78) and setting $k = [xT^{\alpha}/c]$, we get the first assertion of the lemma. Next, we have

$$\mathbf{E}\nu(T) = \sum_{k=1}^{\infty} \mathbf{P}\{\nu(T) \geq k\} = \sum_{k=1}^{\infty} \mathbf{P}\{\xi_1 + \cdots + \xi_k < T\}.$$

Applying Lemma 4.19 again, we see that

$$A(s) = \int_0^{\infty} e^{-sT} \, d\mathbf{E}\nu(T) = \sum_{k=1}^{\infty} [\varphi(s)]^k = \frac{\varphi(s)}{1 - \varphi(s)} \sim \frac{1}{c} s^{-\alpha} \qquad (4.80)$$

as $s \to 0$. The relations (4.79) now follow from (4.80) and Lemma 4.18. $\qquad\qquad\square$

We now proceed to investigate the limiting distribution of the functional $\zeta(T)$ for a time-homogeneous recurrent diffusion process $X(T)$ on the line.

We may assume without loss of generality that the drift coefficient of this process vanishes (otherwise we can perform a transformation of coordinates $x_1 = W(x)$; see Example 3.10). Thus the process $X(t)$ is described by the differential operator

$$\frac{\partial}{\partial t} + \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2} \qquad (\sigma^2(x) > 0). \tag{4.81}$$

The uniqueness and recurrence of the process $X(t)$ associated with the operator (4.81) follow from the results of Chap. 3, provided the function $\sigma(x)$ satisfies a Lipschitz condition in every compact set and moreover $0 < \sigma^2(x) < k(x^2 + 1)$, where k is a constant, for $x \in \mathbb{R}^1$. Throughout Sects. 4.10 and 4.11 we shall assume that these conditions are satisfied.

As in Sect. 4.5, we shall start from the expansion (compare (4.22))

$$\zeta(T) = \sum_{n=1}^{v(T)} \eta_n + \rho(T), \qquad \rho(T) = \int_{\tau_{v(T)}}^T f(X(t))\,dt. \tag{4.82}$$

Here η_k is the increment of the functional $\zeta(T)$ during the k-th cycle. It will be convenient to define now a cycle as the portion of the path of $X(t)$ beginning at $x = 0$ and extending up to the first time of return to zero after the point $x = 1$ has been reached. With this definition of cycles consider the random variables τ_n equal to the time at which the n-th cycle ends ($\tau_0 = 0$). They satisfy

$$\eta_n = \int_{\tau_{n-1}}^{\tau_n} f(X(t))\,dt$$

and they are independent and identically distributed, if $x(0) = 0$. The lengths of the cycles, $\xi_n = \tau_n - \tau_{n-1}$, are also obviously independent and identically distributed. To study their distribution we shall need the following lemma.

Lemma 4.21 *Let $\tau^{(0)}$ ($\tau^{(1)}$) be the time needed by the path of the process $X(t)$ to reach the point $x = 0$ ($x = 1$). Then the function*

$$u_0(s, x) = \mathbf{E}^x \exp\{-s\tau^{(0)}\} \qquad (u_1(s, x) = \mathbf{E}^x \exp\{-s\tau^{(1)}\})$$

is the unique bounded solution of the equation

$$\frac{1}{2}\sigma^2(x)\frac{d^2u}{dx^2} - su = 0 \tag{4.83}$$

in the domain $x > 0$ ($x < 1$) satisfying the condition $u_0(s, 0) = 1$ ($u_1(s, 1) = 1$).

Proof We set $\tau_n^{(0)} = \min\{\tau^{(0)}, \tau^{(n)}\}$, where $\tau^{(n)}$ is the time at which the path reaches the point $x = n$. We also set $\chi_n = 1$ if $X(\tau_n^{(0)}) = 0$ and $\chi_n = 0$ if $X(\tau_n^{(0)}) = n$. In-

equalities for the random variables imply corresponding inequalities for their expectations:

$$v_n(s, x) = \mathbf{E}^x[\exp\{-s\tau_n^{(0)}\}\chi_n] \le u_0(s, x) \le \mathbf{E}^x \exp\{-s\tau_n^{(0)}\} = w_n(s, x).$$

It follows from the Feynman–Kac formula that the functions $v_n(s, x)$ and $w_n(s, x)$ are solutions of (4.83) in the domain $0 < x < n$, satisfying the boundary conditions $v_n(s, 0) = 1$, $v_n(s, n) = 0$ and $w_n(s, 0) = 1$, $w_n(s, n) = 1$. Since the process $X(t)$ is regular and recurrent, we have

$$\lim_{n\to\infty} w_n(s, x) = \lim_{n\to\infty} v_n(s, x) = u_0(s, x).$$

On the other hand, it is clear that any solution of (4.83) which is bounded in $x > 0$ and such that $u(s, 0) = 1$, lies between $v_n(x)$ and $w_n(s, x)$ for $0 \le x \le n$. This completes the proof of the lemma. $\square$

Remark 4.6 The length of a single cycle ξ_n is equal to the sum of two independent random variables: the lengths of the half-cycles. Hence the functions $u_0(s, x)$, $u_1(s, x)$ and the Laplace transform of the distribution of ξ_n satisfy the equality

$$\mathbf{E}e^{-s\xi_n} = \mathbf{E}_0 e^{-s\tau^{(1)}} \mathbf{E}_1 e^{-s\tau^{(0)}} = u_1(s, 0)u_0(s, 1). \tag{4.84}$$

Lemma 4.22 *If the integral*

$$\bar{f} = 2 \int_{-\infty}^{\infty} \frac{f(x)}{\sigma^2(x)} \, dx \tag{4.85}$$

is absolutely convergent, then the expectation of the random variable η_k exists and is equal to $\bar{f}$.

Proof We have

$$\mathbf{E}\eta_k = \mathbf{E}^0 \int_0^{\tau^{(1)}} f(X(t)) \, dt + \mathbf{E}^1 \int_0^{\tau^{(0)}} f(X(t)) \, dt. \tag{4.86}$$

The random variable $\int_0^{\tau^{(0)}} f(X(t)) \, dt$ has an expectation if the sequence $\mathbf{E}^1 \int_0^{\tau_n^{(0)}} |f(X(t))| \, dt$ has a finite limit as $n \to \infty$. By Lemma 3.6, the function

$$u_n(x) = \mathbf{E}^x \int_0^{\tau_n^{(0)}} |f(X(t))| \, dt$$

is a solution in the domain $0 < x < n$ of the problem

$$\frac{\sigma^2(x)}{2} \frac{d^2 u_n}{dx^2} + |f(x)| = 0; \qquad u_n(0) = u_n(n) = 0.$$

This implies easily that

$$\mathbf{E}^x \int_0^{\tau^{(0)}} |f(X(t))|\, dt = \lim_{n\to\infty} u_n(x) = 2 \int_0^x dz \int_z^\infty \frac{|f(y)|}{\sigma^2(y)}\, dy,$$

and similarly

$$\mathbf{E}^x \int_0^{\tau^{(0)}} f(X(t))\, dt = 2 \int_0^x dz \int_z^\infty \frac{f(y)}{\sigma^2(y)}\, dy.$$

In an analogous way one proves the equality

$$\mathbf{E}^x \int_0^{\tau^{(1)}} f(X(t))\, dt = 2 \int_x^1 dz \int_{-\infty}^z \frac{f(y)}{\sigma^2(y)}\, dy$$

valid for $x < 1$. Together with (4.86), these equalities imply the desired assertion. $\square$

4.11 Limit Relations for Null Recurrent Processes (Continued)[6]

Theorem 4.5 *Suppose that the distribution of the length of a single cycle ξ_n belongs to the domain of normal attraction of a stable law with exponent $\alpha < 1$, i.e.,*

$$\mathbf{P}\{\xi_n \geq T\} \sim cT^{-\alpha} \quad (T \to \infty). \tag{4.87}$$

Assume that the integral (4.85) is absolutely convergent and non-zero.

Then the process $X(t)$ defined by the operator (4.81) satisfies the relations, as $T \to \infty$

$$\mathbf{P}^x \left\{ \frac{\int_0^T f(X(t))\, dt}{\bar{f} T^\alpha} < \frac{z}{c\Gamma(1-\sigma)} \right\} \to 1 - G_\alpha(z^{-1/\alpha}), \tag{4.88}$$

$$\mathbf{E}^x \int_0^T f(X(t))\, dt \sim \frac{\bar{f} T^\alpha}{c\Gamma(1+\alpha)\Gamma(1-\alpha)}. \tag{4.89}$$

Proof (1) Since the process $X(t)$ is recurrent, we may assume without loss of generality that $X(0) = 0$. We now observe that the random variables $\nu(T)$ in the expansion (4.82) and the variables ξ_i satisfy the assumptions of Lemma 4.20. Therefore, as $T \to \infty$,

$$\mathbf{P} \left\{ \frac{\nu(T)}{T^\alpha} > \frac{x}{c\Gamma(1-\alpha)} \right\} \sim G_\alpha(x^{-1/\alpha}), \tag{4.90}$$

$$\mathbf{E}\nu(T) \sim \frac{T^\alpha}{\Gamma(1-\alpha)\Gamma(1+\alpha)}. \tag{4.91}$$

---

[6]The theorems of this and next sections are proved in [129, 131, 132].

It follows from the expansion (4.82) that

$$\frac{\zeta(T)}{T^\alpha} = \frac{\eta_1 + \cdots + \eta_{\nu(T)}}{\nu(T)} \frac{\nu(T)}{T^\alpha} + \frac{\rho(T)}{T^\alpha}. \tag{4.92}$$

It is clear that $\nu(T) \to \infty$ almost surely as $T \to \infty$. Hence, by the strong law of large numbers, we get

$$\mathbf{P}\left\{\frac{\eta_1 + \cdots + \eta_{\nu(T)}}{\nu(T)} \to \bar{f} \text{ as } T \to \infty\right\} = 1.$$

This equality, together with (4.90) and (4.92), imply that (4.88) holds, provided that $\rho(T)T^{-\alpha} \to 0$ in probability as $T \to \infty$.

(2) If the function $f(x)$ is non-negative (the general case can be reduced to this by expressing $f(x)$ as the difference of two non-negative functions), then obviously $0 \le \rho(T) \le \eta_{\nu(T)+1}$. It will therefore suffice to prove that $T^{-\alpha}\eta_{\nu(T)+1}$ converges in mean to zero, i.e., that

$$\lim_{T\to\infty} T^{-\alpha}\mathbf{E}\eta_{\nu(T)+1} = 0. \tag{4.93}$$

For any n, the random variables η_{n+1}, ξ_{n+1} are independent of the history of the process up to time τ_n. Hence

$$\mathbf{E}\eta_{\nu(T)+1} = \mathbf{P}\{\tau_1 \ge T\}\mathbf{E}\{\eta_1 \mid \tau_1 \ge T\}$$

$$+ \sum_{n=1}^{\infty} \int_{y=0}^{T} \mathbf{P}\{\tau_n \in dy, \xi_{n+1} > T - y\}\mathbf{E}\{\eta_{n+1} \mid \xi_{n+1} > T - y\}$$

$$= \mathbf{P}\{\tau_1 \ge T\}\mathbf{E}\{\eta_1 \mid \tau_1 \ge T\} + \int_{y=0}^{T} \mathbf{E}(\eta_{n+1}\chi_{n+1})d\left(\sum_{n=1}^{\infty}\mathbf{P}\{\tau_n < y\}\right),$$

where $\chi_{n+1} = 1$ if $\xi_{n+1} \ge T - y$, and $\chi_{n+1} = 0$ if $\xi_{n+1} < T - y$.

The function $\varphi(T - y) = \mathbf{E}(\eta_{n+1}\chi_{n+1})$ is obviously independent of n, positive, bounded by $\mathbf{E}\eta_i = \bar{f}$, and convergent to zero as $T - y \to \infty$. Moreover, $\sum_{n=1}^{\infty}\mathbf{P}\{\tau_n < y\} = \mathbf{E}\nu(y)$ (see Sect. 4.10). Hence, by (4.91), it follows

$$T^{-\alpha} \int_0^T \varphi(T - y)\, d\mathbf{E}\nu(y)$$

$$\le \frac{\mathbf{E}\nu(T - \sqrt{T})}{T^\alpha} \sup_{s>\sqrt{T}} \varphi(s) + \left[\frac{\mathbf{E}\nu(T)}{T^\alpha} - \frac{\mathbf{E}\nu(T - \sqrt{T})}{T^\alpha}\right]\bar{f} \to 0$$

as $T \to \infty$. This proves (4.93), and hence also (4.88).

(3) As before, we may assume without loss of generality that $f(x) \ge 0$. It then follows from (4.82) and (4.93) that

$$\lim_{T\to\infty} \frac{\mathbf{E}^0 \int_0^T f(X(t))\, dt}{T^\alpha} = \lim_{T\to\infty} \frac{\mathbf{E}^0(\eta_1 + \cdots + n_{\nu(T)+1})}{T^\alpha}.$$

We now observe that the event $\{v(T) + 1 = n\}$ depends only on the history of the process $X(t)$ up to time τ_n. Thus, applying Wald's identity (see [146]), we see that

$$\mathbf{E}[\eta_1 + \cdots + \eta_{\nu(T)+1}] = (\mathbf{E}\nu(T) + 1)\,\bar{f}.$$

Together with (4.91), the last two equalities imply (4.89) for $x = 0$. Since the integral of the function $f(X(t))$ up to the time at which the point $x = 0$ is reached has finite expectation, it follows now that (4.89) is valid for any x. This completes the proof of the theorem. $\square$

We now present conditions for the relation (4.87) to hold. By (4.84) and Lemmas 4.19 and 4.21, a sufficient condition is that the solutions $u_0(s, x)$ and $u_1(s, x)$ of (4.83) satisfy

$$u_i(s, x) \sim 1 - c_i(x)s^\alpha, \quad i = 0, 1,$$

as $s \to +0$. We are thus led to consider the asymptotic behavior as $s \to +0$ of the unique bounded for $x \geq 0$ solution of the problem

$$\tilde{y}_s'' - sp(x)\tilde{y}_s = 0, \qquad \tilde{y}_s(0) = 1. \tag{4.94}$$

To simplify the notation, we have set here $p(x) = 2/\sigma^2(x)$. We shall compare the solution of problem (4.94) with the solution of the "model" problem

$$y_s'' - Asx^\beta y_s = 0, \qquad y_s(0) = 1, \tag{4.95}$$

with $\beta > -1$, $A > 0$. Determination of the required asymptotic behavior for (4.95) is trivial. In fact, by the substitution $\xi = (As)^\alpha x$, $\alpha = (\beta + 2)^{-1}$ we reduce (4.95) to the problem ($\xi \geq 0$)

$$\frac{d^2 Z_\beta}{d\xi^2} - \xi^\beta Z_\beta = 0, \qquad Z_\beta(0) = 1. \tag{4.96}$$

Integrating (4.96), we easily see that if $\beta > -1$, the function $Z_\beta(\xi)$ has a derivative at zero, which must be bounded because of the boundedness of the function Z_β and the structure of (4.96). Therefore $Z_\beta(\xi) \sim 1 - c_\alpha\xi$ ($\xi \to 0$). Hence we conclude that for $s \to 0$

$$y_s(x) = Z_\beta(A^\alpha s^\alpha x) \sim 1 - c_\alpha A^\alpha x s^\alpha, \quad c_\alpha = -Z_\beta'(0). \tag{4.97}$$

It is known (see [113, Chap. 3.2]) that the solution of (4.96) can be written as

$$Z_\beta(x) = cx^{1/2} K_\alpha(2\alpha x^{1/(2\alpha)}); \tag{4.98}$$

here $K_\alpha(x)$ is the Bessel function. We note also that formula (4.98) provides an easy way of computing constant c_α in (4.97):

$$c_\alpha = \frac{\alpha^{2\alpha}\,\Gamma(1 - \alpha)}{\Gamma(1 + \alpha)}. \tag{4.99}$$

Theorem 4.6 *Assume that for some constants $\beta > -1$, $\tilde{p}_+$ and $\tilde{p}_-$ such that $\tilde{p}_+ + \tilde{p}_- > 0$ the following condition is valid*:

$$\lim_{T \to \pm\infty} \frac{1}{T} \int_0^T |x|^{-\beta} p(x)\, dx = \tilde{p}_{\pm}. \tag{4.100}$$

Then for the length of a single cycle ξ_n we have

$$\mathbf{P}\{\xi_n > T\} \sim \frac{\alpha^{2\alpha}(\tilde{p}_+ + \tilde{p}_-)}{\Gamma(1+\alpha)} T^{-\alpha} \quad \text{as } T \to \infty. \tag{4.101}$$

Proof It follows from (4.84) and Lemmas 4.19, 4.21 that it is enough to establish asymptotic properties for $s \to 0$ of the unique bounded solutions of problems

$$\frac{d^2 u_0(s, x)}{dx^2} - sp(x)u_0(s, x) = 0; \qquad u_0(s, 0) = 1, \quad x \geq 0 \tag{4.102}$$

and

$$\frac{d^2 u_1(s, x)}{dx^2} - sp(x)u_1(s, x) = 0; \qquad u_1(s, 1) = 1, \quad x \leq 1. \tag{4.103}$$

It follows from the definition of $u_i(s, x)$ that

$$0 < u_i(s, x) \leq 1. \tag{4.104}$$

Making substitutions $x = z/s^\alpha$ and $U_{s,i}(z) = u_i(s, z/s^\alpha)$ and using $1 - 2\alpha = \alpha\beta$ lead to

$$\frac{d^2 U_{s,i}(z)}{dz^2} - s^{\alpha\beta} p\left(\frac{z}{s^\alpha}\right) U_{s,i}(z) = 0. \tag{4.105}$$

Upon using (4.100), (4.104) and (4.105), for a fixed $z_0 > 0$ and $h > 0$, we have

$$0 < \frac{dU_{s,i}(z_0 + h)}{dz} - \frac{dU_{s,i}(z_0)}{dz} = s^{\alpha\beta} \int_{z_0}^{z_0+h} p(xs^{-\alpha}) U_{s,i}(x)\, dx$$

$$\leq s^{\alpha\beta} \int_{z_0}^{z_0+h} \frac{p(xs^{-\alpha})}{|xs^{-\alpha}|^\beta} |xs^{-\alpha}|^\beta\, dx$$

$$\leq \max(z_0^\beta, |z_0 + h|^\beta) s^\alpha \int_{z_0 s^{-\alpha}}^{(z_0+h)s^{-\alpha}} \frac{p(x)}{|x|^\beta}\, dx$$

$$\leq K \max(z_0^\beta, |z_0 + h|^\beta) h + \rho_s(h, z_0), \tag{4.106}$$

where $\rho_s(h, z_0) \to 0$ uniformly in $h > 0$ as $s \to 0$. Note that in the above and hereafter K is a generic positive constant; its values may be different for different appearances.

It is also important to study the increment of $(d/dz)U_{s,0}$ at $z_0 = 0$. In fact,

$$\left| \frac{dU_{s,0}(h)}{dz} - \frac{dU_{s,0}(0)}{dz} \right| \le s^{\alpha\beta} \int_0^h p(xs^{-\alpha})\, dx = A(s,h).$$

Let $0 < h \le s^\alpha$. Then $A(s,h) \le Ks^{\alpha\beta}h$. Choosing $\gamma < 1$ satisfying the inequality $\gamma + \beta > 0$ (recalling that $\beta > -1$), we arrive at

$$A(s,h) \le Ks^{\alpha(\gamma+\beta)}h^{1-\gamma} \le Kh^{1-\gamma}.$$

Next let $h > s^\alpha$, and $\beta < 0$. By virtue of (4.100)

$$A(s,h) = s^{\alpha\beta}\left(\int_0^{s^\alpha} p(xs^{-\alpha})\, dx + \int_{s^\alpha}^h p(xs^{-\alpha})\, dx \right)$$

$$\le Ks^{\alpha\beta+\alpha} + s^{\alpha\beta} \int_{s^\alpha}^h \frac{p(xs^{-\alpha})}{|xs^{-\alpha}|^\beta}\, dx \le ch^{1+\beta}.$$

It is clear that the same inequality also holds for $h > s^\alpha$, $\beta \ge 0$.

It follows from the above arguments that $\{U_{s,0}(z)\}$ and $\{(d/dz)U_{s,0}(z)\}$ are uniformly (in s) bounded on any compact subset of $\mathbb{R}^+ = \{z \ge 0\}$. Using this fact together with (4.106), we can extract convergent subsequence $\{U_{s_n,0}(\cdot)\}$ and obtain $\{(d/dz)U_{s_n,0}(\cdot)\}$ such that the subsequence converge uniformly on any compact set to $U_{0,0}(\cdot)$ and $(d/dz)U_{0,0}(\cdot)$, respectively. Using the integral mean value theorem, we obtain

$$\frac{dU_{s_n,0}(z_0+h)}{dz} - \frac{dU_{s_n,0}(z_0)}{dz} = z_0^\beta \int_{z_0}^{z_0+h} \left(\frac{z}{s_n^\alpha} \right)^{-\beta} p(zs_n^{-\alpha})\, dz\, U_{s_n,0}(z_0) + o(h)$$

as $h \to 0$. Make a change of variable $z/s_n^\alpha = x$ leading to

$$\frac{dU_{s_n,0}(z_0+h)}{dz} - \frac{dU_{s_n,0}(z_0)}{dz} = z_0^\beta s_n^\alpha \int_{z_0/s_n^\alpha}^{(z_0+h)/s_n^\alpha} x^{-\beta} p(x)\, dx\, U_{s_n,0}(z_0) + o(h).$$

Sending $s_n \to 0$ and using (4.100) yield

$$\frac{dU_{0,0}(z_0+h)}{dz} - \frac{dU_{0,0}(z_0)}{dz} = z_0^\beta \tilde{p}_+ U_{0,0}(z_0)h + o(h).$$

Letting $h \to 0$, and noting that z_0 is arbitrary, we obtain the following limit differential equation

$$\frac{d^2 U_{0,0}(z)}{dz^2} - z^\beta \tilde{p}_+ U_{0,0}(z) = 0, \quad \text{for } z > 0. \tag{4.107}$$

Equation (4.107) has a unique bounded solution satisfying the initial condition $U_{0,0}(0) = 1$. This yields that for any $A > 0$, the limits

$$\lim_{s\to 0} U_{s,0}(z) = U_{0,0}(z), \qquad \lim_{s\to 0} \frac{dU_{s,0}(z)}{dz} = \frac{dU_{0,0}(z)}{dz}, \tag{4.108}$$

uniformly on $[0, A]$ and the solution is precisely the solution of (4.107) with initial condition $U_{0,0}(0) = 1$. It follows from (4.95)–(4.97) that

$$\frac{dU_{0,0}(0)}{dz} = -c_\alpha \tilde{p}^\alpha_+, \quad \text{with } c_\alpha = \frac{\alpha^{2\alpha}\Gamma(1-\alpha)}{\Gamma(1+\alpha)}. \tag{4.109}$$

By virtue of (4.108) and (4.109), as $s \to 0$,

$$u_{s,0}(1) = U_{s,0}(s^\alpha) = 1 - c_\alpha \tilde{p}^\alpha_+ s^\alpha + o(s^\alpha). \tag{4.110}$$

Analogously as $s \to 0$,

$$u_{s,1}(0) = U_{s,1}(s^\alpha) = 1 - c_\alpha \tilde{p}^\alpha_- s^\alpha + o(s^\alpha). \tag{4.111}$$

Assertion of theorem follows from (4.84), (4.110) and (4.111). $\qquad\qquad\square$

Theorems 4.5 and 4.6 imply

Corollary 4.5 *Let $p(x) = 2\sigma^{-2}(x)$ satisfy the condition (4.100). Then the assertion of Theorem 4.5 is valid with $c = \frac{\alpha^{2\alpha}(\tilde{p}^\alpha_+ + \tilde{p}^\alpha_-)}{\Gamma(1+\alpha)}$.*

Corollary 4.6 *Theorems 4.5, 4.6 and probabilistic representation for solutions of Cauchy problem for parabolic differential equations imply useful results on asymptotic behavior of the solution $u(t, x)$ and $v(t, x)$ of problems*

$$\frac{\partial u}{\partial t} = \frac{1}{2}a(x)\frac{\partial^2 u}{\partial x^2}, \qquad u(0, x) = \varphi(x)$$

and

$$\frac{\partial v}{\partial t} = \frac{1}{2}a(x)\frac{\partial^2 v}{\partial x^2} + f(x), \qquad v(0, x) = \varphi(x)$$

as $t \to \infty$. See [139] for details.

Remark 4.7 Making use of the reduction of the recurrent process with the generator $\frac{1}{2}\sigma^2(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx}$ to the process with the generator $\frac{1}{2}\sigma^2_1(y)\frac{d^2}{dy^2}$ (see Example 3.10), analogous results can be obtained for the null recurrent process with drift, see details in [139].

4.12 Arcsine Law and One Generalization

P. Levy in his brilliant paper [182] found the probability distribution for the occupation time of the set $\{x : x > 0\}$ by the Brownian motion $W(t)$. He found that for any $T > 0$, $0 \le x \le 1$,

$$\mathbf{P}\left\{\frac{1}{T}\int_0^T \mathbb{1}_{\{W(t)>0\}}(W(t))\,dt < x\right\} = \frac{2}{\pi}\arcsin\sqrt{x}. \tag{4.112}$$

Later it was proved that this distribution law (arcsine law) is also the limiting distribution for the number of positive sums of independent random variables with zero mean, see for instance [67].

We start from the proof (4.112) based on the Feynman–Kac formula (Lemma 3.7).

Theorem 4.7 *The distribution law of random variable*

$$\zeta_T = \frac{1}{T} \int_0^T \mathbb{1}_{\{W(t)>0\}}(W(t))\, dt$$

does not depend on T and is given by (4.112).

Proof It is well known (see, e.g. [254, Sect. 4.1]) that for standard Wiener process $W(t)$ the process $W_1(t) = \sqrt{T}\, W(t/T)$ is also standard Wiener for any $T > 0$. Therefore

$$\frac{1}{T} \int_0^T \mathbb{1}_{\{W_1(t)>0\}}(W_1(t))\, dt = \frac{1}{T} \int_0^T \mathbb{1}_{\{W_1(t)>0\}}(\sqrt{T}\, W(t/T))\, dt$$

$$= \frac{1}{T} \int_0^T \mathbb{1}_{\{W_1(t)>0\}}(W(t/T))\, dt$$

$$= \int_0^1 \mathbb{1}_{\{W(t_1)>0\}}(W(t_1))\, dt_1. \qquad (4.113)$$

Denote

$$u(t, x, s) = \mathbf{E} \exp\left\{ -s \int_0^t \mathbb{1}_{\{x+W(t_1)>0\}}(x + W(t_1))\, dt_1 \right\}.$$

Due to Lemma 3.7 this function is a solution of problem

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2} - s\mathbb{1}_{\{x>0\}}(x)u, \qquad u(0, x, s) = 1. \qquad (4.114)$$

It follows from (4.114) that the Laplace transform

$$v(\lambda, x, s) = \int_0^\infty e^{-\lambda t} u(t, x, s)\, dt$$

is a bounded solution of the equation

$$\lambda v(\lambda, x, s) - 1 = \frac{1}{2} \frac{d^2 v(\lambda, x, s)}{dx^2} - \mathbb{1}_{\{x>0\}}(x)sv(\lambda, x, s).$$

The last equation can be written at the equivalent form

$$\begin{aligned}
\frac{1}{2} \frac{d^2 v}{dx^2} - \lambda v &= -1, &&\text{as } x < 0, \\
\frac{1}{2} \frac{d^2 v}{dx^2} - (\lambda + s)v &= -1, &&\text{as } x > 0,
\end{aligned} \qquad (4.115)$$

with gluing conditions

$$v(\lambda, +0, s) = v(\lambda, -0, s); \qquad \frac{dv}{dx}(\lambda, +0, s) = \frac{dv}{dx}(\lambda, -0, s). \qquad (4.116)$$

It is easy to see that the unique bounded solution of problem (4.115), (4.116) has the explicit form:

$$v(\lambda, x, s) = \begin{cases} \dfrac{1}{\lambda} - \dfrac{s}{\lambda\sqrt{\lambda+s}(\sqrt{\lambda}+\sqrt{\lambda+s})}\, e^{\sqrt{2\lambda}x}, & \text{as } x \le 0, \\[3mm] \dfrac{1}{\lambda+s} + \dfrac{s}{(\lambda+s)\sqrt{\lambda}(\sqrt{\lambda}+\sqrt{\lambda+s})}\, e^{-\sqrt{2(\lambda+s)}x}, & \text{as } x > 0. \end{cases}$$

In particular,

$$v(\lambda, 0, s) = \int_0^\infty e^{-\lambda t} u(t, 0, s)\, dt = \frac{1}{\sqrt{\lambda(\lambda+s)}}. \qquad (4.117)$$

But $u(t, 0, s) = \mathbf{E}\exp\{-s\int_0^t \mathbb{1}_{\{x>0\}}(W(t_1))\, dt_1\}$, and due to (4.113)

$$u(t, 0, s) = \mathbf{E}\exp\{-st\zeta\} \qquad (4.118)$$

with $\zeta = \int_0^1 \mathbb{1}_{\{x>0\}}(W(t))\, dt$. It follows from (4.117) and (4.118) that

$$\frac{1}{\sqrt{\lambda(\lambda+s\zeta)}} \int_0^\infty e^{-\lambda t} \mathbf{E}e^{-st\zeta} = \mathbf{E}\int_0^\infty e^{-t(\lambda+s\zeta)}\, dt = \mathbf{E}\frac{1}{\lambda+s\zeta}.$$

Denote $\lambda/s = z$. Then we have the following expression for the Stieltjes transform of ζ:

$$\mathbf{E}\frac{1}{\zeta+z} = \frac{1}{\sqrt{z(z+1)}}. \qquad (4.119)$$

It follows from (4.119) (see details in [254, Sect. 4.1]) that

$$\mathbf{P}\{\zeta < x\} = \frac{2}{\pi}\arcsin\sqrt{x}, \quad \text{as } 0 \le x \le 1.$$

Assertion of theorem follows now from (4.113). $\square$

Theorem 4.8 [7] *Let $X(t)$ be a diffusion process with the generator $\frac{a(x)}{2}\frac{d^2}{dx^2}$, and $a(x) \ge a_0 > 0$. Denote $p(x) = 2a(x)^{-1}$ and assume that the following conditions are fulfilled.*

1. *For some constants p_+, p_- such that $p_+ + p_- > 0$*

$$\lim_{X\to\pm\infty} \frac{1}{X}\int_0^X p(x)\, dx = p_\pm. \qquad (4.120)$$

---

[7]This theorem was proved in [131], see also [174], [271] some results on limiting distribution of occupation time for the set $\{x > 0\}$.

2. *For piecewise continuous and bounded function $f(x)$, $x \in \mathbb{R}^1$, there exist constants f_+ and f_- such that $f_+ \neq f_-$ and*

$$\lim_{X \to \pm\infty} \frac{\int_0^X f(x)p(x)\,dx}{\int_0^X p(x)\,dx} = f_{\pm}. \tag{4.121}$$

Then the random variable

$$\eta_x(T) = \frac{\zeta_x(T) - f_-}{f_+ - f_-} = \frac{T^{-1}\int_0^T f(X^x(t))\,dt - f_-}{f_+ - f_-}$$

has the proper limit distribution as $T \to \infty$. For $p_+ = p_-$ this limit distribution is the arcsine law:

$$\lim_{T \to \infty} \mathbf{P}\{\eta_x(T) < z\} = \frac{2}{\pi} \arcsin\sqrt{z}, \quad 0 \leq z \leq 1.$$

For $p_+ \neq p_-$ the limit distribution coincides with distribution of a random variable δ such that for all $z > 0$, $A = \sqrt{p_+/p_-}$

$$\mathbf{E}\frac{1}{z + \delta} = \frac{\sqrt{1 + z} + A\sqrt{z}}{\sqrt{(1 + z)z}(\sqrt{z} + A\sqrt{1 + z})}. \tag{4.122}$$

Distribution of δ is uniquely determined by (4.122).

Remark 4.8 It follows from [272] the formula for the distribution law of δ:

$$\mathbf{P}\{\delta \leq x\} = \frac{2}{\pi} \arcsin\sqrt{\frac{x}{A^2 + (1 - A^2)x}}.$$

Proof Consider the random variable

$$\eta_{t,T}(x) = \frac{1}{T(f_+ - f_-)} \int_0^{tT} [f(X^x(\tau)) - f_-]\,d\tau.$$

It follows from Lemma 3.7 that the function

$$u_{s,T}(t, x) = \mathbf{E}\exp\{-s\eta_{t,T}(x)\} \tag{4.123}$$

is a solution of the problem

$$\frac{1}{T}\frac{\partial u_{s,T}}{\partial t} = \frac{1}{2}a(x)\frac{\partial^2 u_{s,T}}{\partial x^2} - \frac{s(f(x) - f_-)}{T(f_+ - f_-)}u_{s,T}; \qquad u_{s,T}(0, x) = 1. \tag{4.124}$$

Consider the new function

$$v_{s,T}(t, x) = u_{s,T}(t, x\sqrt{T}). \tag{4.125}$$

Then it is easy to see from (4.124) that $v_{s,T}(t,x)$ is a solution of the problem

$$\frac{\partial v_{s,T}}{\partial t} = \frac{1}{2} a(x\sqrt{T}) \frac{\partial^2 v_{s,T}}{\partial x^2} - \frac{s(f(x\sqrt{T}) - f_-)}{f_+ - f_-} v_{s,T}; \qquad v_{s,T}(0,x) = 1.$$

Similar to the proof of Theorem 4.7 we consider once more the Laplace transform

$$w_{s,\lambda}^{(T)}(x) = \int_0^\infty e^{-\lambda t} v_{s,T}(t,x)\, dt. \qquad (4.126)$$

Then, from (4.125) and (4.126), we have the ODE for $w_{s,\lambda}^{(T)}(x)$

$$\lambda w_{s,\lambda}^{(T)}(x) - 1 = \frac{1}{2} a(x\sqrt{T}) \frac{d^2 w_{s,\lambda}^{(T)}(x)}{dx^2} - \frac{s(f(x\sqrt{T}) - f_-)}{f_+ - f_-} w_{s,\lambda}^{(T)}(x).$$

This equation is equivalent to the equation

$$\frac{d^2 w_{s,\lambda}^{(T)}(x)}{dx^2} = p(x\sqrt{T})\left\{ \left[\lambda + \frac{s(f(x\sqrt{T}) - f_-)}{f_+ - f_-} \right] w_{s,\lambda}^{(T)}(x) - 1 \right\} \qquad (4.127)$$

and $w_{s,\lambda}^{(T)}(x)$ is a bounded for all $\lambda > 0$ and positive solution of this equation. It follows from (4.120), (4.121), (4.127) and boundedness $w_{s,\lambda}$ that the second-order derivative of $w_{s,\lambda}^{(T)}(x)$ is also bounded. Therefore, the families of functions $w_{s,\lambda}^{(T)}(x)$, $dw_{s,\lambda}^{(T)}(x)/dx$ are compact families with respect to T in any compact set K ($x \in K$). Choose some sequence $T_n \to \infty$ so that uniformly in K

$$\lim_{n\to\infty} w_{s,\lambda}^{(T_n)}(x) = w_{s,\lambda}^{(0)}(x), \qquad \lim_{n\to\infty} \frac{d}{dx} w_{s,\lambda}^{(T_n)}(x) = \frac{d}{dx} w_{s,\lambda}^{(0)}(x).$$

Integrating (4.127) in x for $T = T_n$ and again using Assumptions (4.120), (4.121), we can conclude that for any $x_1, x_2 \in K$

$$\frac{dw_{s,\lambda}^{(0)}(x_2)}{dx} - \frac{dw_{s,\lambda}^{(0)}(x_1)}{dx}$$
$$= \int_{x_1}^{x_2} \{[\lambda(p_+ \mathbb{1}_{\{x>0\}}(x) + p_- \mathbb{1}_{\{x<0\}}(x)) + sp_+ \mathbb{1}_{\{x>0\}}(x)] w_{s,\lambda}^{(0)}(x)$$
$$- p_+ \mathbb{1}_{\{x>0\}}(x) - p_- \mathbb{1}_{\{x<0\}}(x)\} dx.$$

In other words, the function $w_{s,\lambda}^{(0)}(x)$ is a bounded solution of the problem

$$\begin{aligned}
\frac{d^2 w_{s,\lambda}^{(0)}(x)}{dx^2} &= p_+(\lambda + s) w_{s,\lambda}^{(0)}(x) - p_+, \quad x > 0, \\
\frac{d^2 w_{s,\lambda}^{(0)}(x)}{dx^2} &= \lambda p_- w_{s,\lambda}^{(0)}(x) - p_-, \qquad x < 0,
\end{aligned} \qquad (4.128)$$

with the gluing conditions

$$w_{s,\lambda}^{(0)}(-0) = w_{s,\lambda}^{(0)}(+0); \qquad \frac{d}{dx} w_{s,\lambda}^{(0)}(+0) = \frac{d}{dx} w_{s,\lambda}^{(0)}(-0). \tag{4.129}$$

The solution of this problem is unique and can be written in the explicit form analogously the solution of problem (4.115), (4.116):

$$w_{s,\lambda}^{(0)}(x) = c_1(\lambda, s) \exp(-\sqrt{p_+(\lambda + s)}x) + \frac{1}{\lambda + s}, \qquad x \ge 0,$$

$$w_{s,\lambda}^{(0)}(x) = -c_1(\lambda, s)\sqrt{\frac{p_+}{p_-}\left(1 + \frac{s}{\lambda}\right)} \exp(-\sqrt{\lambda p_-}x) + \frac{1}{\lambda}, \quad x < 0, \tag{4.130}$$

$$c_1(\lambda, s) = \frac{s}{\lambda(\lambda + s)(1 + \sqrt{p_+(1 + s/\lambda)/p_-})}.$$

It follows from the uniqueness of solution of the problem (4.128), (4.129) that there exists limit

$$\lim_{T \to \infty} w_{s,\lambda}^{(T)}(x) = w_{s,\lambda}^{(0)}(x). \tag{4.131}$$

The function $v_{s,T}(t, x)$ is uniformly in T continuous with respect to t. Moreover, from the probabilistic meaning of $v_{s,T}(t, x)$ (see (4.123), (4.125)) follows the inequality

$$|v_{s,T}(t + h, x) - v_{s,T}(t, x)| < Ch \tag{4.132}$$

with independent of T constant C. It follows from (4.126), (4.131) and (4.132) that

$$\lim_{T \to \infty} v_{s,T}(t, x) = v_s^{(0)}(t, x) \tag{4.133}$$

and

$$w_{s,\lambda}^{(0)}(x) = \int_0^\infty e^{-\lambda t} v_s^{(0)}(t, s) \, dt. \tag{4.134}$$

It is clear from (4.134) and (4.128) that $v_s^{(0)}(t, x)$ is a solution of the problem

$$\frac{\partial v_s^{(0)}}{\partial t} = \frac{1}{p_+} \frac{\partial^2 v_s^{(0)}}{\partial x^2} - s v_s^{(0)}, \quad x > 0,$$

$$\frac{\partial v_s^{(0)}}{\partial t} = \frac{1}{p_-} \frac{\partial^2 v_s^{(0)}}{\partial x^2}, \quad x < 0,$$

$$v_s^{(0)}(0, x) = 1, \qquad v_s^{(0)}(t, +0) = v_s^{(0)}(t, -0),$$

$$\frac{\partial}{\partial x} v_s^{(0)}(t, +0) = \frac{\partial}{\partial x} v_s^{(0)}(t, -0).$$

So, applying Feynman–Kac lemma again, we have

$$v_s^{(0)}(t, x) = \mathbf{E} \exp\left\{-\frac{s}{T} \int_0^{tT} \mathbb{1}_{\{x > 0\}}(X_0^{x\sqrt{T}}(\tau)) \, d\tau\right\}, \tag{4.135}$$

where $X_0^x(t)$ is a diffusion process with the generator

$$\left(\frac{1}{p_+}\mathbb{1}_{(x>0)}(x) + \frac{1}{p_-}\mathbb{1}_{(x<0)}(x)\right)\frac{d^2}{dx^2}$$

and $X_0^x(0) = x$. In particular it follows from (4.135) that the distribution law of the random variable

$$\delta(T) = \frac{1}{T}\int_0^T \mathbb{1}_{(x>0)}(X_0^0(t))\,dt$$

does not depend on T and coincides with arcsine law for $p_+ = p_-$.

We can conclude from (4.123), (4.125) and (4.133) that the distribution law of $\eta_x(T)$ converges to the distribution law of $\delta(1)$. So in order to finish the proof we need only to check the characterization (4.122) of the distribution law $\delta = \delta(1)$ for $p_+ \neq p_-$ and the uniqueness of reconstruction of this distribution by (4.122). We have from (4.135)

$$v_s^{(0)}(t,0) = \mathbf{E}\exp\left\{-s\int_0^t \mathbb{1}_{\{x>0\}}(X_0^0(\tau))\,d\tau\right\} = \mathbf{E}\exp\{-st\delta(1)\}.$$

Therefore,

$$w_{s,\lambda}^{(0)}(0) = \int_0^\infty e^{-\lambda t}v_s^{(0)}(t,0)\,dt = \int_0^\infty e^{-\lambda t}\mathbf{E}\exp\{-st\delta(1)\}\,dt = \mathbf{E}\frac{1}{\lambda + s\delta(1)}. \tag{4.136}$$

On the other hand we have from (4.130)

$$w_{s,\lambda}^{(0)}(0) = \frac{s + \lambda + \lambda\sqrt{\frac{p_+}{p_-}(1+\frac{s}{\lambda})}}{\lambda(\lambda+s)(1+\sqrt{\frac{p_+}{p_-}(1+\frac{s}{\lambda})})}$$

$$= \frac{\sqrt{1+\frac{\lambda}{s}} + \sqrt{\frac{p_+}{p_-}}\sqrt{\frac{\lambda}{s}}}{s\sqrt{(1+\frac{s}{\lambda})\frac{\lambda}{s}}(\sqrt{\frac{\lambda}{s}} + \sqrt{\frac{p_+}{p_-}(1+\frac{\lambda}{s})})}. \tag{4.137}$$

Denote $A = (p_+/p_-)^{1/2}$, $\lambda/s = z$. Then we have from (4.136) and (4.137)

$$\mathbf{E}\frac{1}{z+\delta(1)} = \frac{\sqrt{1+z} + A\sqrt{z}}{\sqrt{(1+z)z}(\sqrt{z} + A\sqrt{1+z})}. \tag{4.138}$$

The characterization (4.122) is proved. From the Taylor expansion in the neighborhood of $z = 1$ (for instance), we can find all moments of the random variable $\eta = (\delta(1)+1)^{-1}$. Uniqueness of the construction probability distribution by (4.122) follows from the above argument and boundedness of η. $\qquad\square$

Remark 4.9 With $r = z^{-1}$, (4.138) can be rewritten in the form

$$\mathbf{E}(1 + r\delta(1))^{-1} = \frac{A + \sqrt{1+r}}{\sqrt{1+r}(1 + A\sqrt{1+r})}.$$

All moments of $\delta(1)$ can be found from the Taylor expansion at $r = 0$. For instance

$$\mathbf{E}\delta(1) = \frac{A}{A+1} = \frac{\sqrt{p_+}}{\sqrt{p_+} + \sqrt{p_-}},$$

$$\operatorname{var}\delta(1) = \frac{A}{2(A+1)^2} = \frac{\sqrt{p_+ p_-}}{2(\sqrt{p_+} + \sqrt{p_-})^2}. \tag{4.139}$$

It follows from (4.139) that

$$\lim_{T \to \infty} T^{-1} \int_0^T \mathbf{P}\{X^x(t) > 0\}\, dt = \frac{\sqrt{p_+}}{\sqrt{p_+} + \sqrt{p_-}}$$

when assumption (4.120) is valid for a Markov process with the generator $\frac{1}{2}a(x)\frac{d^2}{dx^2}$.

We describe now some further results on the generalized arcsine law for the one-dimensional null recurrent diffusion process in the canonic scale (with the generator $\frac{1}{2}a(x)\frac{d^2}{dx^2}$).

J. Lamperti [174] introduced more general two-parameter class of random variables $Y_{\alpha,A}$, $0 \le \alpha \le 1$, $A \ge 0$ with values in [0, 1] and the Stieltjes transform given by

$$\mathbf{E}\frac{1}{z + Y_{\alpha,A}} = \frac{z^{\alpha-1} + A(1+z)^{\alpha-1}}{z^\alpha + A(1+z)^\alpha}, \quad 0 < z < 1. \tag{4.140}$$

He noticed that for $0 < \alpha < 1$, $A > 0$ the probability distribution $F_{\alpha,A}$ of $Y_{\alpha,A}$ has the density

$$f_{\alpha,A}(x) = \frac{\sin\alpha\pi}{\pi} \frac{Ax^{\alpha-1}(1-x)^{\alpha-1}}{A^2(1-x)^{2\alpha} + 2Ax^\alpha(1-x)^\alpha\cos\alpha\pi + x^{2\alpha}} \mathbb{1}_{[0,1]}(x). \tag{4.141}$$

It was proved in [174] that the distributions $F_{\alpha,A}$ are all possible limit distributions for the occupation time of some sets for some stochastic processes with discrete times (even not necessarily Markovian). Necessary and sufficient conditions for convergence to the distribution $F_{\alpha,A}$ with given A, α are also given there.

S. Watanabe in [271] proved the necessary and sufficient conditions for convergence of the occupation time $\tau_T = \int_0^T \mathbb{1}_{\{x>0\}}(X^x(t))\, dt$ of the set $\{x > 0\}$ for the one-dimensional (generalized) diffusion process in the canonical scale. For non-generalized process in the canonical scale (the process with generator $\frac{1}{2}a(x)\frac{d^2}{dx^2}$) Watanabe's conditions for

$$\frac{\tau_T}{T} \xrightarrow{\text{distr.}} Y_{\alpha,A}, \quad 0 < \alpha < 1,$$

can be written in the form

$$\left| \int_0^{\pm x} p(y)\, dy \right| = |x|^{\beta+1} K_{\pm}(|x|) \tag{4.142}$$

with slowly varying in the Karamata sense functions $K_{\pm}(|x|)$, satisfying the condition

$$\lim_{x \to \infty} \frac{K_+(x)}{K_-(x)} = A^{1/\alpha}. \tag{4.143}$$

(Here and below $\alpha(2+\beta) = 1$.)

The sufficient conditions guaranteeing the convergence in distribution of the integral functional $Y_{\alpha, A}$ for one-dimensional null recurrent diffusion in the canonical scale were found in [132].

The following generalization of Theorem 4.8 was proven there.

Theorem 4.9 *Suppose that the following conditions are valid*:

1. $\lim_{X \to \pm\infty} \frac{1}{X} \int_0^X |x|^{-\beta} p(x)\, dx = p_{\pm}$ *for some constants* $\beta > -1$, $p_+ + p_- > 0$.
2. *For any* $\varepsilon > 0$, $\sup_{|x| > \varepsilon}\{|x|^{-\beta} p(x)\} < \infty$.
3. $f(x)$ *is a piecewise continuous bounded function and*

$$\lim_{X \to \pm\infty} \frac{\int_0^X f(x)|x|^{-\beta} p(x)\, dx}{\int_0^X |x|^{-\beta} p(x)\, dx} = f_{\pm},$$

where the constants f_+ *and* f_- *satisfy the condition* $f_+ - f_- \neq 0$.

Denote $A = (p_+/p_-)^{\alpha}$. *Then*

$$\lim_{T \to \infty} \mathbf{P}\left\{ \frac{\int_0^T f(X^x(t))\, dt - f_-}{T(f_+ - f_-)} < x \right\} = F_{\alpha, A}(x).$$

Proof of this theorem is analogous to the proof of Theorem 4.8, see details in [132].

Chapter 5
Stability of Stochastic Differential Equations

5.1 Statement of the Problem

In Chap. 1 we studied problems of stability under random perturbations of the parameters. We noted there that no significant results can be expected unless the random perturbations possess sufficiently favorable mixing properties. Fortunately, in practical applications one may often assume that the "noise" has a "short memory interval." The natural limiting case of such noise is of course white noise. Thus it is very important to study the stability of solutions of Itô equations since this is equivalent to the study of stability of systems perturbed by white noise.

Any result concerning conditions for the stability of stochastic systems is apparently itself "stable", in the sense that it carries over to systems perturbed by noise which is "almost white" in a well-defined sense. Some relevant results were obtained by the author in [124]. However, as yet, no complete investigation has been made of conditions under which the stability (instability) of a system of Itô's equations implies the stability (instability) of a "similar" system.

Below we present a theory of stability for Itô equations. The main stimulus for setting up this theory was the extremely suggestive paper of Kac and Krasovskii [111]. These authors investigate the stability of the solution $X(t) = 0$ of the equation $dX/dt = f(X, t, Y(t))$, where $Y(t)$ is a time-homogeneous Markov chain with finitely many states. They solve the problem in terms of Lyapunov functions, but instead of the derivative d^0V/dt along the sample path they consider, roughly speaking, the expectation LV of this derivative. The paper also contains important results, amenable to generalization in various directions, concerning the stability of linear systems and stability in the first approximation.

In my paper [118] I remarked that by suitable interpretation these results yield information about the stability of the invariant set (in this case, the hyperplane $x = 0$) of the multi-dimensional Markov process $(X(t), Y(t))$. This interpretation will be given in Sect. 7.6 below.

A brief survey of Chaps. 5–7 follows.

Consider the system

$$dX(t) = b(t, X)\, dt + \sum_{r=1}^{k} \sigma_r(t, X)\, d\xi_r(t). \tag{5.1}$$

We shall assume that $X(t)$, $b(t, x)$ and $\sigma_r(t, x)$ are vectors in $\mathbb{R}^l$, and $\xi_r(t)$ are independent Wiener processes. We assume moreover that the coefficients of (5.1) satisfy the following Lipschitz condition in every domain which is bounded in x, i.e.,

$$\sum_{r=1}^{k} |\sigma_r(t, x) - \sigma_r(t, y)| + |b(t, x) - b(t, y)| < B|x - y|. \tag{5.2}$$

In some cases the Lipschitz constant B will be assumed to be independent of the domain, i.e. inequality (5.2) will be assumed to hold throughout $E = \{t > 0\} \times \mathbb{R}^l$. We shall limit ourselves to conditions for stability of the trivial solution $X(t) = 0$. Accordingly, we assume that

$$b(t, 0) \equiv 0, \qquad \sigma_r(t, 0) \equiv 0. \tag{5.3}$$

In the present chapter, the solution of stability problems for systems of SDEs will be given primarily in the same terms as above. That is, we shall use the same methods, as employed for the qualitative behavior of solutions of such systems in Chap. 3.

Theorems 5.3, 5.5 and 5.7 proved below are natural generalizations to stochastic systems of Lyapunov's second method. One feature of these theorems should be mentioned. They all require that the Lyapunov function is sufficiently smooth in t and x in a neighborhood of $x = 0$, except possibly at the point $x = 0$ itself. Unlike for a deterministic system, for a stochastic system there often does not even exist a Lyapunov function which is smooth at the origin. This will be clear from the example at the end of Sect. 5.3.

In this connection, we introduce the following definition.

Let U be a domain with closure $\widetilde{U}$ in the space $E = I \times \mathbb{R}^l$, and set $U^\varepsilon(0) = \{(t, x) : |x| < \varepsilon\}$. We shall say that a function $V(t, x)$ is in class $\mathbf{C}_2^0(U)$ ($V(t, x) \in \mathbf{C}_2^0(U)$) if it is twice continuously differentiable with respect to x and continuously differentiable with respect to t throughout U, except possibly for the set $x = 0$, and continuous in the closed set $\widetilde{U} \setminus U^\varepsilon(0)$ for any $\varepsilon > 0$.

As in Chap. 1, we shall consider the stability of the moment of different orders (p-stability) and stability in probability. However, whereas in Chap. 1 we were able to derive conditions only for weak stability in probability (see Sect. 1.5), here we shall consider stability in probability in a stronger sense. To be precise, we shall present conditions under which not only does $|X(t)|$ tend to zero in probability uniformly in t, but also $\sup_{t>0} |X(t)|$ tends to zero in probability as $|X(0)| \to 0$. With this definition, the stability or instability of the equilibrium position is determined by the behavior of the coefficients of the equation only in a neighborhood of the

equilibrium position. It is therefore natural to expect that for a broad range of cases the full system (5.1) will be stable provided the first-approximation system

$$dX(t) = \frac{\partial b(t,0)}{\partial x} X(t) \, dt + \sum_{r=1}^{k} \frac{\partial \sigma_r(t,0)}{\partial x} X(t) \, d\xi_r(t) \qquad (5.4)$$

is stable.

The question as to when stability of the system (5.4) implies that of (5.1) will be answered in Chap. 7. We shall prove that if $\partial b/\partial x$ and $\partial \sigma_r/\partial x$ are independent of t, then it suffices that (5.4) is asymptotically stable in probability. If the coefficients depend on time, one must also assume that some uniformity in time condition is valid.

Of major importance in this connection is the problem of stability of linear stochastic systems. In the general case, this problem is rather difficult. We shall solve it in Chap. 6, though not in a very effective way. We shall prove there that the stability or instability of a linear stochastic system with time-independent coefficients is determined by the sign of the expectation of a certain random variable, given the stationary distribution of a certain Markov process on the l-dimensional sphere. It is shown that this expectation is equal to the $\lim_{t\to\infty} \frac{\ln |X(t)|}{t}$ which is known as Lyapunov exponent for the linear system. For $l = 2$ the density of this distribution can be computed by quadratures, so the Lyapunov exponent can be found in closed form.

The question of instability conditions is even more complicated. The analogs of the instability theorems of Lyapunov and Chetaev do not hold in for stochastic systems. Roughly speaking, the reason for this is that the sample paths of stochastic systems may leave the "instability set" because of purely random forces. This is made particularly clear by the second example in Sect. 7.3. In this example, the unstable deterministic system $dx_1/dt = x_1$, $dx_2/dt = -x_2$ is "impaired" by the addition of a small drift and a small diffusion

$$\left. \begin{aligned} dX_1(t) &= (X_1 + b(X_1, X_2)) \, dt + \sigma(X_1, X_2) \, d\xi_1(t), \\ dX_2(t) &= -X_2 \, dt + \sigma(X_1, X_2) \, d\xi_2(t). \end{aligned} \right\} \qquad (5.5)$$

In this case, for any small number $\varepsilon > 0$, the functions b and σ may be chosen so that

$$|b(x_1, x_2)| + |\sigma(x_1, x_2)| < \varepsilon |x|,$$

and the system (5.5) is asymptotically stable in the large. It is well known that this phenomenon is impossible if $\sigma \equiv 0$.

In Chap. 5 (Theorem 5.6) we shall present a sufficient condition for instability which at first glance is relevant only for very special cases. However, it will follow from the discussion in Chaps. 6 and 7 that if the diffusion of the system is nondegenerate (in a fairly weak sense), then it is often possible to construct a function satisfying the assumptions of Theorem 5.6.

On this basis, we shall consider in Chap. 7 the problem of instability in the first approximation. In the same chapter we shall also consider the problem of stability

under damped random perturbations and some applications to statistics. At the end of Chap. 7 we shall discuss the possibility of generalizing the results to a wider class of processes.

After publication of the first Russian edition of this book in 1969 the huge literature was devoted to the problem of evaluation the Lyapunov exponent for the linear stochastic systems. The concepts of moment Lyapunov exponents and stability index were proposed and well studied. Some results at this area are stated in Appendix B to this edition.

5.2 Some Auxiliary Results

This chapter will make systematic use of properties of martingales and supermartingales. The following definitions of these important classes of stochastic process are due to Doob.

Let $(\Omega, \mathfrak{A}, \mathbf{P})$ be a probability space, $\mathcal{M}_t \subset \mathfrak{A}$ a family of σ-algebras of events in Ω, defined for each $t \geq 0$, such that $\mathcal{M}_s \subset \mathcal{M}_t$ for $s < t$. Let $y(t, \omega)$, $t \geq 0$, be a stochastic process with finite expectation $\mathbf{E} y(t, \omega)$, such that $y(t, \omega) = y(t)$ is an $\mathcal{M}_t$-measurable random variable for each t. The family $(y(t, \omega), \mathcal{M}_t)$ is called a *supermartingale* if for any $s < t$

$$\mathbf{E}(y(t) \mid \mathcal{M}_s) \leq y(s) \quad (\textbf{P}\text{-a.s.}). \tag{5.6}$$

If we replace the inequality sign in (5.6) by equality, we get the definition of a *martingale*.

The following examples of martingales and supermartingales are important for the sequel.

Example 5.1 The Wiener process $\xi(t)$ is a martingale with respect to the system of σ-algebras $\widetilde{\mathcal{N}}_t$, since

$$\mathbf{E}(\xi(t) \mid \widetilde{\mathcal{N}}_f) = \mathbf{E}([\xi(s) + (\xi(t) - \xi(s))] \mid \widetilde{\mathcal{N}}_s) = \xi(s) \quad (\text{a.s.}).$$

An analogous argument shows that the more general process

$$y(t) = \int_0^t \sigma(s) \, d\xi(s)$$

with $\widetilde{\mathcal{N}}_s$-measurable process $\sigma(s) = \sigma(s, \omega)$ is also a martingale.

Example 5.2 Let $V(t, x)$ be a function twice continuously differentiable with respect to x and continuously differentiable with respect to t in $I \times U$, where $U \subset \mathbb{R}^l$ is a bounded closed domain. Suppose that in this domain

$$LV(t, x) = \frac{\partial V}{\partial t} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(t, x) \frac{\partial^2 V}{\partial x_i \partial x_j} + \sum_{i=1}^{l} b_i(t, x) \frac{\partial V}{\partial x_i}$$

$$\equiv \frac{\partial V}{\partial t} + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t, x), \frac{\partial}{\partial x} \right)^2 V + \left(b(t, x), \frac{\partial}{\partial x} \right) V \leq 0$$

holds. Set $\tau(t) = \min(\tau, t)$, where τ is the first exit time from U at the sample path of the process $X(t)$ defined by (5.1). Then the process $y(t) = V(\tau(t), X(\tau(t)))$ is a supermartingale with respect to the system $\widetilde{\mathcal{N}}_t$. In fact, under the above assumptions it follows from Lemma 3.2 that

$$\mathbf{E}[V(\tau(t), X(\tau(t))) \mid \mathcal{N}_s] \leq V(s, X(s)) \quad \text{(a.s.)}.$$

Hence condition (5.6) is satisfied for almost all paths such that $\tau > s$, and consequently $X(\tau(s)) = y(s) \in U$. For almost all paths such that $\tau \leq s$ we have $\mathbf{E}(y(t) \mid \widetilde{\mathcal{N}}_s) = y(s)$, since then $\tau(s) = \tau(t) = \tau$. If $LV \leq 0$ for all $x \in \mathbb{R}^l$, $t \geq 0$, and $\mathbf{E}^{s,x} V(t, X(t))$ exists, an analogous argument shows that the process $V(t, X(t))$ is also a supermartingale.

These properties generally fail to hold if the condition $LV \leq 0$ is not satisfied on some set (even at one point). In certain cases, however, the supermartingale property of the process remains valid even when this happens. Let us call the random variable $\tau^\Gamma = \inf\{t : X(t) \in \Gamma\}$ the first time at which the set Γ is reached. A closed set Γ is said to be *inaccessible* to a process $X(t)$ if $\mathbf{P}\{\tau^\Gamma < \infty\} = 0$. Since the sample paths of the process are continuous, a set Γ is inaccessible if and only if

$$\mathbf{P}\{\tau(U_\delta(\Gamma)) \to \infty \text{ as } \delta \to 0\} = 1.$$

Here $U_\delta(\Gamma)$ is the δ-neighborhood of the set Γ.

Lemma 5.1 *Let $V(t, x)$ be a function twice continuously differentiable with respect to x, continuously differentiable with respect to t on the set $I \times \{U \setminus \Gamma\}$ and bounded in $I \times U$, where U is a bounded domain in $\mathbb{R}^l$ and $\Gamma \subset U$ is a set inaccessible to the process $X(t)$ defined by (5.1). Assume that $LV \leq 0$ on the set $I \times (U \setminus \Gamma)$. Then the process $V(\tau_U(t), X(\tau_U(t)))$ is a supermartingale.*

Proof Let $\tau(U, \delta)$ denote the first exit time from the set $U \setminus U_\delta(\Gamma)$, $\tau_{U,\delta}(t) = \min(\tau(U, \delta), t)$. Since Γ is inaccessible, it follows that for all t, we have

$$\tau_{U,\delta}(t) \to \tau_U(t) \quad \text{(a.s.)} \tag{5.7}$$

as $\delta \to 0$. On the other hand, it is clear from Example 5.2 that

$$\mathbf{E}(V(\tau_{U,\delta}(t), X(\tau_{U,\delta}(t))) \mid \widetilde{\mathcal{N}}_s) \leq V(\tau_{U,\delta}(s), X(\tau_{U,\delta}(s))) \quad \text{(a.s.)}.$$

Letting $\delta \to 0$ in this inequality and using (5.7) and the fact that V is bounded, we get the required assertion. $\square$

Lemma 5.2 [1] *Let the coefficients b and σ_r of* (5.1) *satisfy condition* (5.3). *Assume further that condition* (5.2) *holds throughout $E = I \times \mathbb{R}^l$. Then for any real β, $t \geq s$, $x \neq 0$,*

$$\mathbf{E}|X^{s,x}(t)|^\beta \leq |x|^\beta \exp\{k(t-s)\}, \tag{5.8}$$

where k is a constant depending only on β and the constant B of (5.2).

Proof The function $V(x) = |x|^\beta$ is twice continuously differentiable in the domain $|x| > \delta$ for any $\delta > 0$. Applying Itô's formula (3.30) in this domain, we get for $Y^{s,x}(t) = |X^{s,x}(t)|^\beta$ the formula

$$Y^{s,x}(\tau_\delta(t))$$
$$= Y^{s,x}(s) + \beta \int_s^{\tau_\delta(t)} |X^{s,x}(u)|^{\beta-2} \Bigg[(b(u, X^{s,x}(u)), X^{s,x}(u))\, du$$
$$+ \frac{1}{2} \sum_{i=1}^l a_{ii}(u, X^{s,x}(u))\, du + \sum_{r=1}^k (\sigma_r(u, X^{s,x}(u)), X^{s,x}(u))\, d\xi_r(u) \Bigg]$$
$$+ \frac{1}{2}\beta(\beta-2) \int_s^{\tau_\delta(t)} |X^{s,x}(u)|^{\beta-4}(A(u, X^{s,x}(u))X^{s,x}(u), X^{s,x}(u))\, du, \tag{5.9}$$

where τ_δ denotes the first exit time from the set $|x| > \delta$ and $\tau_\delta(t) = \min(\tau_\delta, t)$. It is obvious that the random variable $Y^{s,x}(\tau_\delta(t))$ has an expectation. (If $\beta \leq 0$, this follows from the fact that it is bounded, and if $\beta > 0$ we may conclude this from Example 3.1.) Calculating expectations in (5.9) and using (5.2) and (5.3), we easily obtain that

$$\mathbf{E}Y^{s,x}(\tau_\delta(t)) \leq |x|^\beta + k\mathbf{E} \int_s^{\tau_\delta(t)} Y^{s,x}(u)\, du \tag{5.10}$$

for some $k = k(\beta, B, l)$. Since $\tau_\delta(u) = u$ for $u < \tau_\delta(t)$, it follows from (5.10) that

$$\mathbf{E}Y^{s,x}(\tau_\delta(t)) \leq |x|^\beta + k\mathbf{E} \int_s^{\tau_\delta(t)} Y^{s,x}(\tau_\delta(u))\, du$$
$$\leq |x|^\beta + k \int_s^t \mathbf{E}Y^{s,x}(\tau_\delta(u))\, du.$$

Applying the Gronwall–Bellman lemma to this inequality, we get the estimate

$$\mathbf{E}|X^{s,x}(\tau_\delta(t))|^\beta \leq |x|^\beta \exp\{k(t-s)\}. \tag{5.11}$$

---

[1] See [221].

Setting $\beta = -1$ in (5.11) and using Chebyshev's inequality, we get

$$\mathbf{P}^{s,x}\{\tau_\delta(t) < t\} < \frac{\delta}{|x|}\,e^{k(t-s)}.$$

This implies

$$\mathbf{P}^{s,x}\{\tau_\delta < t\} \to 0 \quad \text{as } \delta \to 0 \tag{5.12}$$

for every $s < t$. Letting $\delta \to 0$ in (5.11) and using (5.12), we get (5.8). $\square$

Remark 5.1 It follows from (5.12) that under the assumptions of the lemma the point $x = 0$ is inaccessible to the process $X^{s,x}(t)$. For this assertion to hold it suffices that condition (5.2) is satisfied in every cylinder $I \times K$, where $K \subset \mathbb{R}^\ell$ is compact and that $X(t)$ is regular. This is intuitively obvious (and easily proved rigorously), since whether the path of a regular process can hit $x = 0$ depends only on the behavior of the coefficients of the equation in the neighborhood of that point. For further reference, we state this result as a lemma.

Lemma 5.3 *Suppose that the coefficients of* (5.1) *satisfy* (5.3), *condition* (5.2) *holds in every domain bounded with respect to* x, *and the process* $X^{s,x_0}(t)$ *is regular. Then the point* $x = 0$ *is inaccessible to any sample path of the process if* $x_0 \neq 0$.

Lemmas 5.1 and 5.3 imply

Lemma 5.4 *Let* $V(t, x)$ *be a function in class* $\mathbf{C}_2^0((t > 0) \times U)$, *bounded in the domain* $(t > 0) \times U$, *where* U *is a neighborhood of the origin, and suppose that* $LV(t, x) \leq 0$ *in this domain. Then the process* $V(\tau_U(t), X(\tau_U(t)))$ *is a supermartingale, so that*

$$\mathbf{E}V(\tau_U(t), X^{s,x}(\tau_U(t))) \leq V(s, x)$$

for $x \in U$.

Remark 5.2 By virtue of (5.12) we have

$$\mathbf{P}^{s,x}\{\tau_\delta(t) \to t \text{ as } \delta \to 0\} = 1,$$

and hence, letting $\delta \to 0$ in (5.9), we see that Itô's formula (3.30) is applicable to the function $|x|^\beta$ on the whole of $\mathbb{R}^l$, despite the fact that if $\beta < 2$, this function does not satisfy the assumption of Theorem 3.3 at zero. This conclusion holds true for any function $V(t, x) \in \mathbf{C}_2^0(E)$ such that $0 \leq V(t, x) < k|x|^\beta$.

Our subsequent applications of martingale theory to stability problems are based on the following theorem which we give here without proof.

Theorem 5.1 [56] *If* $(y(t, \omega), \mathcal{M}_t, t \geq 0)$ *is a positive supermartingale, then the limit* $y_\infty = \lim_{t\to\infty} y(t, \omega)$ *almost surely exists and is finite. Moreover* $\mathbf{E}y_\infty = \lim_{t\to\infty} \mathbf{E}y(t, \omega)$.

In Chap. 6 we shall need also the following

Theorem 5.2 [56] *If $(y(t, \omega), \mathcal{M}_t, t \geq 0)$ is a continuous a.s. martingale, then for any $k > 0$, $p \geq 1$*

$$\mathbf{P}\left\{ \sup_{t_0 \leq t \leq T} |y(t, \omega)| > k \right\} \leq \frac{E|y(T, \omega)|^p}{k^p}.$$

5.3 Stability in Probability

A solution $X(t, \omega) \equiv 0$ of (5.1) is said to be *stable in probability* for $t \geq 0$ if for any $s \geq 0$ and $\varepsilon > 0$

$$\lim_{x \to 0} \mathbf{P}\left\{ \sup_{t > s} |X^{s,x}(t)| > \varepsilon \right\} = 0.$$

This definition is considerably stronger[2] than (1.63). It says that the sample path of the process issuing from a point x at time s will always remain within any pre-scribed neighborhood of the origin with probability tending to one as $x \to 0$. The importance of the definition will be clear from Theorem 5.3 below.

Before we state the theorem, we recall that a function $V(t, x)$ is said to be *positive definite* (in Lyapunov's sense) in a neighborhood of the set $x = 0$ if $V(t, 0) = 0$ and in this neighborhood $V(t, x) > W(x)$, where $W(x) > 0$ for $x \neq 0$ and continuous.

Theorem 5.3 is analogous to the well-known theorem of Lyapunov for deterministic systems. For nondegenerate processes, it was first proved by the author in [118]. Subsequently it has been generalized in various directions.[3]

Theorem 5.3 *Let $\{t > 0\} \times U = U_1$ be a domain containing the line $x = 0$, and assume there exists a function $V(t, x) \in \mathbf{C}_2^0(U_1)$ which is positive definite in Lyapunov's sense and satisfies*

$$LV = \frac{\partial V}{\partial t} + \sum_{i=1}^{l} b_i(t, x) \frac{\partial V}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(t, x) \frac{\partial^2 V}{\partial x_i \partial x_j} \leq 0$$

for $x \neq 0$. Then the trivial solution of (5.1) is stable in probability.

Proof Let r be a number such that the r-neighborhood U_r of the point $x = 0$ is contained in U together with its boundary. We set $V_r = \inf_{x \in U \setminus U_r} V(t, x)$ ($V_r > 0$ by assumption). By Lemma 5.4, we have

$$\mathbf{E}V(\tau_{U_r}(t), X^{s,x}(\tau_{U_r}(t))) \leq V(s, x)$$

---

[2]The relation between this definition and definition (1.63) is discussed in Sect. 6.11.

[3]Gikhman [89, 90] has given another proof, making no use of nondegeneracy, but using only Lyapunov functions which are smooth at zero. A similar result was obtained by Kushner in [166].

for $|x| < r$. Using this and Chebyshev's inequality, we get

$$\mathbf{P}\left\{\sup_{s \leq u \leq t} |X^{s,x}(u)| > r\right\} \leq \frac{\mathbf{E}V(\tau_{U_r}(t), X^{s,x}(\tau_{U_r}(t)))}{V_r} \leq \frac{V(s,x)}{V_r}.$$

Letting $t \to \infty$, we finally have

$$\mathbf{P}\left\{\sup_{u \geq s} |X^{s,x}(u)| > r\right\} \leq \frac{V(s,x)}{V_r}.$$

Since $V(s,0) = 0$ and the function $V(s,x)$ is continuous, this implies the desired assertion. □

Remark 5.3 We shall say that the solution $X(t) \equiv 0$ of (5.1) is *uniformly stable in probability* for $t > 0$ if for any $\varepsilon > 0$ the function $\mathbf{P}\{\sup_{t>s} |X^{s,x}(t)| \geq \varepsilon\}$ tends to zero as $x \to 0$, uniformly in $s \geq 0$. An examination of the proof of Theorem 5.3 immediately reveals that a sufficient condition for uniform stability in probability is that the function $V(t,x)$ satisfies the assumptions of Theorem 5.3 and that it has an infinitesimal upper limit, i.e.,

$$\lim_{x \to 0} \sup_{t > 0} V(t,x) = 0.$$

A question of major theoretical and practical interest is whether there exists for every system which is stable in probability, a Lyapunov function satisfying the assumptions of Theorem 5.3. For simplicity's sake, we shall confine ourselves to the time-homogeneous case, assuming moreover that the "noise" is nondegenerate everywhere, except at $x = 0$.

Theorem 5.4 [4] *Assume that the coefficients b and σ_r of (5.1) are independent of time, and that its solution $X(t) \equiv 0$ is stable in probability. Suppose that in a neighborhood of $x = 0$ condition (5.2) holds and also the nondegeneracy condition*

$$\sum_{i,j=1}^{l} a_{ij}\lambda_i\lambda_j > m(x) \sum_{i=1}^{l} \lambda_i^2, \tag{5.13}$$

where $m(x)$ is a continuous function such that $m(x) > 0$ for $x \neq 0$.

Then in a neighborhood of $x = 0$ there exists a positive definite function $V(x)$, twice continuously differentiable except perhaps at $x = 0$, such that $LV = 0$.

Proof Let $U_r = \{|x| < r\}$ be a sufficiently small neighborhood of $x = 0$. Let $u_\delta(x)$ denote a solution in the domain $U_r \setminus U_\delta$ of the problem

$$Lu = 0; \qquad u|_{|x|=r} = 1; \qquad u|_{|x|=\delta} = 0.$$

---

[4] See [118, 169].

It follows from Lemma 3.4 that

$$u_\delta(x) = \mathbf{P}\{|X^x(\tau_{r,\delta})| = r\},$$

where $\tau_{r,\delta}$ is the first time at which the sample path reaches the set $\{|x| = r\} \cup \{|x| = \delta\}$.

It is clear that the sequence $u_\delta(x)$ of L-harmonic functions is monotone increasing as $\delta \to 0$. Its limit $V(x)$ is also L-harmonic. Let τ_0 denote the first time at which the path of the process reaches the point 0. Then it follows from the obvious relations between events that

$$\left\{\sup_{t>0}|X^x(t)| \geq r\right\} \subset \bigcup_{\delta>0}\{|X^x(\tau_{r,\delta})| = r\} \cup \{\tau_0 < \infty\},$$

$$\bigcup_{\delta>0}\{|X^x(\tau_{r,\delta})| = r\} \subset \left\{\sup_{t>0}|X^x(t)| \geq r\right\},$$

and from Lemma 5.3 we obtain

$$\mathbf{P}\left\{\sup_{t>0}|X^x(t)| \geq r\right\} = \lim_{\delta\to 0}\mathbf{P}\{|X^x(\tau_{r,\delta})| = r\} = V(x).$$

Since the solution $X \equiv 0$ is stable in probability, it follows from this equality that $V(x) \to 0$ as $x \to 0$. Finally, the strong maximum principle implies that the function $u_\delta(x)$ and hence also $V(x)$, is positive for $|x| > \delta_1 > \delta$. Thus the function $V(x)$ is positive definite in Lyapunov's sense, and $LV = 0$. This we wished to prove.   □

Remark 5.4 Malkin [190] showed that the analog of Theorem 5.4 for deterministic systems does not hold. It follows that the nondegeneracy condition (5.13) cannot be dropped (though it can be weakened).

Remark 5.5 The Lyapunov function constructed in Theorem 5.4 at zero is only continuous. It is readily shown that in general a Lyapunov function which is smooth at zero may not exist. This will be clear from the following example.

Let $X(t)$ be a one-dimensional process, described by the equation

$$dX = bX\,dt + \sigma X\,d\xi(t), \tag{5.14}$$

where b and σ are constants. The generator of this process is

$$L = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + bx \frac{\partial}{\partial x}.$$

If $b < \sigma^2/2$, the solution $X(t) \equiv 0$ of the system (5.14) is stable, since the function $V(x) = |x|^{1-2b/\sigma^2}$ satisfies the assumptions of Theorem 5.3. If $b \geq 0$, this function is not differentiable at zero. Using the maximum principle for elliptic equations, one readily shows that any function $V_1(x)$ such that $V_1(0) = 0$, $V_1(\varepsilon) \geq \delta$, satisfies

$V_1(x) \geq \delta(|x|/|\varepsilon|)^{1-2b/\sigma^2}$ in the domain $0 < x < \varepsilon$. Hence it is clear that when $b > 0$, then there can be no Lyapunov function smooth at the origin and independent of t. A similar argument shows that there does not even exist a Lyapunov function smooth at zero which depends on t but has an infinitesimal upper limit.

5.4 Asymptotic Stability in Probability and Instability[5]

The solution $X(t) \equiv 0$ of (5.1) is said to be *asymptotically stable in probability* if it is stable in probability and moreover

$$\lim_{x \to 0} \mathbf{P}\left\{ \lim_{t \to \infty} X^{s,x}(t) = 0 \right\} = 1. \tag{5.15}$$

In this section we shall frequently assume that the following condition is satisfied[6]:

Condition D *Any solution of* (5.1), *beginning in the domain* $\varepsilon < |x| < r$, *almost surely reaches the boundary of this domain in a finite time, for any sufficiently small r and $\varepsilon > 0$.*

It follows from Theorem 3.9 that Condition D is satisfied if there exists in the domain $0 < |x| < r$ a function $W(t,x) \in \mathbf{C}_2^0(\{t > 0\} \times U_r)$, such that for any ε, $0 < \varepsilon < r$,

$$W(t,x) \geq 0, \qquad LW(t,x) < -c_\varepsilon < 0, \quad \text{if } |x| > \varepsilon. \tag{5.16}$$

In the following theorem, $U \subset \mathbb{R}^l$ is some neighborhood of the origin.

Theorem 5.5 *Suppose that there exists a positive definite function $V(t,x) \in \mathbf{C}_2^0(\{t > 0\} \times U)$, which has an infinitesimal upper limit and satisfies $LV \leq 0$. Let Condition D hold. Then the solution $X(t) \equiv 0$ of (5.1) is asymptotically stable in probability.*

Proof By Lemma 5.4, the stochastic process $V(\tau_U(t), X^{s,x}(\tau_U(t)))$ is a super-martingale. By Theorem 5.1, this implies that almost surely:

$$\lim_{t \to \infty} V(\tau_U(t), X^{s,x}(\tau_U(t))) = \xi. \tag{5.17}$$

---

[5]The conditions for asymptotic stability and instability in Theorems 5.5 and 5.6 generalize corresponding results of Khasminskii [118]. Conditions for stability in the large have been considered by Nevelson [215], to whom, in particular, Theorem 5.8 is due. Stability in the large of stochastic systems in a different setting has been investigated by Kac [108].

[6]An analogous condition for deterministic systems was considered by Krasovskii [155, p. 23] in connection with the inversion of Lyapunov's theorems on asymptotic stability and instability.

Let B_x denote the set of sample paths of $X^{s,x}(t)$ such that $\tau_U = \infty$. Since the function V satisfies the assumptions of Theorem 5.3, the solution $X(t) \equiv 0$ is stable in probability, and consequently

$$P(B_x) \to 1 \quad \text{as } x \to 0. \tag{5.18}$$

It follows from Condition D that for all paths contained in the set B_x, except for a set of paths of probability zero, we have $\inf_{t>0} |X^{s,x}(t)| = 0$, and in view of Lemma 5.3 we have also the stronger relation

$$\lim_{t \to \infty} |X^{s,x}(t)| = 0.$$

Since the function V has an infinitesimal upper limit, it follows that also $\underline{\lim}_{t \to \infty} V(t, X^{s,x}(t)) = 0$. But by (5.17) the limit

$$\lim_{t \to \infty} V(\tau_U(t), X^{s,x}(\tau_U(t))) = \lim_{t \to \infty} V(t, X^{s,x}(t))$$

exists for almost all paths in B_x. By the above reasoning this limit is equal to zero. Since the function $V(t, x)$ is positive definite for paths in B_X, this implies that

$$\lim_{t \to \infty} |X(t)| = 0.$$

The assertion of the theorem follows now from this relation and (5.18). $\qquad\square$

Corollary 5.1 *As it was mentioned above, Condition D may be replaced by the requirement that there exists a function $W(t, x)$ satisfying the inequalities (5.16). The function $V(t, x)$ itself satisfies these inequalities if LV is negative definite. We have thus proved the following generalization of Lyapunov's theorem on asymptotic stability of deterministic systems*: The solution $X(t) \equiv 0$ of (5.1) is asymptotically stable in probability if there exists in the domain $\{t > 0\} \times U$ a positive definite function $V(t, x) \in \mathbf{C}_2^0(\{t > 0\} \times U)$, which has an infinitesimal upper limit, such that the function LV is negative definite in this domain.

Corollary 5.2 *Condition D always holds if the matrix $A(t, x)$ satisfies the nondegeneracy condition (5.13). Indeed, then the function $W = k - |x|^n$ satisfies conditions (5.16) for a suitable choice of k and n. This means that, if condition (5.13) holds, then the existence of a function $V(t, x)$ satisfying the assumptions of Theorem 5.3 and having an infinitesimal upper limit is also sufficient for asymptotic stability in probability of the solution $X(t) = 0$ of (5.1). This fact and Theorem 5.4 yield the following proposition*: Assume that the coefficients b and σ_r are independent of t and that the nondegeneracy condition (5.13) is satisfied. Then, if the solution of (5.1) is stable in probability, it is also asymptotically stable in probability. *This proposition can be generalized to time-non-homogeneous systems. The example of deterministic systems shows that condition (5.13) cannot be dropped (though it can be weakened).*

As before, we let U_r denote the subset $\{|x| < r\}$ of $\mathbb{R}^l$.

Theorem 5.6 *Assume that there exists a function* $V(t, x) \in \mathbf{C}_2^0(\{t > 0\} \times U_r)$ *such that*

$$LV \leq 0 \quad as \ x \in U_r, \ x \neq 0, \tag{5.19}$$

$$\liminf_{\substack{x \to 0 \\ t > 0}} V(t, x) = \infty. \tag{5.20}$$

Let Condition D *hold. Then the solution* $X(t) \equiv 0$ *of* (5.1) *is not stable in probability. Moreover, in this case the event*

$$\left\{ \sup_{t > 0} |X^{s,x}(t)| < r \right\}$$

has probability zero for all $s > 0$, $x \in U_r$.

Proof Let $\tau_{r,\varepsilon}$ denote the first time of reaching the set $\{|x| = r\} \cup \{|x| = \varepsilon\}$, $\tau_{r,\varepsilon}(t) = \min(\tau_{r,\varepsilon}, t)$. By (5.19) and Lemma 3.2

$$\mathbf{E}V(\tau_{r,\varepsilon}(t), X^{s,x}(\tau_{r,\varepsilon}(t))) \leq V(s, x)$$

holds in the domain $U_r \setminus U_\varepsilon$ for any $\varepsilon < r$. Letting $t \to \infty$ and using Condition D, we conclude that $\mathbf{E}V(\tau_{r,\varepsilon}, X^{s,x}(\tau_{r,\varepsilon})) \leq V(s, x)$. Chebyshev's inequality implies now the estimate

$$\inf_{|x| < \varepsilon, t > 0} V(t, x) \mathbf{P}\left\{ \sup_{s < t < \tau^\varepsilon} |X^{s,x}(t)| < r \right\} < V(s, x),$$

where τ^ε is the first time the set $|x| = \varepsilon$ is reached. Since by Lemma 5.3, $\tau^\varepsilon \to \infty$ almost surely as $\varepsilon \to 0$, we infer the required assertion from the last inequality and (5.20), letting $\varepsilon \to 0$. $\qquad\square$

Remark 5.6 Arguments similar to those used to deduce the corollary from Theorem 5.5 yield the following sufficient conditions for instability.

(1) *The solution* $X(t) \equiv 0$ *of* (5.1) *is unstable if conditions* (5.19), (5.20) *and* (5.13) *hold in the domain* $\{t > 0\} \times U_r$.
(2) *The solution* $X(t) \equiv 0$ *of* (5.1) *is unstable if condition* (5.20) *holds and moreover* $\sup_{\varepsilon < |x| < r} LV < 0$ *for any* $\varepsilon > 0$.

Definition 5.1 The solution $X(t) \equiv 0$ of (5.1) is said to be (*asymptotically*) *stable in the large* if it is stable in probability and also for all s, x

$$\mathbf{P}\left\{ \lim_{t \to \infty} X^{s,x}(t) = 0 \right\} = 1.$$

Theorem 5.7 *A sufficient condition for the solution* $X(t) \equiv 0$ *of* (5.1) *to be stable in the large is that it is uniformly stable in probability, and moreover the process* $X(t)$ *is recurrent relative to the domain* $|x| < \varepsilon$ *for any* $\varepsilon > 0$.

Proof Since the solution $X(t) \equiv 0$ is uniformly stable in probability, it follows that for any $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\sup_{s>0,\,|y|<\delta} \mathbf{P}\Big\{ \sup_{t>s} |X^{s,y}(t)| > \varepsilon \Big\} < \varepsilon.$$

Let τ_δ denote the first time at which the path of the process reaches the set $|x| \le \delta$. By assumption, $\tau_\delta < \infty$ almost surely. Using the strong Markov property of the process and choosing $\delta > 0$ such that $|x| > \delta$, we get

$$\mathbf{P}\Big\{ \varlimsup_{t\to\infty} |X^{s,x}(t)| > \varepsilon \Big\}$$
$$= \int_{u=s}^{\infty} \int_{|y|=\delta} \mathbf{P}\{\tau_\delta \in du,\, X^{s,x}(\tau_\delta) \in dy\} \mathbf{P}\Big\{ \varlimsup_{t\to\infty} |X^{u,y}(t)| > \varepsilon \Big\}$$
$$= \int_{u=s}^{\infty} \int_{|y|=\delta} \mathbf{P}\{\tau_\delta \in du,\, X^{s,x}(\tau_\delta) \in dy\} \mathbf{P}\Big\{ \sup_{t>u} |X^{u,y}(t)| > \varepsilon \Big\}$$
$$\le \varepsilon.$$

This implies the required assertion. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

From Theorem 5.7 one readily derives various sufficient conditions for stability in the large in terms of Lyapunov functions. The following theorem generalizes to stochastic equations a well-known theorem of Barbashin and Krasovskii [18].

Theorem 5.8 *A sufficient condition for the solution $X(t) \equiv 0$ of (5.1) to be stable in the large is that there exists a positive definite function $V(t, x) \in \mathbf{C}_2^0(E)$ with an infinitesimal upper limit such that the function LV is negative definite and*

$$\inf_{t>0} V(t, x) \to \infty \quad as\ |x| \to \infty.$$

Proof We observe that under these assumptions the solution $X(t) \equiv 0$ is uniformly stable in probability, by virtue of the Remark 5.3. Moreover, by Lemma 3.9 and Theorem 3.9, this solution is recurrent relative to the domain $|x| < \varepsilon$ for any $\varepsilon > 0$. $\quad\square$

Theorem 5.9 *The following conditions are sufficient for the solution $X(t) \equiv 0$ of (5.1) to be stable in the large:*

(1) *the process $X(t)$ is regular;*
(2) *there exists a nonnegative function $V_1(t, x) \in \mathbf{C}_2^0(E)$ such that the function LV_1 is negative definite;*
(3) *there exists a positive definite function $V_2(t, x) \in \mathbf{C}_2^0(E)$, having an infinitesimal upper limit, such that $LV_2 \le 0$.*

Proof The proof follows from Theorem 5.7 and from the above mentioned theorems of Chaps. 3 and 5. Note that by Theorem 3.3 we can replace condition (1) by

the following condition (1′): There exists a nonnegative function $V_3(t, x) \in \mathbf{C}_2^0(E)$ such that $LV_3 < kV_3$ for some positive constant k and $\lim_{R \to \infty} \inf_{|x|>R} V_3 = \infty$. Similarly, (2) can be replaced by (2′): The nondegeneracy condition (5.13) holds in $U_R \setminus U_\varepsilon$ for any R and $\varepsilon < R$. (2) may be also replaced by the even weaker condition that $a_{ii}(t, x) > a_{R,\varepsilon} > 0$ for some i. □

5.5 Examples

Example 5.3 Consider the one-dimensional process described by the following Itô equation in $\mathbb{R}^l$:

$$dX(t) = b(t, X)\,dt + \sigma(t, X)\,d\xi(t). \tag{5.21}$$

Here the generator is

$$L = \frac{\partial}{\partial t} + b(t, x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(t, x)\frac{\partial^2}{\partial x^2}. \tag{5.22}$$

Suppose that the expansions

$$b(t, x) = b(t)x + o(|x|); \qquad \sigma(t, x) = \sigma(t)x + o(|x|) \tag{5.23}$$

hold in a neighborhood of $x = 0$ where, in accordance with the conditions of Sect. 5.1, the functions $b(t)$ and $\sigma(t)$ are bounded and the relations (5.23) hold uniformly in $t > 0$.

Assume that

$$\int_0^t \left[b(s) - \frac{\sigma^2(s)}{2} + \varepsilon \right] ds < k \tag{5.24}$$

holds for some $\varepsilon > 0$, $k > 0$ and all $t > 0$. Then for sufficiently small $\nu > 0$ the auxiliary function

$$V_1(t, x) = |x|^\nu \exp\left\{ -\nu \int_0^t \left(b(s) - \frac{\sigma^2(s)}{2} + \varepsilon \right) ds \right\} = |x|^\nu V(t)$$

satisfies all the assumptions of Theorem 5.3. Indeed, that $V_1(t, x)$ is positive definite follows from (5.24). Moreover, by (5.22) and (5.23),

$$LV_1(t, x) = \nu|x|^\nu V(t)[-\varepsilon + \nu\sigma^2(t)/2] + o(|x|^\nu).$$

Thus, if $\nu < \varepsilon / \sup_{t>0} \sigma^2(t)$, then the function $LV_1(t, x)$ is negative definite in a sufficiently small neighborhood of $x = 0$. Consequently, *the solution $X \equiv 0$ is stable in probability if condition (5.24) holds.*

Let us now assume that

$$\int_0^t \left[b(s) - \frac{\sigma^2(s)}{2} - \varepsilon \right] ds > -k \tag{5.25}$$

holds for some $\varepsilon > 0$, $k > 0$ and all $t > 0$. Then the auxiliary function

$$V_2(t, x) = -\ln|x| + \int_0^t \left[b(s) - \frac{\sigma^2(s)}{2} - \varepsilon \right] ds$$

obviously satisfies condition (5.20). Moreover

$$LV_2(t, x) \leq -\varepsilon + o(1) \quad (x \to 0).$$

Hence, by Remark 5.6, it follows that the trivial solution of (5.21) is unstable if condition (5.25) holds.

In the cases considered above, the stability (or instability) of the linearized system

$$dX(t) = b(t)X \, dt + \sigma(t)X \, d\xi(t)$$

implies the stability (or instability) of the full system (5.21). In the general case, however, this is not so. In Chap. 7 we shall consider in greater detail the question of conditions under which the theorem on stability in the first approximation is valid.

We mention one other peculiar consequence of Example 5.3. Condition (5.24) is satisfied, in particular, by a system (5.21) in which the function $b(s)$ is positive, provided the difference $b(s) - \sigma^2(s)/2$ is smaller than a negative constant. Thus, the system $dx/dt = b(t, x)$, which is unstable (even in the linear approximation!) can be "stabilized" by introducing an additive stochastic term $\sigma(t, x) \, d\xi(t)$, if the "intensity" of the noise $\sigma^2(t, x)$ is sufficiently high. For example, the linear system with constant coefficients $dX = bX \, dt + \sigma X \, d\xi(t)$ is stable for $b < \sigma^2/2$. Writing this equation as

$$\dot{X} = (b + \sigma\xi)X, \tag{5.26}$$

one is tempted to interpret the results as follows: An unstable first-order deterministic system is stabilized if white noise of sufficiently high intensity is superimposed on its coefficient. This assertion is in conflict with physical intuition. Neither is it correct, if we define a solution of (5.26) to be the limit of a sequence of solutions $x_n(t)$ of the equations $\dot{x}_n = (b + \sigma\xi_n(t))x_n$, where $\xi_n(t)$ is a sequence of Gaussian processes whose autocorrelation functions converge to a Dirac δ-function. It can be shown (this was established heuristically by Stratonovich [256] and rigorously by the author [124]) that in this case (and also under much more general conditions) the above limit procedure leads not to an Itô equation but to the analogous stochastic equation

$$dX(t) = bX(t) \, dt + \sigma X(t)d^*\xi(t),$$

where the stochastic differential $d^*\xi(t)$ is to be understood in the sense of Stratonovich [255].

The solution of the Itô equation (5.1) can be constructed as the limit in mean square as $h \to 0$ of solutions of the stochastic difference equations

$$X(t_n + h) - X(t_n) = b(t_n, X(t_n))h + \sum_{r=1}^{k} \sigma_r(t_n, X(t_n))[\xi_r(t_n + h) - \xi_r(t_n)].$$

On the other hand, the solution of the analogous Stratonovich equation

$$dX(t) = b(t, X(t)) dt + \sum_{r=1}^{k} \sigma_r(t, X(t)) d^* \xi_r(t) \tag{5.27}$$

is defined as the limit in mean square as $h \to 0$ of the solutions of the finite-difference equations

$$X(t_n + h) - X(t_n) = b(t_n, X(t_n))h$$
$$+ \sum_{r=1}^{k} \sigma_r \left(t_n, \frac{X(t_n) + X(t_n + h)}{2} \right) (\xi_r(t_n + h) - \xi_r(t_n)).$$

$$\tag{5.28}$$

Stratonovich [255] showed that if the functions $\sigma_r(t, x)$ are continuously differentiable with respect to x, then the above implicit difference scheme is equivalent, as $h \to 0$, to the explicit scheme

$$X(t_n + h) - X(t_n) = \left[b(t_n, X(t_n)) + \frac{1}{2} \sum_{r=1}^{k} \frac{\partial \sigma_r(t_n, X(t_n))}{\partial x} \sigma_r(t_n, X(t_n)) \right] h$$
$$+ \sum_{r=1}^{k} \sigma_r(t_n, X(t_n))[\xi_r(t_n + h) - \xi_r(t_n)]. \tag{5.29}$$

Consequently, the stochastic equation (5.27) is equivalent to the Itô equation

$$dX(t) = \left[b(t, X(t)) + \frac{1}{2} \sum_{r=1}^{k} \frac{\partial \sigma_r}{\partial x} \sigma_r(t, X(t)) \right] dt + \sum_{r=1}^{k} \sigma_r(t, X(t)) d\xi_r(t). \tag{5.30}$$

Thus, the generator of the process defined by (5.27) is

$$L = \frac{\partial}{\partial t} + \left(b(t, x) + \frac{1}{2} \sum_{r=1}^{k} \frac{\partial \sigma_r}{\partial x} \sigma_r(t, x), \frac{\partial}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t, x), \frac{\partial}{\partial x} \right)^2$$

$$= \frac{\partial}{\partial t} + \left(b(t, x), \frac{\partial}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t, x), \frac{\partial}{\partial x} \left(\sigma_r(t, x), \frac{\partial}{\partial x} \right) \right). \tag{5.31}$$

The reason for the equivalence of the difference schemes (5.28) and (5.29) may be explained briefly as follows. Since the vectors $\sigma_r(t, x)$ are continuously differentiable with respect to x, it follows that

$$
\sigma_r\left(t_n, \frac{X(t_n) + X(t_n + h)}{2}\right)
$$

$$
= \sigma_r(t_n, X(t_n)) + \frac{1}{2}\left(\frac{\partial\sigma_r(t_n, X(t_n))}{\partial x} + o(1)\right)(X(t_n + h) - X(t_n))
$$

as $h \to 0$. Substituting this relation into (5.28) and using the equality

$$
\left(J - \sum_{r=1}^{k}\frac{\partial\sigma_r}{\partial x}\Delta\xi_r(t)\right)^{-1} = J + \sum_{r=1}^{k}\frac{\partial\sigma_r}{\partial x}\Delta\xi_r(t) + o(\Delta\xi_r(t)) \quad \text{as } \Delta\xi_r(t) \to 0
$$

(where J is the identity matrix and $\Delta\xi_r(t) = \xi_r(t + h) - \xi_r(t)$), we conclude that

$$
X(t_n + h) - X(t_n)
$$

$$
= b(t_n, X(t_n))h + \sum_{r=1}^{k}\sigma_r(t_n, X(t_n))\Delta\xi_r(t_n)
$$

$$
+ \sum_{r=1}^{k}\sum_{j=1}^{k}\frac{\partial\sigma_r}{\partial x}(t_n, X(t_n))\sigma_j(t_n, X(t_n))\Delta\xi_r(t_n)\Delta\xi_j(t_n) + o(h). \quad (5.32)
$$

It can be shown further that the terms corresponding to $j \neq r$ in the last sum are in a certain sense small quantities of higher order than h as $h \to 0$, owing to the mutual independence of $\xi_r(t)$ and $\xi_j(t)$ for $r \neq j$. Moreover, as $h \to 0$, the expression $(\Delta\xi_r(t_n))^2$ in (5.32) can be replaced by its expectation h. It should now be clear from (5.32) and these relations why (5.28) and (5.29) are equivalent as $h \to 0$.

It follows from the above considerations that in dealing with physical problems in which white noise is an idealization of a real process with small time correlation, it is often natural to regard the equation as a stochastic equation of type (5.27).

In particular, we see from (5.31) that in the one-dimensional case considered above the operator associated with (5.26), regarded as a Stratonovich equation, is

$$
L = \frac{\partial}{\partial t} + bx\frac{\partial}{\partial x} + \frac{\sigma^2}{2}x\frac{\partial}{\partial x}\left(x\frac{\partial}{\partial x}\right) = \frac{\partial}{\partial t} + \left(b + \frac{\sigma^2}{2}\right)x\frac{\partial}{\partial x} + \frac{\sigma^2}{2}x^2\frac{\partial^2}{\partial x^2}.
$$

Combining this with Example 5.1, we get the result that the process $X(t)$ defined by the equation

$$
dX = bX\,dt + \sigma X d^*\xi(t)
$$

is stable for $b < 0$ and unstable for $b > 0$. This accords with physical intuition.

Thus, the unstable one-dimensional deterministic system $\dot{x} = bx$ ($b > 0$ is a constant) cannot be stabilized by a "physically"[7] feasible perturbation of its parameter. This was noticed by Leibowitz [180], who also conjectured that an analogous result holds in the multi-dimensional case. However, the example presented below in Sect. 6.9 shows that this is not so.

Example 5.4 Let the deterministic system

$$\frac{dx}{dt} = b(t, x) \tag{5.33}$$

be exponentially stable in the sense of Remark 1.10, and suppose that the function $b(t, x)$ has bounded first and second derivatives with respect to the space variables. Then, slightly modifying the proof of Theorem 11.1 in [155], we easily see that there exists a function $W(t, x)$ for the system (5.33) such that

$$k_i |x|^2 < W(t, x) < k_2 |x|^2,$$

$$\frac{d^0 W}{dt} = \frac{\partial W}{\partial t} + \sum_{i=1}^{l} b_i(t, x) \frac{\partial W}{\partial x_i} < -k_3 |x|^2;$$

$$\left| \frac{\partial W}{\partial x} \right| < k_4 |x|; \qquad \left| \frac{\partial^2 W}{\partial x_i \partial x_j} \right| < k_5.$$

Using this Lyapunov function to investigate the stability of the "perturbed" system

$$dX(t) = b(t, X) \, dt + \sum_{r=1}^{k} \sigma_r(t, x) \, d\xi_r(t) + F(t, X) \, dt, \tag{5.34}$$

we get

$$LW = \frac{d^0 W}{dt} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(t, x) \frac{\partial^2 W}{\partial x_i \partial x_j} + \sum_{i=1}^{l} F_i(t, x) \frac{\partial W}{\partial x_i}$$

$$\leq -k_3 |x|^2 + k_5 \|A(t, x)\| + k_4 |x| |F(t, x)|.$$

Hence it follows by Theorem 5.5 that the solution $X(t) \equiv 0$ of (5.34) is asymptotically stable in probability if the system (5.33) is exponentially stable in some neighborhood of the origin, and in this neighborhood

$$\sum_{r=1}^{k} |\sigma_r(t, x)| + |F(t, x)| < \varepsilon |x| \tag{5.35}$$

---

[7]We shall continue to use this term in this sense in the sequel, though it is somewhat vague. For example, if a stochastic process $X(t)$ is a "continuous approximation" to a discrete Markov chain which is the solution of a finite-difference equation, it is natural to use an Itô equation (see [251]).

for sufficiently small ε. If the system (5.33) is exponentially stable in the large and condition (5.35) holds everywhere, it follows from Theorem 5.7 that the system (5.34) is stable in the large. Therefore in this example sufficiently strong stability of the deterministic system (5.33) implies stability of the system (5.34), provided condition (5.35) holds for $\varepsilon < \varepsilon_0$. It is not difficult to find effective estimates for ε_0. In particular, the above conclusion holds true for "physically feasible" random perturbations of the parameters of the system by white noise.

Example 5.5 In our discussion of Example 5.1 we saw that the one-dimensional system $x' = bx$ $(b < 0)$ remains stable when its parameter is perturbed by white noise of arbitrary intensity. This holds true both for perturbations "of Itô's kind" and for physically feasible perturbations. We shall now show that if the dimension of the space is greater than 2, or in the case of physically feasible perturbations, greater than 1, then sufficiently strong isotropic noise will destroy the stability. To this end it will suffice to show that if σ is a sufficiently large constant, then for the systems

$$dX_i(t) = b_i(t, X)\,dt + \sigma \sum_{j=1}^{l} X_j\,d\xi_{(i-1)l+j}(t) \quad (i = 1, \ldots, l),$$

$$d\widetilde{X}_i(t) = b_i(t, \widetilde{X})\,dt + \sigma \sum_{j=1}^{l} \widetilde{X}_j d^*\xi_{(i-1)l+j}(t) \quad (i = 1, \ldots, l)$$

(where $\xi_1(t), \ldots, \xi_{2l}(t)$ are independent Wiener processes), there exists a function $V(x)$ satisfying the assumptions of Theorem 5.6, provided the functions $b_i(t, x)$ satisfy conditions (5.2), (5.3). It is readily seen that the generators of the processes $X(t)$ and $\widetilde{X}(t)$ are

$$L = \frac{\partial}{\partial t} + \sum_{i=1}^{l} b_i(t, x)\frac{\partial}{\partial x_i} + \frac{\sigma^2}{2}|x|^2 \sum_{i=1}^{l} \frac{\partial^2}{\partial x_i^2},$$

$$\widetilde{L} = L + \frac{\sigma^2}{2} \sum_{i=1}^{l} x_i \frac{\partial}{\partial x_i}.$$

Considering the auxiliary function

$$V(x) = -\ln|x|^2 = -\ln(x_1^2 + \cdots + x_l^2)$$

and assuming that the constant σ is sufficiently large, we see by (5.2) and (5.3) that

$$LV = -\frac{2(x, b(t, x))}{|x|^2} - \sigma^2(l - 2) < 0 \quad \text{for } l > 2,$$

$$\widetilde{L}V = -\frac{2(x, b(t, x))}{|x|^2} - \sigma^2(l - 1) < 0 \quad \text{for } l > 1.$$

Applying Theorem 5.6 and the subsequent Remark 5.6, we get the above assertions.

We shall now prove that the "Itô perturbations" considered in this example do not destroy the stability of the asymptotically stable system $dx_i = b_i x_i \, dt$ ($b_i < 0$, $i = 1, 2$) for any value of σ. In this case,

$$L = \frac{\partial}{\partial t} + b_1 x_1 \frac{\partial}{\partial x_1} + b_2 x_2 \frac{\partial}{\partial x_2} + \frac{\sigma^2}{2} |x|^2 \sum_{i=1}^{2} \frac{\partial^2}{\partial x_i^2}.$$

Considering the auxiliary function

$$V(x) = |x|^\alpha = (x_1^2 + x_2^2)^{\alpha/2}$$

for sufficiently small positive α, we readily obtain the inequality

$$LV(x) = \alpha |x|^{\alpha-2} \left(b_1 x_1^2 + b_2 x_2^2 + \frac{\alpha \sigma^2}{2} |x|^2 \right) < 0.$$

Hence, it follows by Theorems 5.5 and 5.7 that the system is asymptotically stable in the large.

Example 5.6 Consider the system

$$dX_1 = X_2 \, dt + \sigma(X_1, X_2) \, d\xi_1(t);$$
$$dX_2 = -X_1 \, dt + \sigma(X_1, X_2) \, d\xi_2(t).$$

It is clear that in the absence of random perturbations ($\sigma \equiv 0$) the equilibrium position of this system is stable, but not asymptotically stable. The generator is

$$L = x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2} + \frac{1}{2} \sigma^2(x) \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right).$$

It is obvious that the function $W(x) = -\ln(x_1^2 + x_2^2)$ satisfies conditions (5.19), (5.20). Consequently, the system is unstable if $\sigma(x) \neq 0$ for $x \neq 0$. This example shows that a non-asymptotically stable deterministic system may become unstable when driven by white noise whose intensity tends arbitrarily fast to zero as $x \to 0$.

5.6 Differentiability of Solutions of Stochastic Equations with Respect to the Initial Conditions

We have already seen (Sect. 5.3) that Lyapunov functions satisfying the assumptions of the stability theorems can be constructed as expectations of certain functionals of the relevant processes. However, only in the nondegenerate case can one guarantee, using the theory of partial differential equations, the necessary smoothness of these expectations. Gikhman [88] and Blagoveshchenskii and Freidlin [33] have demonstrated an alternative approach: One first proves that the solution $X^{s,x}(t)$ of

the stochastic equation is smooth with respect to s, x, and then the smoothness of the corresponding expectations follows as a corollary. This approach is applicable to processes with diffusion of arbitrary degree of degeneracy; for this reason it imposes stringent restrictions on the smoothness of the coefficients b and σ_r.

In this section we shall present Gikhman's theorem on the differentiability of the solutions of stochastic equations with respect to the initial conditions; we shall then establish certain auxiliary relations which will be used in Sect. 5.7 to prove the existence of Lyapunov functions for certain stable systems.

Theorem 5.10 *Let the coefficients of the equation*

$$dX^{s,x}(t) = b(t, X^{s,x}) dt + \sigma(t, X^{s,x}) d\xi(t) \qquad (5.36)$$

in $\mathbb{R}^l$ be continuous in t, x and with continuous bounded derivatives of order up to and including 2 with respect to $x_1, \ldots, x_l$. Then the solution $X^{s,x}(t)$ of (5.36) is twice continuously differentiable in mean square with respect to x.[8] The derivatives

$$\frac{\partial}{\partial x_i} X^{s,x}(t), \qquad \frac{\partial^2}{\partial x_i \partial x_j} X^{s,x}(t)$$

are then continuous in x in mean square. They are defined by the system obtained by formally differentiating (5.36) with respect to x.

We shall not give here all the details of the proof of Theorem 5.10; the interested reader can find them in [92]. We describe the idea of the proof, making more precise some of the arguments in [92] which will be needed later.

To avoid cumbersome notation, we limit ourselves to the case in which the dimension l of the space $\mathbb{R}^l$ is 1. It is easy to see that the stochastic process

$$Y_{x,\Delta x}(t) = \frac{1}{\Delta x} [X^{s,x+\Delta x}(t) - X^{s,x}(t)]$$

is a solution of the equation

$$X_{z,\Delta x}(t) = 1 + \int_s^t A(x, \Delta x, u) Y_{x,\Delta x}(u) \, du + \int_s^t B(x, \Delta x, u) Y_{x,\Delta x}(u) \, d\xi(u),$$
$$(5.37)$$

---

[8] If $\Phi(x_1, \ldots, x_l, t)$ is a random variable depending on the parameters $x_1, \ldots, x_l, t$, its partial derivative in mean square with respect to x_i is defined as the random variable $\frac{\partial \Phi}{\partial x_i}(x_1, \ldots, x_l, t)$ such that

$$\mathbf{E}\left\{ \frac{1}{\Delta x_i}[\Phi(x_1, \ldots, x_i + \Delta x_i, \ldots, x_l, t) - \Phi(x_1, \ldots, x_l, t)] - \frac{\partial \Phi}{\partial x_i}(x_1, \ldots, x_l, t) \right\}^2 \to 0$$

as $\Delta x_i \to 0$.

where

$$A(x, \Delta x, t) = \frac{b(t, X^{s,x+\Delta x}(t)) - b(t, X^{s,x}(t))}{X^{s,x+\Delta x}(t) - X^{s,x}(t)},$$

$$B(x, \Delta x, t) = \frac{\sigma(t, X^{s,x+\Delta x}(t)) - \sigma(t, X^{s,x}(t))}{X^{s,x+\Delta x}(t) - X^{s,x}(t)}.$$

By the assumptions of the theorem, the functions $|A|$ and $|B|$ are almost surely bounded by some constant k.

For any $n \geq 1$, we apply Itô's formula (3.30) to the process $Z(x, \Delta x, t) = [Y_{x,\Delta x}(t)]^{2n}$ and thus get from (5.37) the relation

$$Z(x, \Delta x, t) = 1 + n \int_s^t Z(x, \Delta x, u)[2A(x, \Delta x, u) + (2n - 1)B^2(x, \Delta x, u)] \, du$$

$$+ 2n \int_s^t Z(x, \Delta x, u)B(x, \Delta x, u) \, d\xi(u).$$

As in the proof of Lemma 5.2, we now calculate expectations on both sides of this equality and apply the Gronwall–Bellman lemma, to get the inequality

$$\mathbf{E}[Y_{x,\Delta x}(t)]^{2n} \leq e^{k(t-s)}, \tag{5.38}$$

where the constant k depends only on the lowest upper bounds of σ_x' and b_x' and the number n. In particular it follows from (5.38) that $X^{s,x+\Delta x}(t) \to X^{s,x}(t)$ in probability as $\Delta x \to 0$. Hence the coefficients A and B of (5.37) converge in probability as $\Delta x \to 0$ to the functions $b_x'(u, X^{s,x}(u))$ and $\sigma_x'(u, X^{s,x}(u))$, respectively. Since the functions A, B, b_x' and σ_x' are also bounded, it follows that all moments of the differences $A - b_x'$ and $B - \sigma_x'$ converge to zero. Hence, as before, we readily conclude that $Y_{x,\Delta x}(t)$ converges in mean square as $\Delta x \to 0$ to a solution of the equation

$$\zeta_x(t) = 1 + \int_s^t b_x'(u, X^{s,x}(u))\zeta_x(u) \, du + \int_s^t \sigma_x'(u, X^{s,x}(u))\zeta_x(u) \, d\xi(u). \tag{5.39}$$

By definition, the process $\zeta_x(t)$ is equal to $\partial X^{s,x}(t)/\partial x$. It is also easy to see on the basis of (5.37), (5.38) and (5.39) that for any integer $n \geq 1$

$$\left. \begin{array}{l} \mathbf{E}[\zeta_x(t)]^{2n} \leq e^{k(t-s)}, \\ \mathbf{E}[Y_{x,\Delta x}(t) - \zeta_x(t)]^{2n} \to 0 \quad \text{as } \Delta x \to 0. \end{array} \right\} \tag{5.40}$$

Similar arguments prove the existence and continuity of the second derivatives.

Lemma 5.5 *Let the coefficients of* (5.36) *be continuous in* t, x *and satisfy the conditions*

$$\sigma(t, 0) \equiv 0, \qquad b(t, 0) \equiv 0. \tag{5.41}$$

Suppose also that they have continuous bounded first and second partial derivatives with respect to $x_1, \ldots, x_l$. Then for any real β the function $u(s, x) = \mathbf{E}|X^{s,x}(t)|^\beta$ is twice continuously differentiable with respect to $x_1, \ldots, x_l$, except perhaps at $x = 0$. We also have then

$$
\left.
\begin{aligned}
\left| \frac{\partial u(s, x)}{\partial x} \right| &\le k_1 |x|^{\beta-1} e^{k_2(t-s)}, \\
\left| \frac{\partial^2 u(s, x)}{\partial x_i \partial x_j} \right| &\le k_1 |x|^{\beta-2} e^{k_2(t-s)}
\end{aligned}
\right\}
\tag{5.42}
$$

for some $k_1 > 0$, $k_2 > 0$.

Proof As before, we consider only the case $l = 1$, leaving it to the reader to extend the arguments to more dimensions. By formal differentiation, we get

$$
u'_x(s, x) = \beta \mathbf{E}\left[|X^{s,x}(t)|^{\beta-2} X^{s,x}(t) \frac{\partial X^{s,x}(t)}{\partial x} \right].
\tag{5.43}
$$

The existence of the expectation on the right of (5.43) follows from Lemma 5.2, Theorem 5.10 and the estimate

$$
|X^{s,x}(t)|^{\beta-1} \left| \frac{\partial X^{s,x}(t)}{\partial x} \right| < |X^{s,x}(t)|^{2\beta-2} + \left| \frac{\partial X^{s,x}(t)}{\partial x} \right|^2.
$$

We first consider the case $\beta = 2$. Then

$$
\left| \frac{u(s, x + \Delta x) - u(s, x)}{\Delta x} - 2\mathbf{E}\left[X^{s,x}(t) \frac{\partial X^{s,x}(t)}{\partial x} \right] \right|
$$

$$
\le \left(\mathbf{E}\left[\frac{(X^{s,x+\Delta x}(t))^2 - (X^{s,x}(t))^2}{\Delta x} - 2X^{s,x}(t) \frac{\partial X^{s,x}(t)}{\partial x} \right]^2 \right)^{1/2}
$$

$$
= \left\{ \mathbf{E}\left\{ 2X^{s,x}(t)\left(Y_{x,\Delta x}(t) - \frac{\partial X^{s,x}(t)}{\partial x} \right) + [Y_{x,\Delta x}(t)]^2 \Delta x \right\}^2 \right\}^{1/2}
$$

$$
\le \left\{ 32\left[\mathbf{E}(X^{s,x}(t))^4 \mathbf{E}\left(Y_{x,\Delta x}(t) - \frac{\partial X^{s,x}(t)}{\partial x} \right)^4 \right]^{1/2} + 2\mathbf{E}[Y_{x,\Delta x}(t)]^4 (\Delta x)^2 \right\}^{1/2}.
$$

Hence, by the estimates (5.38) and (5.40), it follows that the derivative $\frac{\partial}{\partial x}\mathbf{E}|X^{s,x}(t)|^2$ exists, and also that

$$
\sup_{s < t \le s+T} \mathbf{E}\left(\frac{\Delta y}{\Delta x} \right)^2 < K,
$$

$$
\mathbf{E}\left[\frac{\Delta y}{\Delta x} - 2X^{s,x}(t) \frac{\partial X^{s,x}(t)}{\partial x} \right]^2 \to 0 \quad \text{as } \Delta x \to 0,
\tag{5.44}
$$

where T is any positive number and we have set

$$
y = [X^{s,x}(t)]^2; \qquad y + \Delta y = [X^{s,x+\Delta x}(t)]^2.
\tag{5.45}
$$

Using (5.38), (5.40) and Lemma 5.2, we readily obtain

$$\mathbf{E}(\Delta y)^4 \to 0 \quad \text{as } \Delta x \to 0. \tag{5.46}$$

Now let $\beta \neq 2$. Then, again using the notation (5.45), we get

$$\left| \frac{u(s, x + \Delta x) - u(s, x)}{\Delta x} - \beta \mathbf{E}\left[|X^{s,x}(t)|^{\beta-2} X^{s,x}(t) \frac{\partial X^{s,x}(t)}{\partial x} \right] \right|$$

$$= \left| \mathbf{E}\left\{ \left[\frac{(y + \Delta y)^{\beta/2} - y^{\beta/2}}{\Delta y} - \frac{\beta}{2} y^{\beta/2-1} \right] \frac{\Delta y}{\Delta x} \right. \right.$$

$$\left. \left. + \frac{\beta}{2} y^{\beta/2-1} \left[\frac{\Delta y}{\Delta x} - 2X^{s,x}(t) \frac{\partial X^{s,x}(t)}{\partial x} \right] \right\} \right|$$

$$\leq \left\{ \mathbf{E}\left[\frac{(y + \Delta y)^{\beta/2} - y^{\beta/2}}{\Delta y} - \frac{\beta}{2} y^{\beta/2-1} \right]^2 \mathbf{E}\left(\frac{\Delta y}{\Delta x} \right)^2 \right\}^{1/2}$$

$$+ \frac{\beta}{2} \left\{ \mathbf{E} y^{\beta-2} \mathbf{E}\left(\frac{\Delta y}{\Delta x} - 2X^{s,x}(t) \frac{\partial X^{s,x}(t)}{\partial x} \right)^2 \right\}^{1/2}. \tag{5.47}$$

By Lemma 5.2 and (5.44), the second term on the right of (5.47) tends to zero as $\Delta x \to 0$. It also follows from (5.44) that the function $E(\Delta y/\Delta x)^2$ is bounded. We claim that the function

$$\mathcal{D} = \frac{(y + \Delta y)^{\beta/2} - y^{\beta/2}}{\Delta y} - \frac{\beta}{2} y^{\beta/2-1}$$

converges to zero in mean square.

Since $y > 0$, $y + \Delta y > 0$, it follows that

$$|\mathcal{D}| < k|\Delta y|[(y + \Delta y)^{\beta/2-2} + y^{\beta/2-2}],$$

for some $k > 0$[9] and hence the relation

$$\mathbf{E}\mathcal{D}^2 \leq k_1^2 (\mathbf{E}|\Delta y|^4)^{1/2} \{[\mathbf{E}(y + \Delta y)^{2\beta-8}]^{1/2} + (\mathbf{E}y^{2\beta-8})^{1/2}\}$$

holds. The last estimate, (5.46) and Lemma 5.2, imply that $\lim_{\Delta x \to 0} \mathbf{E}\mathcal{D}^2 = 0$. Hence, by (5.47), we conclude that the first derivative of $u(x)$ with respect to x exists and is given, by formula (5.43). Using (5.40) and (5.8), we see from (5.43) that also

$$|u'_x(s, x)| < |\beta| \mathbf{E}\left\{ |X^{s,x}(t)|^{\beta-1} \left| \frac{\partial X^{s,x}(t)}{\partial x} \right| \right\}$$

---

[9]This estimate follows from the obvious inequality ($u > 0$)

$$\frac{|u^{\beta/2} - 1 - (\beta/2)u^{\beta/2-1}(u - 1)|}{(u - 1)^2} < k(u^{\beta/2-2} + 1),$$

if we set $u = (y + \Delta y)/y$.

$$\leq |\beta| (\mathbf{E}|X^{s,x}(t)|^{2\beta-2})^{1/2} \left(\mathbf{E} \left| \frac{\partial X^{s,x}(t)}{\partial x} \right|^2 \right)^{1/2}$$

$$\leq k_1 |x|^{\beta-1} e^{k_2(t-s)}.$$

Similarly one proves the existence and continuity of the second derivative u''_{xx} and the estimate (5.42). $\square$

Lemma 5.6 *Under the assumptions of Lemma* 5.5, *the function* $u(s, x) = \mathbf{E}|X^{s,x}(t)|^{\beta}$ *is differentiable with respect to* s. *Moreover, for* $x \neq 0$

$$Lu = \frac{\partial u}{\partial s} + \frac{1}{2} \sum_{i,j=1}^{l} a_{ij}(s, x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{l} b_i(s, x) \frac{\partial u}{\partial x_i} = 0.$$

Proof The proof differs only in details from that of Theorem 5.1, Chap. VIII of [92]. We shall therefore confine the present discussion to those parts of the proof which are new. As before, we consider only the case $l = 1$. Expressing the difference $u(s + \Delta s, X^{s,x}(s + \Delta s)) - u(s + \Delta s, x)$ by means of Itô's formula and using the identity $\mathbf{E}u(s + \Delta s, X^{s,x}(s + \Delta s)) = u(s, x)$, we get

$$u(s, x) - u(s + \Delta s, x) = \mathbf{E} \int_s^{s+\Delta s} \left[\frac{\partial}{\partial x} u(s + \Delta s, X^{s,x}(t)) b(t, X^{s,x}(t)) \right.$$

$$\left. + \frac{1}{2} \sigma^2(t, X^{s,x}(t)) \frac{\partial^2}{\partial x^2} u(s + \Delta s, X^{s,x}(t)) \right] dt.$$

$$(5.48)$$

Next, using the explicit expressions for u, $\partial u / \partial x$ and $\partial^2 u / \partial x^2$ and proceeding as in the case of (5.42), we derive the estimates

$$|u_x^{(i)}(s + \Delta s, x) - u_x^{(i)}(s, x)| \leq k \Delta s |x|^{\beta-i} \quad (i = 0, 1, 2).$$

Applying these estimates and (5.42), we easily deduce the assertion of the lemma from (5.48) by letting $\Delta s \to 0$. $\square$

Corollary 5.3 *The function*

$$V(s, x) = \int_s^{s+T} \mathbf{E}|X^{s,x}(t)|^{\beta} dt$$

is in class $\mathbf{C}_2^0(E)$, *and*

$$LV(s, x) = \mathbf{E}|X^{s,x}(s + T)|^{\beta} - |x|^{\beta}.$$

Proof We set $u_t(s, x) = \mathbf{E}|X^{s,x}(t)|^\beta$. Differentiating V with respect to s and applying Lemma 5.6, we get

$$LV(s, x) = \mathbf{E}|X^{s,x}(s + T)|^\beta - \mathbf{E}|X^{s,x}(s)|^\beta + \int_s^{s+T} Lu_t(s, x)\, dt$$

$$= \mathbf{E}|X^{s,x}(s + T)|^\beta - |x|^\beta,$$

as required.   □

5.7   Exponential *p*-Stability and *q*-Instability[10]

The solution $X(t) \equiv 0$ of the system

$$dX(t) = b(t, X)\, dt + \sum_{r=1}^k \sigma_r(t, X)\, d\xi_r(t) \tag{5.49}$$

in $\mathbb{R}^l$ is said to be

(1) *p-stable* ($p > 0$) for $t \geq 0$, if

$$\sup_{|x| \leq \delta,\, t \geq s} \mathbf{E}|X^{s,x}(t)|^p \to 0 \quad \text{as } \delta \to 0 \ (s \geq 0);$$

(2) *asymptotically p-stable*, if it is *p*-stable and moreover $\mathbf{E}|X^{s,x}(t)|^p \to 0$ as $t \to \infty$;

(3) *exponentially p-stable*, if for some positive constants A and α

$$\mathbf{E}|X^{s,x}(t)|^p \leq A|x|^p \exp\{-\alpha(t - s)\}. \tag{5.50}$$

The case most frequently considered in the literature to date is that of *p*-stability for $p = 1$ (stability in the mean) and for $p = 2$ (stability in mean square).[11]

The following two theorems give necessary and sufficient conditions for exponential *p*-stability of stochastic systems in terms of Lyapunov functions. They may be regarded as generalizations of well-known theorems for deterministic systems (see [155, Sect. 11]).

Theorem 5.11 *The trivial solution of the system* (5.49) *is exponentially p-stable for* $t \geq 0$ *if there exists a function* $V(t, x)$ *of class* $\mathbf{C}_2^0(E)$ *such that*

---

[10]Theorems 5.11 and 5.12 are due to Nevelson and Khasminskii [221] (the second theorem is proved there under slightly more restrictive conditions).

[11]After the first edition of this book the study of *p*-stability became very popular, see Appendix B, and references therein.

$$k_1|x|^p \le V(t, x) \le k_2|x|^p, \tag{5.51}$$

$$LV(t, x) \le -k_3|x|^p \tag{5.52}$$

for certain positive constants k_1, k_2, k_3.

Proof Conditions (5.51) and (5.52) are sufficient for the process $X(t)$ to be regular, since the function $V(t, x)$ satisfies the assumptions of Theorem 3.5. It follows from the same theorem that $\mathbf{E}V(t, X^{s,x}(t))$ exists for all $t > s$. Expressing the difference $V(t, X^{s,x}(t)) - V(s, x)$ by means of Itô's formula (3.35), calculating expectations and using conditions (5.51) and (5.52), we get

$$\mathbf{E}V(t, X^{s,x}(t)) - V(s, x) = \int_s^t \mathbf{E}LV(u, X^{s,x}(u)) \, du.$$

Differentiating this equality with respect to t and using (5.51), (5.52), we see that

$$\frac{d}{dt} \mathbf{E}V(t, X^{s,x}(t)) \le -\frac{k_3}{k_2} \mathbf{E}V(t, X^{s,x}(t)).$$

This implies the estimate

$$\mathbf{E}V(t, X^{s,x}(t)) \le V(s, x) \exp\left\{-\frac{k_3}{k_2}(t - s)\right\}.$$

Together with (5.51), this estimate yields (5.50). The proof is complete. $\square$

Theorem 5.12 *If the solution $X(t) \equiv 0$ of the system (5.49) is exponentially p-stable and the coefficients b and σ_r have continuous bounded derivatives with respect to x up to second order, then there exists a function $V(t, x) \in \mathbf{C}_2^0(E)$ satisfying inequalities (5.51), (5.52) and also*

$$\left|\frac{\partial V}{\partial x_i}\right| < k_4|x|^{p-1}, \qquad \left|\frac{\partial^2 V}{\partial x_i \partial x_j}\right| < k_4|x|^{p-2} \tag{5.53}$$

for some $k_4 > 0$.

Proof We claim that the function

$$V(t, x) = \int_t^{t+T} \mathbf{E}|X^{t,x}(u)|^p \, du \tag{5.54}$$

satisfies all the conditions of the theorem for suitable choice of the constant $T > 0$. Indeed, by (5.50),

$$V(t, x) \le \int_t^{t+T} A|x|^p \exp\{-\alpha(u - t)\} \, du = k_1|x|^p.$$

Since the coefficients b and σ_r have bounded partial derivatives with respect to x_i, while $\sigma_r(t, 0) = 0$, $b(t, 0) = 0$, we have

$$|a_{ij}(t, x)| < k_5 |x|^2, \qquad |b_i(t, x)| < k_5 |x|.$$

Hence it follows that

$$|L(|x|^p)| < k_6 |x|^p. \tag{5.55}$$

Applying Itô's formula to the function $|x|^p$ and using (5.55), we get

$$\mathbf{E}|X^{t,x}(t+T)|^p - |x|^p = \int_t^{t+T} \mathbf{E}L(|X^{t,x}(u)|^p)\, du$$

$$\geq -k_6 \int_t^{t+T} \mathbf{E}|X^{t,x}(u)|^p\, du = -k_6 V(t, x).$$

Choosing T so that

$$\mathbf{E}|X^{t,x}(t+T)|^p < \frac{1}{2}|x|^p, \tag{5.56}$$

we thus get the inequality $V(t, x) > |x|^p/(2k_6)$. This proves (5.51). To prove the required smoothness of $V(t, x)$ and to verify (5.52), we apply the Corollary 5.3 and (5.56). Finally, using (5.42), we derive the estimate

$$\left| \frac{\partial V(t, x)}{\partial x_i} \right| = \left| \int_t^{t+T} \frac{\partial}{\partial x_i} \mathbf{E}|X^{t,x}(u)|^p\, du \right|$$

$$\leq \int_t^{t+T} k_1 |x|^{p-1} \exp\{k_2(u - t)\}\, du = k_4 |x|^{p-1}.$$

Proof of the second part of (5.53) is similar.   □

The next lemma is useful in investigations of the stabilization of stochastic systems (see Chap. 8).

Lemma 5.7 *Assume that the coefficients $b(t, x)$ and $\sigma_r(t, x)$ satisfy the conditions of Theorem 5.12, and moreover that*

$$\int_s^\infty \mathbf{E}|X^{s,x}(t)|^p\, dt < \infty. \tag{5.57}$$

Then

$$\lim_{t \to \infty} \mathbf{E}|X^{s,x}(t)|^p = 0. \tag{5.58}$$

Proof By Remark 5.2, we can apply Itô's formula (3.30) to the function $|x|^p$. Doing this and using the estimate (5.55), which follows from the assumptions of our

lemma, we see that

$$\left| \mathbf{E}|X^{s,x}(t+h)|^p - \mathbf{E}|X^{s,x}(t)|^p \right| < k \int_t^{t+h} \mathbf{E}|X^{s,x}(u)|^p \, du$$

holds for some constant $k > 0$. Thus,

$$\left| \frac{\partial}{\partial t} \mathbf{E}|X^{s,x}(t)|^p \right| \le k\mathbf{E}|X^{s,x}(t)|^p. \qquad (5.59)$$

Inequalities (5.57) and (5.59) obviously imply (5.58).   □

We now consider the concept of q-instability.

The trivial solution of the system (5.49) is said to be *exponentially q-unstable* $(q > 0)$ if

$$\mathbf{E}|X^{s,x}(t)|^{-q} < A|x|^{-q} \exp\{-\alpha(t-s)\}$$

for some positive constants A and α. Similarly, we modify the other definitions at the beginning of this section; we replace p by $-q$ and a neighborhood of 0 by a neighborhood of the point at infinity, to get the definitions of *q-instability* and asymptotic *q-instability*.

It is clear that asymptotic q-instability for some $q > 0$ implies instability in probability, since, by Chebyshev's inequality,

$$\mathbf{P}\{|X^{s,x}(t)| < R\} < R^q \mathbf{E}|X^{s,x}(t)|^{-q}$$

for any $R > 0$. In order to avoid the difficulties created by possible irregularity of the process $X(t)$, we shall assume till the end of this section that the coefficients b and σ_r of (5.49) have bounded derivatives with respect to the space variables.

Later, in connection with the problem of instability in the first approximation, we shall be especially interested in the investigation of exponential q-instability. The proofs of the following two theorems are almost word-for-word repetitions of those of Theorems 5.11 and 5.12.

Theorem 5.13 *The solution $X(t) \equiv 0$ of the system (5.49) is exponentially q-unstable for $t \ge 0$ if there exists a function $V(t, x)$ of class $\mathbf{C}_2^0(E)$ such that*

$$\begin{aligned} k_1|x|^{-q} &\le V(t, x) \le k_2|x|^{-q}; \\ LV(t, x) &\le -k_3|x|^{-q}. \end{aligned} \qquad (5.60)$$

Theorem 5.14 *If the coefficients b and σ_r have continuous bounded derivatives with respect to x up to second order, and the solution $X(t) \equiv 0$ of the system (5.49) is exponentially q-unstable, then there exists a function $V(t, x)$ satisfying inequalities (5.60) and the inequalities*

$$\left| \frac{\partial V}{\partial x_i} \right| \le k_4|x|^{-q-1}; \qquad \left| \frac{\partial^2 V}{\partial x_i \partial x_j} \right| < k_4|x|^{-q-2}.$$

Remark 5.7 It follows from Theorems 5.12 and 5.8 that the solution $X(t) \equiv 0$ of the system (5.49) is asymptotically stable in the large if it is exponentially p-stable for some $p > 0$, and the functions $b(t, x)$ and $\sigma_r(t, x)$ have continuous bounded derivatives with respect to x of order up to 2 inclusive.

Remark 5.8 Let the function $V(t, x) \in \mathbf{C}_2^0(E)$ be positive definite and such that $V(t, x) < k|x|^p$. Suppose moreover that $LV(t, x) \geq 0$. Then the system (5.49) is not asymptotically p-stable. Indeed, we deduce from Lemma 5.4 and Remark 5.2, by applying Itô's formula and then taking expectations, that

$$k\mathbf{E}|X^{s,x}(t)|^p > \mathbf{E}V(t, X^{s,x}(t)) \geq V(s, x).$$

5.8 Almost Sure Exponential Stability

Kozin [152] raises the question of finding conditions under which almost all solutions of the system (5.49) are exponentially stable. He proves that a sufficient condition is that the trivial solution be exponentially stable in mean square. Using a different method, we shall prove both this and a more general result.

Theorem 5.15 *Under the assumptions of Theorem 5.11, there exists a constant $\gamma > 0$ such that, if $x \in \mathbb{R}^l$, $s \geq 0$, the inequality $|X^{s,x}(t)| < K_{s,x}e^{-\gamma t}$ holds almost surely for $t \geq s$, where the random variable $K_{s,x}$ is almost surely finite.*

Proof Setting

$$W(t, x) = V(t, x)\exp\left\{\frac{k_3 t}{k_2}\right\}$$

we see by (5.51) and (5.52) that for $x \neq 0$

$$LW = \frac{k_3}{k_2}\exp\left\{\frac{k_3 t}{k_2}\right\}V + \exp\left\{\frac{k_3 t}{k_2}\right\}LV \leq 0.$$

Hence the process $W(t, X^{s,x}(t))$ is a supermartingale. Since it is positive, it follows from Theorem 5.1 that for all s, x the process $W(t, X^{s,x}(t))$ converges almost surely to a finite limit as $t \to \infty$. Consequently,

$$\sup_t W(t, X^{s,x}(t)) = A_{s,x} < \infty$$

almost surely. Therefore,

$$V(t, X^{s,x}(t)) \leq A_{s,x}e^{-\gamma t}.$$

This together with condition (5.51) implies the required assertion. □

In exactly the same way one proves the following

Theorem 5.16 *Under the assumptions of Theorem 5.13, there exists a constant* $\gamma > 0$ *such that for any* $x \neq 0$, $s \geq 0$

$$|X^{s,x}(t)| > K_{s,x}e^{\gamma t}$$

holds almost surely for $t \geq s$, *where the random variable* $K_{s,x}$ *is almost surely positive.*

Chapter 6
Systems of Linear Stochastic Equations

6.1 One-Dimensional Systems

In this chapter we shall study a linear homogeneous system of equations whose coefficients are perturbed by Gaussian white noise $\dot{\eta}_i^j(t)$. A system of this type can be written as

$$\frac{dX_i}{dt} = \sum_{j=1}^{l}(b_i^j(t) + \dot{\eta}_i^j(t))X_j(t), \quad i = 1, \ldots, l. \tag{6.1}$$

The white noise processes $\dot{\eta}_i^j(t)$ figuring here are generalized Gaussian stochastic processes with zero mean and covariance matrix

$$\mathbf{E}[\dot{\eta}_i^j(s)\dot{\eta}_n^m(t)] = k_{ij}^{mn}(t)\delta(t-s),$$

where $\delta(t)$ is the Dirac δ-function, $i, i, m, n = 1, \ldots, l$. It is well known that the dependent white noise processes $\dot{\eta}_i^j(t)$ may be replaced by linear combinations of at most l^2 independent processes.

We append a nonrigorous justification of the last statement (it can be made rigorous within the framework of the theory of generalized stochastic processes of Gelfand [86] and Itô [105]). Let $\dot{\eta}_1(t), \ldots, \dot{\eta}_N(t)$ be Gaussian white noise processes with covariance matrix with elements $k_{ij}(t)\delta(t-s)$. Let $\lambda_1(t), \ldots, \lambda_N(t)$ denote the eigenvalues and $f_1(t), \ldots, f_N(t)$ the associated normalized eigenvectors of the matrix $((k_{ij}(t)))$. Since the matrix $((k_{ij}(t)))$ is symmetric, the vectors $f_i(t)$ are orthogonal. Now let $\dot{\xi}_1(t), \ldots, \dot{\xi}_N(t)$ be independent Gaussian white noise processes with unit spectral density, so that

$$\mathbf{E}[\dot{\xi}_i(s)\dot{\xi}_j(t)] = \delta_{ij}\delta(t-s).$$

Define new (generalized) stochastic processes by

$$\dot{\tilde{\eta}}_i(t) = \sum_{k=1}^{N} \sqrt{\lambda_k(t)} f_k^{(i)}(t)\dot{\xi}_k(t), \quad i = 1, \ldots, N,$$

where $f_k^{(i)}(t)$ is the i-th component of the vector $f_k(t)$. Then, using the relations

$$\sum_{i=1}^{N} f_k^{(i)}(t) f_l^{(i)}(t) = \delta_{kl},$$

$$\lambda_k(t) f_k^{(i)}(t) = \sum_{j=1}^{N} k_{ij}(t) f_k^{(j)}(t),$$

which follow from the definition of $\lambda_k(t)$ and $f_k(t)$, we get

$$\mathbf{E}[\dot{\tilde{\eta}}(s)\dot{\tilde{\eta}}_j(t)] = \sum_{k=1}^{N} \sqrt{\lambda_k(s)\lambda_k(t)} f_k^{(i)}(s) f_k^{(j)}(t)\mathbf{E}[\dot{\xi}_k(s)\dot{\xi}_k(t)]$$

$$= k_{ij}(t)\delta(t-s).$$

Thus, the correlation matrices, and hence also the probability distributions of the processes $\dot{\eta}_1(t), \ldots, \dot{\eta}_N(t)$ and $\tilde{\eta}_1(t), \ldots, \tilde{\eta}_N(t)$ coincide. We may therefore set $((\dot{\eta}_i^j(t))) = \sum_{r=1}^{N} \sigma_r(t)\dot{\xi}_r(t)$, $N \leq l^2$, in (6.1).

We shall treat (6.1) as a system of Itô equations

$$dX(t) = B(t)X(t)\,dt + \sum_{r=1}^{N} \sigma_r(t)X(t)\,d\xi_r(t). \tag{6.2}$$

Here $B(t)$ and $\sigma_r(t) = ((\sigma_{ir}^j(t)))$ are $l \times l$ matrices, $l^2 \geq N$, and

$$\mathbf{E}\left[\sum_{r=1}^{N} \sigma_{ir}^j(s)\xi_r(s) \sum_{r=1}^{N} \sigma_{nr}^m(t)\xi_r(t)\right] = k_{ij}^{mn}(t)\delta(t-s).$$

In accordance with (5.2), we shall assume that $\|B(t)\|$, $\|\sigma_r(t)\|$ are bounded functions of time on any finite interval.

A linear stochastic equation with stochastic differentials in the sense of Stratonovich (see Chap. 5) can also be reduced to the form (6.2). To be precise, it follows from (5.30) that a linear system with Stratonovich differentials $d^*\xi(t)$ is equivalent to a system involving Itô differentials. Both systems have the same coefficients $\sigma_r(t)$ and the new drift coefficients are related to the old ones by

$$\widetilde{B}(t) = B(t) + \frac{1}{2}\sum_{r=1}^{N} \sigma_r^2(t).$$

We mention some properties of solutions of the system (6.2), which follow easily from the properties of stochastic integrals and the uniqueness of the solution.

1. If $X^{(1)}(t), \ldots, X^{(l)}(t)$ are the solutions of the system (6.2), then the function

$$Y(t) = \sum_{i=1}^{l} k_i X^{(i)}(t), \tag{6.3}$$

is also a solution of the system (6.2) for any constants $k_1, \ldots, k_l$.

2. If $X^{(i)}(t)$ are solutions such that the determinant of the matrix $(X_j^{(i)}(t_0))$ does not vanish, then the solution of the system (6.2) with initial condition $X(t_0) = x_0$ can be expressed as a sum (6.3) for suitable constants k_i (the system of solutions $X^{(1)}(t), \ldots, X^{(l)}(t)$ is fundamental).

One consequence of these properties is that a linear stochastic system which is asymptotically stable in probability is asymptotically stable in the large. We leave the verification to the reader.

It is well known that the solution of a deterministic linear system for $l = 1$ can be determined by quadratures. An analogous statement holds for the one-dimensional stochastic system

$$dX(t) = b(t)X(t)\,dt + \sigma(t)X(t)\,d\xi(t). \tag{6.4}$$

Indeed, a direct check shows that the function[1]

$$X(t) = x_0 \exp\left\{ \int_0^t \left[b(s) - \frac{\sigma^2(s)}{2} \right] ds + \int_0^t \sigma(s)\,d\xi(s) \right\}, \tag{6.5}$$

satisfies (6.4) and the initial condition $X(0) = x_0$. (To verify this one calculates $dX(t)$ using Itô's formula (3.30), treating $X(t)$ as a function of t and the process $y(t) = \int_0^t \sigma(s)\,d\xi(s)$.)

The representation (6.5) enables us to obtain conditions for the stability of solutions of (6.4) which are an improvement on those derived above (see (5.24) and (5.25)). First, it follows from (6.5) that the solution $X(t) \equiv 0$ of the system (6.4) is asymptotically stable if the process

$$\eta(t) = \int_0^t \left[b(s) - \frac{1}{2}\sigma^2(s) \right] ds + \int_0^t \sigma(s)\,d\xi(s) \tag{6.6}$$

satisfies the condition $\mathbf{P}\{\eta(t) \to -\infty \text{ as } t \to \infty\} = 1$.

Similarly, the solution $X(t) \equiv 0$ is stable if $\overline{\lim}_{t \to \infty} \eta(t) < \infty$ almost surely, and unstable if $\overline{\lim}_{t \to \infty} \eta(t) = \infty$ with positive probability.

The process $\eta(t)$ is clearly Gaussian and has independent increments. The study of its growth as $t \to \infty$ can be reduced to that of the growth of the Wiener process, by means of the following simple lemma.

---

[1]The representation of solutions of (6.4) in the form (6.5) is well known; see, e.g., [256], [89], [90].

Lemma 6.1 *Let $\int_{t_0}^{t} \sigma(s) \, d\xi(s)$ be an Itô stochastic integral with respect to a Wiener process. Then there exists another Wiener process $\tilde{\xi}(t)$ such that*

$$\int_{t_0}^{t} \sigma(s) \, d\xi(s) = \tilde{\xi}\left(\int_{t_0}^{t} \sigma^2(s) \, ds\right) \quad (a.s.) \tag{6.7}$$

for all $t \geq 0$.

Proof Let $\hat{t}(\tau)$ denote the smallest number such that $\tau = \int_{t_0}^{\hat{t}(\tau)} \sigma^2(s) \, ds$. Let us investigate some properties of the process $\tilde{\xi}(\tau) = \int_{t_0}^{\hat{t}(\tau)} \sigma(s) \, d\xi(s)$. By the properties of stochastic integrals it is easy to see that this process has independent increments and is Gaussian and moreover $\mathbf{E}\tilde{\xi}(\tau) = 0$, $\mathbf{E}[\tilde{\xi}(\tau)]^2 = \tau$. This means that $\tilde{\xi}(\tau)$ is a Wiener process. The assertion of the lemma now follows easily, since for any t we have almost surely:

$$\int_{t_0}^{t} \sigma(s) \, d\xi(s) = \int_{t_0}^{\hat{t}} \sigma(s) \, d\xi(s); \qquad \tilde{\xi}\left(\int_{t_0}^{t} \sigma^2(s) \, ds\right) = \tilde{\xi}\left(\int_{t_0}^{\hat{t}} \sigma^2(s) \, ds\right),$$

where $\hat{t}$ is the smallest number such that

$$\int_{t_0}^{t} \sigma^2(s) \, ds = \int_{t_0}^{\hat{t}} \sigma^2(s) \, ds. \qquad \square$$

In the sequel we shall need the following theorem of Khinchin [141].
Law of the iterated logarithm. If $\xi(t)$ is a Wiener process, then almost surely

$$\varlimsup_{t \to \infty} \frac{\xi(t)}{\sqrt{2t \ln \ln t}} = 1.$$

We set

$$\tau(t) = \int_0^t \sigma^2(s) \, ds, \qquad J(t) = \frac{\int_0^t [b(s) - \frac{1}{2}\sigma^2(s)] \, ds}{[2 \int_0^t \sigma^2(s) \, ds \ln \ln(\int_0^t \sigma^2(s) \, ds)]^{1/2}}.$$

Theorem 6.1 *If $\tau(\infty) < \infty$, then the inequality*

$$\varlimsup_{t \to \infty} \int_0^t b(s) \, ds < \infty$$

is a necessary and sufficient condition for the trivial solution of the system (6.4) *to be stable; and the condition*

$$\lim_{t \to \infty} \int_0^t b(s) \, ds = -\infty$$

is necessary and sufficient for asymptotic stability.

On the other hand, if $\tau(\infty) = \infty$, then $\overline{\lim}_{t\to\infty} J(t) < -1$ is a sufficient condition for asymptotic stability and $\underline{\lim}_{t\to\infty} J(t) > -1$ a sufficient condition for instability of the trivial solution.

Proof By Lemma 6.1, we can write the process $\eta(t)$ as

$$\eta(t) = \int_{t_0}^{t} \left[b(s) - \frac{1}{2}\sigma^2(s) \right] ds + \tilde{\xi}\left(\int_{t_0}^{t} \sigma^2(s)\,ds \right).$$

Hence, by virtue of the fact that almost surely

$$\sup_{0 \leq \tau < \tau_0} \tilde{\xi}(\tau) < \infty,$$

we at once obtain the first part of the theorem.

Now let $\tau(\infty) = \infty$. Then

$$\underline{\lim_{t\to\infty}} J(t) + \overline{\lim_{t\to\infty}} \frac{\tilde{\xi}(\tau(t))}{(2\tau(t)\ln\ln\tau(t))^{1/2}} \leq \overline{\lim_{t\to\infty}} \frac{\eta(t)}{[2\tau(t)\ln\ln\tau(t)]^{1/2}}$$

$$\leq \overline{\lim_{t\to\infty}} J(t) + \overline{\lim_{t\to\infty}} \frac{\tilde{\xi}(\tau(t))}{[2\tau(t)\ln\ln\tau(t)]^{1/2}}.$$

Hence, using the law of the iterated logarithm, we obtain the second part of the theorem. $\qquad\square$

Remark 6.1 It is readily seen from the relation (5.30) between Itô and Stratonovich stochastic equations, that the assertions of Theorem 6.1 for the case $\tau(\infty) < \infty$ remain valid for the Stratonovich variant

$$dX(t) = b(t)X(t)\,dt + \sigma(t)X(t)\,d^*\xi(t).$$

The assertions for $\tau(\infty) = \infty$ also remain valid if the function $J(t)$ is replaced by

$$J_1(t) = \frac{\int_{t_0}^{t} b(s)\,ds}{[2\tau(t)\ln\ln\tau(t)]^{1/2}}.$$

Hence, in particular, it follows that the unstable solution $X \equiv 0$ of the equation $\dot{x} = bx$ cannot be stabilized by physically feasible (see Sect. 5.5) perturbations of the parameter b. For a constant b this was demonstrated in Sect. 5.5.

Remark 6.2 It follows from Theorem 6.1 and Remark 6.1 that random noise does not affect the stability properties of the system $\dot{x} = bx$ if $\int_0^\infty \sigma^2(s)\,ds < \infty$.

The representation (6.5) may also be used to derive conditions for *p*-stability and *q*-instability in the one-dimensional case. In fact, for any real α one easily sees from

(6.5) that

$$\mathbf{E}|X(t)|^{\alpha} = |x_0|^{\alpha} \exp\left\{\alpha \int_{t_0}^{t}\left[b(s) - \frac{\sigma^2(s)}{2}\right]ds + \frac{\alpha^2}{2}\int_{t_0}^{t}\sigma^2(s)\,ds\right\}. \quad (6.8)$$

For example, let b and σ be constants. Then, as we already have mentioned in Sect. 5.5, the necessary and sufficient condition for asymptotic stability is $b < \sigma^2/2$, and formula (6.8) can be rewritten as

$$\mathbf{E}|X(t)|^{\alpha} = |x_0|^{\alpha} \exp\left\{\alpha\left[\left(b - \frac{\sigma^2}{2}\right) + \frac{\alpha}{2}\sigma^2\right](t - t_0)\right\}.$$

Hence we see that an asymptotically stable one-dimensional linear system with constant coefficients is p-stable for sufficiently small p. Later, in Sect. 6.4 we shall see that this important property carries over to multi-dimensional systems. The analogous statement for instability does not hold. For example, if $b = \sigma^2/2 \neq 0$, the system is unstable, but it is not q-unstable for any $q > 0$.

6.2 Equations for Moments[2]

It is well known that the solution of a linear homogeneous deterministic system with constant coefficients can be determined from the roots of an auxiliary algebraic equation. Unfortunately, there is apparently no analogous reduction procedure for stochastic systems. However, as has been pointed out by many authors (see [180], [89], [90] and others), the problem of determining the moments of orders $1, 2, 3, \ldots$ can be reduced to solving an auxiliary deterministic system of linear differential equations.

Let us examine the situation more closely. Expressing the system (6.2) in integral form and calculating the conditional expectation, given $X(t_0) = x_0$, we easily derive the following equation for the vector $m_1(t) = \mathbf{E}X^{t_0, x_0}(t)$:

$$\frac{dm_1}{dt} = B(t)m_1, \qquad B(t) = ((b_i^j(t))) \tag{6.9}$$

with initial condition

$$m_1(t_0) = x_0.$$

Systems of equations for the second, third, etc. moments can be obtained by using Itô's formula (3.30).

Applying this formula to the function $x_i x_j$, we get

$$d(X_i(t)X_j(t)) = X_i(t)\,dX_j(t) + X_j(t)\,dX_i(t) + \sum_{r=1}^{k}(\sigma_r(t)X(t))_i\,(\sigma_r(t)X(t))_j\,dt$$

---

[2]The results of this section are derived from [180], [89], [90].

$$= [X_i(t)(B(t)X(t))_j + X_j(t)(B(t)X(t))_i] \, dt$$

$$+ \sum_{r=1}^{k} \{[X_i(t)(\sigma_r(t)X(t))_j + X_j(t)(\sigma_r(t)X(t))_i] \, d\xi_r(t)$$

$$+ (\sigma_r(t)X(t))_i (\sigma_r(t)X(t))_j \, dt\}.$$

Expressing this relation in integral form and taking expectations, we get the system of differential equations:

$$\frac{dm_{ij}}{dt} = \sum_{n=1}^{l} \left\{ b_i^n(t)m_{jn}(t) + b_j^n(t)m_{in}(t) \right.$$

$$\left. + \sum_{s,n=1}^{l} \sum_{r=1}^{k} \sigma_{ir}^s(t)\sigma_{jr}^n(t)m_{sn}(t) \right\} \quad (i, j = 1, \ldots, l), \qquad (6.10)$$

with the unknowns

$$m_{ij}(t) = \mathbf{E}[X_i^{t_0,x_0}(t) X_j^{t_0,x_0}(t)].$$

This system contains $n(n + 1)/2$ independent equations, since $m_{ij}(t) = m_{ji}(t)$.

The same method yields equations for

$$m_{i_1 i_2 i_3}(t) = \mathbf{E}[X_{i_1}^{t_0,x_0}(t) X_{i_2}^{t_0,x_0}(t) X_{i_3}^{t_0,x_0}(t)]$$

and so on.

Remark 6.3 Comparing (6.9) and (6.2), we see that if the system (6.2) is stable in the mean, then the deterministic system obtained by suppressing the "fluctuation" terms is stable.

Remark 6.4 Since

$$\mathbf{E}|X^{t_0,x_0}(t)|^2 = m_{11}(t) + \cdots + m_{ll}(t),$$

it follows that the system (6.2) is stable in mean square (asymptotically or exponentially) if and only if the deterministic system (6.10) is stable in the corresponding sense. If the coefficients b_i^j and σ_{ir}^j are constants, the system (6.1) is asymptotically (exponentially) stable in mean square if and only if the roots of the equation

$$\det(\lambda J - \widetilde{B}) = 0 \qquad (6.11)$$

have negative real parts. (Here $\widetilde{B}$ is the matrix of the system (6.10), J is the $l^2 \times l^2$ identity matrix.) Necessary and sufficient conditions for the real parts of the roots of the algebraic (6.11) to be negative can be found in the form of inequalities for the elements of the matrix B. However, these conditions (the Routh–Hurwitz criterion)

are quite cumbersome, since they involve computation of l^2 determinants of orders up to l^2. In Sect. 6.10 we shall see that these conditions can be simplified in a special case of practical importance.

Remark 6.5 The roots of (6.11) are continuous functions of the coefficients b_i^j and σ_{ir}^j; thus, if the system (6.2) with constant coefficients is asymptotically stable in mean square, the same holds for a system with coefficients deviating slightly from those of (6.2). A more general result can be proved by the method of Lyapunov functions (see Chap. 7).

6.3 Exponential *p*-Stability and *q*-Instability[3]

In this section we shall give some improvements and applications of the theorems of Chap. 5, proving necessary and sufficient conditions for p-stability and q-instability relative to the linear system

$$dX(t) = B(t)X(t)\,dt + \sum_{r=1}^{k} \sigma_r(t)X(t)\,d\xi_r(t). \tag{6.12}$$

We shall assume throughout that the functions $\|B(t)\|$, $\|\sigma_r(t)\|$ are bounded.

Theorem 6.2 *The solution $X(t) \equiv 0$ of the system (6.12) is exponentially p-stable if and only if there exists a function $V(t,x)$, homogeneous of degree p in x, such that for some constants $k_i > 0$*

$$k_1|x|^p \le V(t,x) \le k_2|x|^p; \qquad LV(t,x) \le -k_3|x|^p, \tag{6.13}$$

$$\left|\frac{\partial V}{\partial x_i}\right| \le k_4|x|^{p-1}; \qquad \left|\frac{\partial^2 V}{\partial x_i \partial x_j}\right| \le k_4|x|^{p-2} \quad (i,j=1,\ldots,l).$$

Proof Sufficiency follows from Theorem 5.11. To prove necessity we proceed as in the proof of Theorem 5.12, using the function

$$V(t,x) = \int_t^{t+T} \mathbf{E}|X^{t,x}(u)|^p\,du. \tag{6.14}$$

Since the coefficients of the system (6.12) have bounded derivatives of arbitrary order with respect to x, it follows from Theorem 5.12 that this function satisfies (6.13). We claim that $V(t,x)$ is homogeneous of degree p. In fact, (6.3) implies that

---

[3]The main theorems of this section are due to Nevelson and Khasminskii [221].

the solution $X^{t,x}(u)$ admits the representation

$$X^{t,x}(u) = \sum_{i=1}^{l} x_i X^{(i)}(u),\tag{6.15}$$

where $(x_1, \ldots, x_l) = x$ and $(X^{(1)}(u), \ldots, X^{(l)}(u))$ is a fundamental system of solutions of (6.12) satisfying the initial conditions $X^i_j(t) = \delta^i_j$ (δ^i_j is the Kronecker symbol). Substituting (6.15) into (6.14), we see that $V(t, x)$ is homogeneous of degree p.   □

In applications of Theorem 6.2, one would like to be sure that the p-stability of a stochastic system can be "detected" with the aid of homogeneous functions from a relatively small class. For the general case we have unfortunately no results of this kind. However, for even p the following theorem is valid.

Theorem 6.3 *A necessary condition for exponential p-stability of even order ($p = 2, 4, \ldots$) of the system* (6.12) *is that for every positive definite form $W(t, x)$ of degree p whose coefficients are continuous bounded functions of time there exist a positive definite form $V(t, x)$ of the same degree such that*

$$LV = -W.$$

The same condition, with the phrase "for every ..." replaced by "for some ...", is also sufficient.

The proof is analogous to that of Theorem 6.2; the only difference is that instead of the function (6.14) one considers the function

$$V_1(t, x) = \mathbf{E} \int_t^{\infty} W(u, X^{t,x}(u))\, du.$$

The infinite upper limit of integration here causes no difficulties, since the required smoothness of $V_1(t, x)$ as a function of x follows from the fact that this function is a form in x.

If the matrices B and σ_r in (6.12) are constant, so that we have a linear autonomous system

$$dX(t) = BX(t)\, dt + \sum_{r=1}^{k} \sigma_r X(t)\, d\xi_r(t),\tag{6.16}$$

then the forms $V(t, x)$ and $W(t, x)$ in the statement of Theorem 6.3 may be replaced by forms $V(x)$ and $W(x)$ with constant coefficients. In this case, Theorem 6.3 yields the following algorithm for a construction of algebraic criteria for p-stability for even p. Given some positive definite form $W(x)$ of degree p, we look for a form $V(x)$ of the same degree such that $LV = -W(x)$. Comparing the coefficients at the

monomials $x_1^{k_1} \cdots x_l^{k_l}$ on both sides of this equation ($k_1 + \cdots + k_l = p$), we get a system of linear equations for the coefficients of $V(x)$. It follows from Theorem 6.3 that the system is p-stable if and only if the function $V(x)$ turns out to be positive definite.

This procedure is well known for deterministic systems (see [45]). Its applicability to stability in mean square of stochastic systems was first indicated by Kac and Krasovskii [111].

Theorem 6.4 *The solution* $X(t) \equiv 0$ *of the system* (6.12) *is exponentially* q-*unstable if and only if there exists a function* $V(t, x)$, *homogeneous in* x *of degree* $-q$, *such that for some constants* $k_i > 0$ *we have*

$$k_1|x|^{-q} \le V(t, x) \le k_2|x|^{-q}; \qquad LV(t, x) \le -k_3|x|^{-q};$$

$$\left|\frac{\partial V}{\partial x_i}\right| \le k_4|x|^{-q-1}; \qquad \left|\frac{\partial^2 V}{\partial x_i \partial x_j}\right| \le k_4|x|^{-q-2} \quad (i, j = 1, \ldots, l).$$

The reader should have no difficulty in proving this theorem, using the Theorems 5.13 and 5.14.

A deterministic stationary linear system $dx/dt = Bx$ is q-unstable for any q if and only if all the roots λ_i of the characteristic equation $\det(\lambda J - B) = 0$ have positive real parts. This is easily checked directly, but we shall view this as one of the consequences of Theorem 6.4.

It is well known that the existence of a positive definite quadratic form $W(x)$ such that the form $d^0W/dt = LW$ is also positive definite is a necessary and sufficient condition for the numbers $\operatorname{Re} \lambda_i$, $i = 1, \ldots, l$, to be positive. Set $V(x) = [W(x)]^{-q/2}$. It is readily seen that this function satisfies all the assumptions of Theorem 6.4, and this implies the above assertion.

The following two examples illustrate applications of the theorems proved in this section.

Example 6.1 The generator of the one-dimensional system

$$dX(t) = bX(t)\, dt + \sigma X(t)\, d\xi(t)$$

with constants b and σ is equal to

$$L = bx\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2}.$$

Setting $V(x) = |x|^p$, we get

$$LV = p|x|^{p-1}\left[b + \frac{\sigma^2}{2}(p - 1)\right].$$

It follows from Theorem 6.3 that in this case the inequality $b + \frac{1}{2}\sigma^2(p - 1) < 0$ is a necessary and sufficient condition for exponential p-stability. Of course, this conclusion also follows from the explicit expression (6.5).

Example 6.2 Recall that the expression

$$L = \frac{\partial}{\partial t} + \left(Bx, \frac{\partial}{\partial x} \right) + \frac{1}{2} \left(A(x) \frac{\partial}{\partial x}, \frac{\partial}{\partial x} \right)$$

$$\equiv \frac{\partial}{\partial t} + \left(Bx, \frac{\partial}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r x, \frac{\partial}{\partial x} \right)^2$$

defines the generator of the system (6.16). We now consider some sufficient conditions for p-stability of this system, restricting ourselves to Lyapunov functions which are powers of a positive definite quadratic from $(Wx, x) = \sum W_{ij} x_i x_j$. We introduce the notation: $V(x) = (Wx, x)^{p/2}$, B^* is the matrix adjoint to B, $\lambda_{\min}^D = \lambda_1^D < \lambda_2^D < \cdots < \lambda_l^D = \lambda_{\max}^D$ are the eigenvalues of a symmetric $l \times l$ matrix D. In addition, we set

$$m = \inf_{|x|=1} \lambda_{\min}^{A(x)}; \qquad M = \sup_{|x|=1} \lambda_{\max}^{A(x)}.$$

We first observe that for any positive semi-definite symmetric matrices D_1 and D_2 we have

$$\lambda_{\min}^{D_1} \operatorname{tr} D_2 < \operatorname{tr}(D_1 D_2) < \lambda_{\max}^{D_1} \operatorname{tr} D_2. \tag{6.17}$$

This inequality is easily proved by reducing D_1 to diagonal form.

Next,

$$LV = p(Wx, x)^{p/2-1} \left[\frac{1}{2} \operatorname{tr}(A(x)W) + ((WB + B^*W)x, x) \right.$$

$$\left. + \left(\frac{p}{2} - 1 \right) \operatorname{tr}(A(x)F(x)) \right]$$

$$= p(Wx, x)^{p/2-1} \Phi(x) \tag{6.18}$$

holds, where $F(x) = ((f_{ij}(x)))$ is the matrix with the elements

$$f_{ij}(x) = \frac{(Wx_i)(Wx_j)}{(Wx, x)}.$$

It follows from the preceding theorem that a sufficient condition for the linear system (6.16) to be exponentially p-stable is that the expression $\Phi(x)$ (in square brackets in (6.18)) is negative for $x \neq 0$. But if $\Phi(x) \geq 0$, then (see Remark 5.8) the system is not exponentially p-stable.

It follows from (6.17) and (6.18) that for $p \geq 2$

$$\Phi(x) \leq \frac{1}{2} \lambda_{\max}^{A(x)} \operatorname{tr} W + ((WB + B^*M)x, x) + \left(\frac{p}{2} - 1\right) \lambda_{\max}^{A(x)} \frac{(W^2 x, x)}{(Wx, x)}$$

$$\leq \left[M\left(\frac{1}{2} \operatorname{tr} W + \left(\frac{p}{2} - 1\right) \lambda_{\max}^{W}\right) + \lambda_{\max}^{WB + B^*W} \right](x, x), \qquad (6.19)$$

$$\Phi(x) \geq \left[m\left(\frac{1}{2} \operatorname{tr} W + \left(\frac{p}{2} - 1\right) \lambda_{\min}^{W}\right) + \lambda_{\min}^{WB + B*W} \right](x, x). \qquad (6.20)$$

Hence we see that if for some $p \geq 2$ there exists a positive definite matrix W such that

$$M\left[\frac{1}{2} \operatorname{tr} W + \left(\frac{p}{2} - 1\right) \lambda_{\max}^{W} \right] + \lambda_{\max}^{WB + B^*W} < 0, \qquad (6.21)$$

then the system is exponentially p-stable for this p. On the other hand, if

$$m\left[\frac{1}{2} \operatorname{tr} W + \left(\frac{p}{2} - 1\right) \lambda_{\min}^{W} \right] + \lambda_{\min}^{WB + B^*W} \geq 0, \qquad (6.22)$$

then the system is not exponentially p-stable.

Similarly, if for some $p \leq 2$ there exists a positive definite matrix W such that

$$\frac{1}{2} M \operatorname{tr} W + m\left(\frac{p}{2} - 1\right) \lambda_{\min}^{W} + \lambda_{\max}^{WB + B^*W} < 0, \qquad (6.23)$$

then the system (6.12) is exponentially p-stable, while if

$$\frac{1}{2} m \operatorname{tr} W + M\left(\frac{p}{2} - 1\right) \lambda_{\max}^{W} + \lambda_{\min}^{WB + B^*W} \geq 0, \qquad (6.24)$$

the system is not exponentially p-stable.

Although the conditions furnished by (6.21) through (6.24) are rather weak, we shall show now that in a certain special case they yield necessary and sufficient conditions for exponential p-stability. Let $A(x) = \delta |x|^2 J$ (where J is the identity matrix). Then $M = m = \delta$.

Now suppose that $B + B^* = -\lambda J$, and set $W = J$. It then follows from (6.19)–(6.24) that this system is asymptotically p-stable if and only if $\frac{1}{2} \delta(l + p - 2) < \lambda$. An analogous argument shows that the system is q-unstable if and only if $\frac{1}{2} \delta(l - q - 2) > \lambda$ for some positive q.

6.4 Exponential p-Stability and q-Instability (Continued)

It is well known that an asymptotically stable linear deterministic system with constant coefficients is exponentially stable. This statement holds true for a system with

variable coefficients which is uniformly (in time) asymptotically stable. In this section we shall prove analogs of these properties for linear stochastic systems with constant coefficients. Systems with variable coefficients will be considered in the next section.

Lemma 6.2 *If a linear system with constant coefficients*

$$dX(t) = BX(t)\,dt + \sum_{r=1}^{k} \sigma_r X(t)\,d\xi_r(t) \tag{6.25}$$

is stable in probability, then it is p-stable for sufficiently small p.

Proof Since the system (6.25) is stable in probability, there exists $\alpha > 0$ such that

$$\sup_{|x| \le 2^{-\alpha}} \mathbf{P}\Big\{ \sup_{t \ge 0} |X^x(t)| > 1 \Big\} \le \frac{1}{2}.$$

Further, since the system is linear, we have

$$X^{\gamma x}(t) = \gamma X^x(t). \tag{6.26}$$

Therefore, for any k,

$$\sup_{|x| \le 2^{k\alpha}} \mathbf{P}\Big\{ \sup_{t \ge 0} |X^x(t)| > 2^{\alpha(k+1)} \Big\} \le \frac{1}{2}. \tag{6.27}$$

Let τ denote the first time at which the path of the process reaches the set $|x| = 2^\alpha$. Using the strong Markov property of the process $X(t)$ and (6.27), we get

$$\sup_{|x| \le 1} \mathbf{P}\Big\{ \sup_{t \ge 0} |X^x(t)| > 2^{2\alpha} \Big\}$$

$$= \sup_{|x| \le 1} \int_{u=0}^{\infty} \int_{|y|=2^\alpha} \mathbf{P}\{\tau \in du,\, X^x(u) \in dy\} \mathbf{P}\Big\{ \sup_{t \ge 0} |X^x(t)| > 2^{2\alpha} \Big\}$$

$$\le \frac{1}{2} \sup_{|x| \le 1} \mathbf{P}\{\tau < \infty\} = \frac{1}{2} \sup_{|x| \le 1} \mathbf{P}\Big\{ \sup_{t > 0} |X^x(t)| > 2^\alpha \Big\} \tag{6.28}$$

$$< \frac{1}{2^2}$$

$$\vdots$$

$$\sup_{|x| \le 1} \mathbf{P}\Big\{ \sup_{t \ge 0} |X^x(t)| > 2^{k\alpha} \Big\} < \frac{1}{2^k}.$$

Now let $x \in \mathbb{R}^l$ be such that $|x| = 1$. Then, using (6.28) with $p < 1/\alpha$, we have

$$\mathbf{E}\Big[\sup_{t > 0} |X^x(t)|^p \Big] \le \sum_{k=1}^{\infty} 2^{k\alpha p} \mathbf{P}\Big\{ 2^{(k-1)\alpha} < \sup_{t > 0} |X^x(t)| < 2^{k\alpha} \Big\}$$

$$\leq \sum_{k=1}^{\infty} 2^{k\alpha p} 2^{-(k-1)} = 2 \sum_{k=1}^{\infty} 2^{-k(1-\alpha p)} = K(p)$$

$$< \infty. \tag{6.29}$$

It follows from (6.29) and (6.26) that

$$\sup_{|x|<\delta} \mathbf{E}\left[\sup_{t>0} |X^x(t)|^p \right] \leq \delta^p K(p),$$

and this inequality implies the assertion.   □

Lemma 6.3 *If the system* (6.25) *is asymptotically stable in probability, then it is asymptotically p-stable for sufficiently small* p.

Proof Let x be such that $|x| = 1$. Then (6.29) implies that

$$\mathbf{E}\left\{ \sup_{t\geq 0} |X^x(t)|^p \right\} < K \tag{6.30}$$

holds for some $p > 0$.

Moreover, as was mentioned in Sect. 6.1, a linear system is asymptotically stable in probability if and only if it is asymptotically stable in the large. Therefore, almost surely

$$X^x(t) \to 0 \quad \text{as } t \to \infty. \tag{6.31}$$

It now follows from (6.30) and (6.31) by Lebesgue's bounded convergence theorem [97, Sect. 5.26] that

$$\mathbf{E}|X^x(t)|^p \to 0 \quad \text{as } t \to \infty.$$

This and Lemma 6.2 imply the assertion.   □

Lemma 6.4 *If the system* (6.25) *is asymptotically p-stable for some* p, *then it is also exponentially p-stable.*

Proof [4] We first show that, under the assumptions of the lemma, for every $Q < 1$ there exists a $T > 0$ such that

$$\sup_{|x|=1} \mathbf{E}|X^x(T)|^p \leq Q. \tag{6.32}$$

To do this we use the representation (6.15). It follows from the assumptions that for each $i = 1, \ldots, l$ and sufficiently large T

$$\mathbf{E}|X^{(i)}(T)| < \frac{Q}{l^{p+1}}.$$

---

[4]The proof of Lemma 6.4 essentially uses the same idea as that of Theorem 6.1 in [111].

Next, using the inequality

$$|A_1 + \cdots + A_l|^p \le l^p(|A_1|^p + \cdots + |A_l|^p)$$

and (6.15), we see that if $|x| = 1$, then

$$\mathbf{E}|X^x(T)|^p \le l^p \sum_{i=1}^{l} \mathbf{E}|X^{(i)}(T)|^p \le Q.$$

We now choose T so that (6.32) holds with $Q = e^{-1}$ and, in view of (6.26), we rewrite (6.32) as

$$\mathbf{E}|X^x(T)|^p \le e^{-1}|x|^p. \tag{6.33}$$

Then

$$
\begin{aligned}
\mathbf{E}|X^x(2T)|^p &= \int_{\mathbb{R}^l} P(x, T, dy)\mathbf{E}|X^y(T)|^p \\
&\le e^{-1} \int_{\mathbb{R}^l} P(x, T, dy)|y|^p = e^{-1}\mathbf{E}|X^x(T)|^p \\
&\le e^{-2}|x|^p
\end{aligned}
\tag{6.34}
$$

$$\vdots$$

$$\mathbf{E}|X^x(kT)|^p \le e^{-k}|x|^p.$$

Let $t = nT + t_1$ ($0 \le t_1 < T$), and let

$$K = \sup_{t>0,\ |x|=1} \mathbf{E}|X^x(t)|^p.$$

Here $K < \infty$ by virtue of Lemma 6.2. This and (6.34) imply

$$\mathbf{E}|X^x(p)|^p = \int_{\mathbb{R}^l} P(x, nT, dy)\mathbf{E}|X^y(t_1)|^p$$

$$\le K\mathbf{E}|X^x(nT)|^p \le K|x|^p e^{-n} \le Ke|x|^p e^{-t/T} = K_1|x|^p e^{-t/T}.$$

This completes the proof. $\qquad\qquad\square$

Lemma 6.3 and 6.4 immediately imply

Theorem 6.5 *If the linear system* (6.25) *with constant coefficients is asymptotically stable in probability, then it is exponentially p-stable for all sufficiently small positive p.*

Similar arguments hold for q-instability. Thus one easily proves the following

Theorem 6.6 *If the solutions of the linear system* (6.25) *with constant coefficients satisfy the relation*

$$\mathbf{P}\{|X^x(t)| \to \infty \ as \ t \to \infty\} = 1$$

for $x \neq 0$, *then the system is exponentially* q-*unstable for all sufficiently small positive* q.

Theorems 6.5 and 6.6 fail to hold for linear systems with variable coefficients. For example, the deterministic system $dx/dt = -x/(t+1)$ is asymptotically stable but not exponentially stable. However, if certain additional assumptions are made, then one can prove analogs of Lemmas 6.2 and 6.4, hence also of Theorem 6.5. But first we have to introduce new definitions of stability and instability and then study the properties of systems satisfying these definitions. This we do in the next section.

6.5 Uniform Stability in the Large

The solution $X(t) \equiv 0$ of the system

$$dX(t) = b(t, X(t)) \, dt + \sum_{r=1}^{k} \sigma_r(t, X(t)) \, d\xi_r(t) \tag{6.35}$$

is said to be *stable in the large uniformly in* $t > 0$, if it is uniformly stable in probability and moreover for any $x \in \mathbb{R}^l$, $\varepsilon > 0$,

$$\sup_{s>0} \mathbf{P}\left\{ \sup_{u>s+T} |X^{s,x}(u)| > \varepsilon \right\} \xrightarrow[T \to \infty]{} 0. \tag{6.36}$$

Let us give a few comments on this definition.

1. It follows from (6.36) that the system (6.35) is stable in the large in the sense of the definition of Sect. 5.4. In fact, the equivalence of the events

$$\left\{ \bigcap_{n=1}^{\infty} \sup_{u>s+n} |X^{s,x}(u)| > \varepsilon \right\} = \left\{ \overline{\lim_{t \to \infty}} |X^{s,x}(t)| > \varepsilon \right\}$$

and (6.36) imply that

$$\mathbf{P}\left\{ \overline{\lim_{t \to \infty}} |X^{s,x}(t)| > \varepsilon \right\} = \lim_{n \to \infty} \mathbf{P}\left\{ \sup_{u>s+n} |X^{s,x}(u)| > \varepsilon \right\} = 0.$$

Since ε is arbitrary, this shows that the system is stable in the large.

2. A sufficient condition for the solution $X(t) \equiv 0$ of the system (6.35) to be stable in the large is that it is uniformly stable in probability and, moreover, for any $x \in \mathbb{R}^l$, $\varepsilon > 0$,

$$\sup_{s>0} \mathbf{P}\{|X^{s,x}(s+T)| > \varepsilon\} \xrightarrow[T \to \infty]{} 0. \tag{6.37}$$

This evidently follows from the inequalities

$$\mathbf{P}\left\{ \sup_{u>s+T} |X^{s,x}(u)| > \varepsilon \right\}$$

$$= \left(\int_{|y|<\delta} + \int_{|y|\geq\delta} \right) P(s,x,s+T,dy)\mathbf{P}\left\{ \sup_{u>s+T} |X^{s+T,y}(u)| > \varepsilon \right\}$$

$$\leq \sup_{s>0,\,|y|<\delta} \mathbf{P}\left\{ \sup_{u>s} |X^{s,y}(u)| > \varepsilon \right\} + \mathbf{P}\{|X^{s,x}(s+T)| \geq \delta\},$$

where the first and second terms on the right can be made arbitrarily small by a suitable choice of δ and T.

Yet another sufficient condition for uniform asymptotic stability in the large is given by

Lemma 6.5 *A sufficient condition for the solution $X(t) \equiv 0$ of (6.35) to be stable in the large uniformly in $t > 0$ is that it be uniformly stable in probability, and that the family of processes $X^{s,x}(i)$ associated with different values of the parameters s, x is for any $\varepsilon > 0$ uniformly recurrent relative to the domain $|x| < \varepsilon$, in the sense that*

$$\sup_{s>0} \mathbf{P}\{\tau_\varepsilon^{s,x} - s > T\} \to 0 \quad as\ T \to \infty, \tag{6.38}$$

where $\tau_\varepsilon^{s,x}$ is the first time the path of the process $X^{s,x}(t)$ reaches the set $|x| = \varepsilon$.

Proof It is clear that for any $\varepsilon > 0$, δ and $T > 0$, we have

$$\left\{ \sup_{u>s+T} |X^{s,x}(u)| > \varepsilon \right\} \subset \{\tau_\delta^{s,x} > s+T\} \cup \left\{ \tau_\delta^{s,x} \leq s+T;\ \sup_{u>\tau_\delta^{s,x}} |X^{s,x}(u)| > \varepsilon \right\}.$$

Hence, applying the strong Markov property, we get

$$\mathbf{P}\left\{ \sup_{u>s+T} |X^{s,x}(u)| > \varepsilon \right\}$$

$$\leq \mathbf{P}\{\tau_\delta^{s,x} - s > T\}$$

$$+ \int_{v=s}^{s+T} \int_{|y|=\delta} \mathbf{P}\{\tau_\delta^{s,x} \in ds,\, X^{s,x}(\tau_\delta^{s,x}) \in dy\}\mathbf{P}\left\{ \sup_{u>v} |X^{v,y}(u)| > \epsilon \right\}$$

$$\leq \mathbf{P}\{\tau_\delta^{s,x} - s > T\} + \sup_{v>0,\,|y|=\delta} \left\{ \sup_{u>v} \mathbf{P}|X^{v,y}(u)| > \varepsilon \right\}.$$

For any $\varepsilon > 0$, the second term on the right of this inequality can be made arbitrarily small by suitable choice of δ, and the first, by a choice of T (see (6.38)). This proves the lemma. $\square$

Remark 6.6 For condition (6.38) to hold, it is sufficient that there exists a positive function $V(t, x)$ in the domain $|x| > \varepsilon$ such that

$$\left.\begin{array}{c} \inf\limits_{|x|>R,\, t>0} V(t, x) = V_R \to \infty \quad \text{as } R \to \infty, \\ \sup\limits_{\varepsilon<|x|<\delta,\, t>0} V(t, x) = V^{(\delta)} < \infty, \end{array}\right\} \tag{6.39}$$

$$LV < -k \quad (k > 0). \tag{6.40}$$

In fact, conditions (6.39) and (6.40) imply by Theorem 3.5 that the process $X(t)$ is regular. Applying Theorem 3.9, we get the inequality

$$\mathbf{E}(\tau_\delta^{s,x} - s) < \frac{V(s, x)}{k} \le \frac{V^{(|x|)}}{k}.$$

Hence, using Chebyshev's inequality, we get (6.38).

This remark and the Remark 5.3 imply

Lemma 6.6 *A sufficient condition for the solution $X(t) \equiv 0$ of (6.35) to be stable in the large uniformly in $t > 0$ is that there exists a positive definite function $V(t, x) \in \mathbf{C}_2^0(E)$, with infinitesimal upper limit, such that the function LV is negative definite and (6.39) holds.*

Comparing this lemma with Theorem 5.8, we see that a system is uniformly stable in the large if there exists a Lyapunov function V satisfying the assumptions of Theorem 5.8 and such that $\sup_{t>0} V(t, x)$ is bounded in any bounded (with respect to x) domain.

We now state the analogs of Theorem 6.5 and Lemma 6.2 for nonstationary linear systems.

Theorem 6.7 *If the linear system*

$$dX(t) = B(t)X(t)\, dt + \sum_{r=1}^{k} \sigma_r(t)X(t)\, d\xi_r(t) \tag{6.41}$$

is stable in the large uniformly in $t > 0$, then it is exponentially p-stable for sufficiently small positive p.

Lemma 6.7 *If the system (6.41) is uniformly stable in probability, then it is p-stable for sufficiently small p, and there exists an $\alpha > 0$ such that for all $k = 1, 2, \ldots$*

$$\sup_{s>0,\, |x|\le 1} \mathbf{P}\left\{ \sup_{t\le s} |X^{s,x}(t)| > 2^{k\alpha} \right\} \le \frac{1}{2^k}. \tag{6.42}$$

The proof of inequality (6.42) is an almost word for word repetition of that of inequality (6.28), except that properties of uniform stability are used. From (6.42)

we get the first assertion of the lemma, in the same way as Lemma 6.2 was deduced from (6.28).

Proof of Theorem 6.7 It will suffice to verify that there exists a $T > 0$ such that for all positive $p < \alpha$

$$\sup_{s>0,\,|x|=1} \mathbf{E}|X^{s,x}(s+T)|^p < 1. \tag{6.43}$$

Indeed, the assertion will then follow from (6.43) in the same way as Lemma 6.3 follows from (6.32).

Now, proceeding as in the case of (6.29), we see that for arbitrary $\alpha > 0$, $n > 0$, $p > 0$

$$\mathbf{E}|X^{s,x}(s+T)|^p \leq \mathbf{E}\left[\sup_{t\geq s+T} |X^{s,x}(t)|^p\right]$$

$$\leq \sum_{k=-\infty}^{+\infty} 2^{k\alpha p}\mathbf{P}\left\{2^{(k-1)\alpha} \leq \sup_{t\geq s+T} |X^{s,x}(t)| < 2^{k\alpha}\right\}$$

$$< 2^{-\alpha p} + 2^{n\alpha p}\mathbf{P}\left\{\sup_{t>s+T} |X^{s,x}(t)| > 2^{-\alpha}\right\}$$

$$+ \sum_{k=n+1}^{\infty} 2^{k\alpha p}\left\{\sup_{t>s}|X^{s,x}(t)| > 2^{a(k-1)}\right\}.$$

If we now select α as in Lemma 6.7 and use (6.42), we get the inequality

$$\mathbf{E}|X^{s,x}(s+T)|^p \leq 2^{-n\alpha p} + 2^{n\alpha p}\mathbf{P}\left\{\sup_{t>s+T} |X^{s,x}(t)| > 2^{-\alpha}\right\} + 2\sum_{k=n+1}^{\infty} 2^{k(\alpha p-1)}.$$

Now choose $p < 1/\alpha$, then let n be large enough, so that

$$2\sum_{k=n+1}^{\infty} 2^{k(\alpha p-1)} < \frac{1}{2}(1 - 2^{-\alpha p}).$$

Finally take T large enough, so that

$$2^{n\alpha p}\sup_{s>0,\,|x|=1} \mathbf{P}\left\{\sup_{t>s+T} |X^{s,x}(t)| > 2^{-\alpha}\right\} < \frac{1}{2}(1 - 2^{-\alpha p})$$

(such choice of T is possible by (6.36)). This yields (6.43) and hence the assertion of the theorem. $\qquad\square$

To conclude this section we state the analogous instability theorem (the proof is similar).

Theorem 6.8 *If the linear system* (6.41) *is uniformly unstable in the sense that for any $x \neq 0$, $A > 0$*

$$\sup_{s>0} \mathbf{P}\left\{ \sup_{u>s+T} |X^{s,x}(u)| < A \right\} \xrightarrow[T\to\infty]{} 0,$$

then the system is exponentially q-unstable for sufficiently small positive q.

6.6 Stability of Products of Independent Matrices

Let $A_n = ((a_{ij}^{(n)}))$ $(i, j = 1, \ldots, l;\ n = 1, 2, \ldots)$ be a sequence of identically distributed[5] $l \times l$ matrices and let $P(dA)$ be their common probability distribution. It is easy to see that the sequence

$$x_0;\quad x_1 = A_1 x_0;\quad \ldots;\quad x_n = A_n x_{n-1},\quad \ldots, \tag{6.44}$$

where $x_0 \in \mathbb{R}^l$, is a Markov chain in $\mathbb{R}^l$. In applications (see [112]) one is often interested to find conditions under which $|x_n| = |A_n A_{n-1} \ldots A_1 x_0| \to 0$ for all $x_0 \in \mathbb{R}^l$ (or, what is the same, $\|A_n A_{n-1} \ldots A_1\| \to 0$ in some sense) as $n \to \infty$. In Chap. 1 (Example 1.11) we showed that in certain cases the stability theory of linear systems whose coefficients are step functions can be reduced to this problem.

The solution of this problem will provide yet another illustration of how to apply the methods of this chapter to a discrete model.

1. We first consider the trivial case $l = 1$.

Then

$$|x_n| = |x_0| \exp\left\{ \sum_{i=1}^{n} \ln|A_i| \right\}. \tag{6.45}$$

If $a = \mathbf{E} \ln|A_i|$ exists, then it follows from the law of large numbers that $|x_n| \to 0$ almost surely when $a < 0$, and $|x_n| \to \infty$ when $a > 0$. For $a = 0$ we may have either stability (e.g., if $|A_n| = 1$ almost surely) or instability. However, if the random variable $\ln|A_i|$ has finite non-zero variance, one readily shows that

$$\overline{\lim_{n\to\infty}} |x_n| = \infty, \quad \text{but } \underline{\lim_{n\to\infty}} |x_n| = 0$$

almost surely.

It is also obvious that $\mathbf{E}|x_n|^p \to 0$ if and only if $\mathbf{E}|A_i|^p < 1$. The condition for (exponential) q-instability is similar.

---

[5]That is to say, the joint distributions of the l^2 random variables $a_{ij}^{(n)}$ $(i, j = 1, \ldots, l)$ are independent for different n. The distributions of the elements of the same matrix may of course be dependent.

2. Since the matrix A_n and the vector x_{n-1} are independent, it follows from (6.44) that

$$\mathbf{E}x_n = \mathbf{E}A_n\mathbf{E}x_{n-1} = (\mathbf{E}A)^n x_0.$$

Clearly, $\mathbf{E}x_n \to 0$ if all the eigenvalues of the constant matrix $\mathbf{E}A$ are such that $|\lambda_i| < 1$.

Similar reasoning gives recursive relations for moments of higher orders. This was pointed out by Bellman [25], who also observed that in the determination of the k-th moment, $\mathbf{E}A$ is replaced by the expectation of the k-fold direct product of the matrix A itself. For example, it follows that x_n is asymptotically stable in mean square if and only if the roots of the matrix $\mathbf{E}[A \times A]$ lie inside the unit disk. (Recall that

$$A \times A = \begin{pmatrix} a_{11}A & \dots & a_{1l}A \\ \dots & \dots & \dots \\ a_{l1}A & \dots & a_{ll}A \end{pmatrix}$$

is an $l^2 \times l^2$ matrix.)

3. It is not hard to prove analogs of the theorems in Sect. 6.3 for products of independent identically distributed matrices $A_1, \dots, A_n, \dots$. We present two theorems of this type, where p is any non-zero real number.

Theorem 6.9 *A sufficient condition for*

$$\mathbf{E}\|A_n \dots A_1\|^p \to 0 \quad as \ n \to \infty \tag{6.46}$$

to hold is that there exists a positive definite function $f(x)$, homogeneous of degree p, such that the function $\mathbf{E}f(Ax) - f(x)$ is negative definite.

Proof The assumptions imply that there exist positive constants k_i $(k_3 > k_1)$ for which

$$\mathbf{E}f(Ax) - f(x) \le -k_1|x|^p; \qquad k_2|x|^p < f(x) < k_3|x|^p. \tag{6.47}$$

Hence, setting $q = 1 - (k_1/k_3)$, we get

$$f(Ax) \le f(x)\left(1 - \frac{k_1}{k_3}\right) = qf(x). \tag{6.48}$$

Let $P(x, n, \Gamma)$ denote the transition probability of the Markov chain (6.44). Then, applying (6.48), we get

$$\mathbf{E}f(A_n A_{n-1} \dots A_1 x_0) = \mathbf{E}f(A_n x_{n-1}) = \int P(x_0, n-1, dx)\mathbf{E}f(Ax)$$

$$\le q\mathbf{E}f(A_{n-1} \dots A_1 x_0) \le \dots \le q^n f(x_0),$$

whence, by (6.47) we may conclude that

$$\mathbf{E}|A_n A_{n-1} \ldots A_1 x_0|^p \leq \frac{k_3}{k_2} q^n |x_0|^p. \tag{6.49}$$

Thus,

$$\mathbf{E}\|A_n A_{n-1} \ldots A_1\|^p \leq \frac{k_3}{k_2} q^n.$$

This proves the theorem. $\square$

Remark 6.7 Examining the above proof we easily see that if we replace the condition that the function $\mathbf{E}f(Ax) - f(x)$ be negative definite by the inequality $\mathbf{E}f(Ax) - f(x) \geq 0$, we can prove the existence of a constant k such that

$$\mathbf{E}\|A_n A_{n-1} \ldots A_1\|^p > k > 0 \quad (n = 1, 2, \ldots).$$

Theorem 6.10 *If $A_1, A_2, \ldots$ is a sequence of independent identically distributed matrices satisfying condition (6.46), then, for any positive definite function $g(x)$ which is homogeneous of degree p, there exists a positive definite function $f(x)$, homogeneous of the same degree, such that*

$$\mathbf{E}f(Ax) - f(x) = -g(x). \tag{6.50}$$

The proof is analogous to that of Theorem 6.2. One first applies (6.46) to establish (6.49). Then, using (6.49), one readily shows that (6.50) is satisfied by the function

$$f(x) = g(x) + \sum_{i=1}^{\infty} \mathbf{E}g(A_i A_{i-1} \ldots A_1 x) = g(x) + \sum_{i=1}^{\infty} \mathbf{E}g(x_i)$$

which is homogeneous of degree p. Setting $f(x) = (Fx, x)$, $g(x) = (Gx, x)$ in Theorems 6.9 and 6.10, we get

Corollary 6.1 *A necessary condition for the process x_n defined by (6.44) to be asymptotically stable in mean square is that, for every positive definite matrix G, the solution F of the equation (here A^* is the transposition of A)*

$$\mathbf{E}(A^* F A) - F = -G$$

be a positive definite matrix. The same condition with the phrase "for every ..." replaced by "for some ..." is also sufficient.

Corollary 6.2 *If the assumptions of Theorem 6.9 hold with $p > 0$, then almost surely*

$$\|A_n A_{n-1} \ldots A_1\| \to 0 \quad as\ n \to \infty. \tag{6.51}$$

On the other hand, if the assumptions hold for $p < 0$, then almost surely

$$\|A_n A_{n-1} \ldots A_1\| \to \infty \quad as \; n \to \infty. \tag{6.52}$$

To prove this, we observe that by the assumptions of the theorem the sequence $f(x_n) = f(A_n \ldots A_1 x_0)$ is a positive supermartingale. By Theorem 5.1 the limit

$$\lim_{n \to \infty} f(A_n A_{n-1} \ldots A_1 x_0) = \xi \quad (a.s.),$$

is finite and $\mathbf{E}\xi \leq \lim_{n \to \infty} \mathbf{E}f(x_n) = 0$. Consequently, $\xi = 0$ almost surely. This at once yields both assertions (6.51) and (6.52).

Example 6.3 Let G be the group of orthogonal real $l \times l$ matrices and let $g \in G$, $h \in G$. It is well known (see [97]) that there exists a unique Borel measure μ (Haar measure) such that $\mu(G) = 1$ and, for any nonempty open set U and any $g \in U$, we have $\mu(U) > 0$, $\mu(gU) = \mu(Ug) = \mu(U)$. The integral with respect to μ is invariant in the sense that for any $h \in G$

$$\int_G f(g)\mu(dg) = \int_G f(gh)\mu(dg) = \int_G f(hg)\mu(dg). \tag{6.53}$$

We now assume that $A_i = B_1 g_i B_2$, where B_1 and B_2 are fixed nonsingular matrices and $g_i \in G$ are independent matrices distributed over G in accordance with the Haar measure. We shall determine sufficient conditions for (6.51) and (6.52) to hold in this case.

Set $f(x) = |B_2 x|^p$. Then $f(x)$ is obviously a positive definite and homogeneous function of degree p satisfying

$$\mathbf{E}f(Ax) = \int_G |B_2 B_1 g B_2 x|^p \mu(dg).$$

Let e denote some fixed unit vector in $\mathbb{R}^l$, say $e = (1, 0, \ldots, 0)$. Then it is obvious that for any $x \neq 0$ there exists a matrix $g_0 \in G$ such that $g_0 e = B_2 x / |B_2 x|$. Using this and (6.53), we get

$$\mathbf{E}f(Ax) - f(x) = |B_2 x|^p \left(\int_G |B_2 B_1 g g_0 e|^p \mu(dg) - 1 \right).$$

Applying Theorem 6.9 and Corollary 6.2, we get the following result.
If for some $p > 0$

$$\int |B_2 B_1 g e|^p \mu(dg) - 1 < 0, \tag{6.54}$$

then

$$\|A_n A_{n-1} \ldots A_1\| = \|B_1 g_n B_2 B_1 g_{n-1} B_2 \ldots B_1 g_1 B_2\| \to 0 \quad (a.s.).$$

Similarly, a sufficient condition for $\|A_n \ldots A_1\| \to \infty$ to hold almost surely as $n \to \infty$ is that (6.54) holds for some $p < 0$.

This result may be given in a more convenient form. Using the expansion of a^p in powers of p and the inequalities

$$0 < k_1 < |B_2 B_1 g e| < k_2,$$

we easily see that condition (6.54) holds for sufficiently small $p > 0$ if

$$I = \int_G \ln |B_2 B_1 g e| \mu(dg) < 0, \tag{6.55}$$

and for sufficiently small $p < 0$ if $I > 0$. Thus the condition $I < 0$ guarantees that the chain (6.44) of this example is stable; the condition $I > 0$, implies that it is unstable.[6]

It is natural to call $I = 0$ the "critical" case. The matrix product is then either unstable or nonasymptotically stable (the latter possibility occurs, for instance, if $B_i \in G$).

It is evident that in this example we may replace G by a subgroup of G or by the group of unitary matrices.

In particular, let us consider the group of rotations of the plane. The element of this group corresponding to the rotation by an angle φ_i is the matrix

$$g_i = \begin{pmatrix} \cos \varphi_i & -\sin \varphi_i \\ \sin \varphi_i & \cos \varphi_i \end{pmatrix}.$$

Let φ_i be independent random variables, uniformly distributed on the interval $[0, 2\pi]$, and let $B_2 B_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. It follows from the foregoing formulas that the product

$$g_n \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} g_{n-1} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \cdots g_1$$

almost surely converges to 0 if

$$I = \frac{1}{2\pi} \int_0^{2\pi} \ln(\lambda_1^2 \cos^2 \varphi + \lambda_2^2 \sin^2 \varphi) \, d\varphi < 0$$

and to infinity if $I > 0$.

In this example we have verified condition (6.55) by letting $p \to +0$ in (6.54). In the next section we shall consider another approach where a sufficient stability condition of type (6.55) will be derived by a different method and in greater generality. Instead of the Haar measure $\mu(dg)$ we shall have to consider a different measure, related in a natural manner to the distribution of the matrices A_i.

---

[6]The above example is due to V.N. Tutubalin, who derived condition (6.55) by a different method.

6.7 **Asymptotic Stability of Linear Systems with Constant Coefficients**

1. Let us again consider the Markov chain (6.44). Set $\lambda_n = x_n/|x_n|$. The equality $x_n = A_n x_{n-1}$ is equivalent to

$$\lambda_n = \frac{A_n \lambda_{n-1}}{|A_n \lambda_{n-1}|} . \tag{6.56}$$

It follows from (6.56) that the sequence $\lambda_0, \lambda_1, \ldots$ is also a time-homogeneous Markov chain on the sphere $S_l = \{|x| = 1\}$ in $\mathbb{R}^l$. It is readily seen that the transition probability $P_0(\lambda, A)$ of this chain has the Feller property (i.e., the function $\int P_0(\lambda, dy) f(y)$ is continuous if $f(\lambda)$ is continuous). Proceeding as in the proof of Theorem 3.1, using the compactness of the phase space and the Feller property of the transition probability function, we easily prove that the chain $\lambda_0, \lambda_1, \ldots$ has a stationary probability distribution. Suppose that the chain is ergodic and let $\nu(d\lambda)$ be its unique stationary distribution. Set $\rho_n = \ln |x_n|$. Obviously

$$\rho_n = \rho_{n-1} + \ln |A_n \lambda_{n-1}| = \rho_0 + \sum_{i=1}^{n} \ln |A_i \lambda_{i-1}|. \tag{6.57}$$

Since the matrices $A_1, A_2, \ldots$ are independent, the pairs $X_n = \{A_n, \lambda_{n-1}\}$ $(n = 1, 2, \ldots)$ also form a Markov chain. This chain has the stationary distribution $P(dA)\nu(d\lambda)$ in the phase space $\mathcal{A} \times S_l$, where $\mathcal{A}$ is the set of real $l \times l$ matrices. Under fairly broad assumptions about the function $f(A, \lambda)$, $A \in \mathcal{A}$, $\lambda \in S_l$, we have the following form of the strong law of large numbers (see [56, Sect. 5.6]): If

$$\iint |f(A, \lambda)| \mathbf{P}(dA)\nu(d\lambda) < \infty,$$

then the limit

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f(A_k, \lambda_{k-1})$$

exists almost surely for any x_0 and is equal to

$$\iint f(A, \lambda) \mathbf{P}(dA)\nu(d\lambda).$$

Suppose that

$$\iint |\ln |A\lambda|| \mathbf{P}(dA)\nu(d\lambda) < \infty. \tag{6.58}$$

Then, applying the strong law of large numbers to (6.57), we see that almost surely

$$\lim_{n \to \infty} \frac{\rho_n}{n} = a, \quad \text{where } a = \iint \ln |A\lambda| \mathbf{P}(dA)\nu(d\lambda). \tag{6.59}$$

Constant α is called a Lyapunov exponent for the random dynamical system $x_0, x_1, \ldots$. Due to equality $a = \lim_{n \to \infty} \frac{\ln |x_n|}{n}$ this constant characterizes exponential rate of convergence (for $a < 0$) or divergence (for $a > 0$) of x_n to 0 as $n \to \infty$. Since $\rho_n = \ln |A_n A_{n-1} \ldots A_1 x_0|$, it follows from (6.59) that, almost surely,

$$\|A_n A_{n-1} \ldots A_1\| \to 0 \quad \text{as } n \to \infty, \text{ if } a < 0,$$

$$\|A_n A_{n-1} \ldots A_1\| \to \infty \quad \text{as } n \to \infty, \text{ if } a > 0.$$

$a = 0$ is the critical case. As can be seen from the Example 6.3, the chain (6.44) could then possibly be stable but not asymptotically so. The typical case here is nevertheless instability.

Let us study this case in more detail. It is well known that under certain additional conditions the function $f(A_n, \lambda_{n-1})$ satisfies the central limit theorem (see [56, Sect. 5.7]). Assuming that these conditions are satisfied and using (6.57), we get

$$\mathbf{P}\left\{\frac{\rho_n - \mathbf{E}\rho_n}{\sqrt{n}} < \lambda\right\} \to \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-y^2/2\sigma^2} dy. \tag{6.60}$$

(We are assuming here that $\sigma^2 = \lim_{n \to \infty} \frac{1}{n} \operatorname{var} \rho_n \neq 0$.) Since $a = 0$, we may apply Lemma 7.2 in [56, Chap. 5] to conclude that for some constant $k < \infty$

$$|\mathbf{E}\rho_n| = \left|\sum_{i=1}^{n} \mathbf{E} \ln |A_i \lambda_{i-1}|\right| \leq k.$$

Hence it follows by (6.60) that the probability of the event $\rho_n > \lambda\sqrt{n}$ does not tend to zero as $n \to \infty$.

In particular, one sees from (6.60) that the approach presented here may also be used to obtain further results on products of random matrices (see [116]).

Remark 6.8 The Lyapunov exponent a can be effectively calculated from (6.59) only when an invariant measure of the chain $\lambda_0, \lambda_1, \ldots$ is known. An integral equation for this measure is easily set up, but its general solution is fraught with difficulties. In some special cases one can easily compute the invariant measure of the chain $\tilde{\lambda}_n$ obtained by "projecting" the chain (6.44) onto the ellipsoid $|Bx| = 1$ instead of the sphere $|x| = 1$. Next one replaces (6.56) and (6.57) by the formulas

$$\tilde{\lambda}_n = \frac{x_n}{|Bx_n|}; \qquad \tilde{\rho}_n = \ln |Bx_n| = \tilde{\rho}_{n-1} + \ln |BA_n\tilde{\lambda}_{n-1}|.$$

This method is also readily applicable to the example considered at the end of Sect. 6.6.[7]

---

[7]The construction described in this subsection resembles that used by Kesten and Furstenberg in [116]. Sufficient conditions for ergodicity of the Markov chain λ_n on the sphere $|\lambda| = 1$ with identified antipodal points follow from a theorem of Furstenberg [82]. See also the survey article [247].

2. It is natural to ask questions about the behavior of the norms of products of random matrices when each matrix is close to the unit matrix. Limit theorems for this case are discussed in the interesting book of Grenander [96]. Grenander shows that under certain natural assumptions, the limiting distribution of a product of random matrices close to the unit matrix coincides with the distribution of a homogeneous multiplicative matrix-valued stochastic process. This process is governed by a system of linear stochastic equations with constant coefficients and a matrix initial condition. Thus we are again faced with the need to study stability conditions for linear systems of SDEs

$$dX(t) = BX(t)\,dt + \sum_{r=1}^{k} \sigma_r X(t)\,d\xi_r(t) \tag{6.61}$$

with constant coefficients b_i^j and σ_{ir}^j.

To derive necessary and sufficient conditions for asymptotical stability of the system (6.61) we employ the same method as in the case of matrix products.

As before, we set

$$a_{ij}(x) = \sum_{r=1}^{k} \sum_{n,m=1}^{l} \sigma_{ir}^n \sigma_{jr}^m x_n x_m; \qquad A(x) = ((a_{ij}(x)));$$

$$L = \frac{\partial}{\partial t} + \left(Bx, \frac{\partial}{\partial x} \right) + \frac{1}{2}\left(A(x)\frac{\partial}{\partial x}, \frac{\partial}{\partial x} \right)$$

$$= \frac{\partial}{\partial t} + \left(Bx, \frac{\partial}{\partial x} \right) + \frac{1}{2}\sum_{r=1}^{k}\left(\sigma_r x, \frac{\partial}{\partial x} \right)^2.$$

It follows from the general properties of the operator L that the matrix $A(x)$ is positive semidefinite. To simplify matters, we first assume that this matrix is also non-degenerate in the sense that there is a constant $m > 0$ such that for any vector $\alpha = (\alpha_1, \ldots, \alpha_l)$

$$(A(x)\alpha, \alpha) = \sum_{r=1}^{k}(\sigma_r x, \alpha)^2 \geq m|x|^2|\alpha|^2. \tag{6.62}$$

As in (6.56), we introduce new variables:

$$\lambda = \frac{x}{|x|}; \qquad \rho = \ln|x|.$$

The process $\Lambda(t) = X(t)/|X(t)|$ on the sphere $S_l = \{|x| = 1\}$ is Markovian and time-homogeneous.[8] To verify this, we need only use Itô's formula (3.30). We find

---

[8]It is easy to see that the process obtained from $\Lambda(t)$ by identifying antipodal points $\Lambda(t)$ and $-\Lambda(t)$ is also Markovian.

it from expressions for $d\Lambda_i(t)$, from which it is clear that the coefficients of dt and $d\xi_r(t)$ depend only on $\Lambda_1(t), \ldots, \Lambda_l(t)$. Condition (6.62) is sufficient for the process $\Lambda(t)$ to be ergodic, since it guarantees that the transition probability has an everywhere positive density (see Chap. 4). Let $\nu(d\lambda)$ denote the unique normalized invariant measure of the process on the sphere.

Next, let $\rho(t) = \ln|X(t)|$. Using Itô's formula, we get

$$
d\rho(t) = L\rho(t)\,dt + \sum_{r=1}^{k}(\sigma_r\Lambda(t),\,\Lambda(t))\,d\xi_r(t)
$$

$$
= \left[(B\Lambda(t),\,\Lambda(t)) + \frac{1}{2}\operatorname{tr}A(\Lambda(t)) - (A(\Lambda(t))\Lambda(t),\,\Lambda(t))\right]dt
$$

$$
+ \sum_{r=1}^{k}(\sigma_r\Lambda(t),\,\Lambda(t))\,d\xi_r(t). \tag{6.63}
$$

As expected in view of the analogy with formula (6.57), the increment of the function $\rho(t)$ is a functional of the process $\Lambda(t)$ and the Wiener processes $\xi_r(t)$. We set

$$
Q(\lambda) = (B\lambda,\,\lambda) + \frac{1}{2}\operatorname{tr}A(\lambda) - (A(\lambda)\lambda,\,\lambda), \qquad a^* = \int_{S_l} Q(\lambda)\nu(d\lambda).
$$

Theorem 6.11 *Suppose that condition* (6.62) *is satisfied and* $a^* < 0$. *Then the solution* $X(t) \equiv 0$ *of the system* (6.61) *is almost surely asymptotically stable. On the other hand, if* $a^* > 0$, *then for* $x \neq 0$

$$
\mathbf{P}\left\{|X^x(t)| \xrightarrow[t\to\infty]{} \infty\right\} = 1. \tag{6.64}
$$

To prove this we need

Lemma 6.8 *Let* $\sigma(t,\omega) \in \widetilde{\mathcal{N}}_t$ *be a function such that* $\mathbf{E}\sigma^2(t,\omega) < k^2$ ($k > 0$ *is a constant*). *Then the stochastic integral* $\int_0^t \sigma(s,\omega)\,d\xi(s,\omega)$ *almost surely satisfies the relation*

$$
\frac{1}{T}\int_0^T \sigma(t,\omega)\,d\xi(t,\omega) \to 0 \quad as\ T \to \infty.
$$

Proof Let $A_{n,m}$ denote the event

$$
\left\{\sup_{T>2^n}\frac{1}{T}\left|\int_0^T \sigma(t,\omega)\,d\xi(t)\right| > \frac{1}{m}\right\}, \qquad A^{(m)} = \bigcap_{n=1}^{\infty} A_{n,m}.
$$

By virtue of the relation between events

$$B = \left\{ \varlimsup_{T \to \infty} \frac{1}{T} \left| \int_0^T \sigma(t, \omega) \, d\xi(t) \right| > 0 \right\} = \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} A_{n,m};$$

$$A_{n,m} \supset A_{n+1,m}; \qquad A^{(m+1)} \supset A^{(m)}$$

we get

$$\mathbf{P}(B) = \lim_{m \to \infty} \lim_{n \to \infty} \mathbf{P}(A_{n,m}).$$

The lemma will be proved if we show that $\lim_{n \to \infty} \mathbf{P}(A_{n,m}) = 0$ for any $m > 0$. Since the process $\int_0^t \sigma(s, \omega) \, d\xi(s)$ is a martingale, it follows from Theorem 5.2 that

$$\mathbf{P} \left\{ \sup_{2^r \leq T \leq 2^{r+1}} \left| \int_0^T \sigma(t, \omega) \, d\xi(t) \right| > \varepsilon \right\} \leq \frac{1}{\varepsilon} \mathbf{E} \left| \int_0^{2^{r+1}} \sigma(t, \omega) \, d\xi(t) \right|$$

$$\leq \frac{1}{\varepsilon} \left[\mathbf{E} \left(\int_0^{2^{r+1}} \sigma(t, \omega) \, d\xi(t) \right)^2 \right]^{1/2}$$

$$\leq \frac{k 2^{(r+1)/2}}{\varepsilon}.$$

Setting $\varepsilon = \varepsilon_0 2^r$, we get the estimate

$$\mathbf{P} \left\{ \sup_{2^r \leq T \leq 2^{r+1}} \frac{1}{T} \left| \int_0^T \sigma(t, \omega) \, d\xi(t) \right| > \varepsilon_0 \right\} < \frac{k\sqrt{2}}{\varepsilon_0} 2^{-r/2}.$$

This estimate implies the inequality

$$\mathbf{P}(A_{n,m}) \leq \sum_{r=n}^{\infty} \mathbf{P} \left\{ \sup_{2^r \leq T \leq 2^{r+1}} \frac{1}{T} \left| \int_0^T \sigma(t, \omega) \, d\xi(t) \right| > \frac{1}{m} \right\} \leq km\sqrt{2} \sum_{r=n}^{\infty} 2^{-r/2}.$$

Thus $P(A_{n,m}) \to 0$ as $n \to \infty$. Lemma 6.8 is proved. $\qquad \square$

Proof of Theorem 6.11 We can rewrite (6.63) as

$$\frac{\rho(T) - \rho(0)}{T} = \frac{1}{T} \int_0^T Q(\Lambda(t)) \, dt + \frac{1}{T} \sum_{r=1}^{k} \int_0^T (\sigma_r \Lambda(t), \Lambda(t)) \, d\xi_r(t). \quad (6.65)$$

It follows from the strong law of large numbers for the process $\Lambda(t)$, and from Lemma 6.8, that almost surely

$$\lim_{T \to \infty} \frac{\rho(T)}{T} = a^*. \qquad (6.66)$$

This implies both parts of the theorem. $\qquad \square$

Remark 6.9 It follows from (6.66) that almost surely

$$a^* = \lim_{T \to \infty} \frac{\ln |X^x(T)|}{T}.$$

So a^* is the Lyapunov exponent for the process described by the stochastic (6.61).

Theorem 6.12 *Suppose that condition* (6.62) *is satisfied and* $a^* = 0$. *Then the solution* $X(t) \equiv 0$ *of the system* (6.61) *is neither asymptotically stable nor asymptotically unstable in the sense of* (6.64).

Proof Suppose that the solution $X(t) \equiv 0$ is asymptotically stable in probability.

Then, by Theorem 6.5, this solution is exponentially p-stable for all sufficiently small $p > 0$. Hence, by (6.63) and using Jensen's inequality

$$\mathbf{E}(\exp \xi) \geq \exp\{\mathbf{E}\xi\}$$

(see [56]), we get

$$|x_0|^p \exp\left\{ p \int_0^T \mathbf{E}Q(\Lambda(s)) \, ds \right\} \leq \mathbf{E}|X^x(T)|^p < Ae^{-\alpha T}.$$

These inequalities imply that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \mathbf{E}Q(\Lambda(s)) \, ds < 0,$$

but this contradicts the assumption $a^* = 0$. One proves similarly that the solution $X(t) \equiv 0$ cannot satisfy condition (6.64). $\qquad\square$

6.8 Systems with Constant Coefficients (Continued)

In the preceding section we studied conditions implying asymptotical stability of the system (6.61). We assumed the nondegeneracy condition (6.62). We now consider what modifications must be made in the arguments of Sect. 6.7 if the diffusion matrix is allowed to degenerate on certain curves, surfaces, or even everywhere.

We first observe that the essential point for all arguments in Sect. 6.7 is not so much condition (6.62) as one of its consequences, namely the ergodicity of the Markov process $\Lambda(t)$ on the sphere $|x| = 1$.[9] Now suppose that ergodicity fails

---

[9]It is readily seen that the following weaker than (6.62) condition is sufficient: For all vectors λ and μ such that $|\lambda| = |\mu| = 1$, $\lambda \neq \pm\mu$,

$$\sum_{r=1}^{k}(\sigma_r\lambda - (\sigma_r\lambda, \lambda)\lambda, \mu)^2 > 0. \qquad (6.67)$$

to hold and that the path of the Markov process

$$\Lambda^{\lambda_0}(t) = \frac{X^x(t)}{|X^x(t)|} \quad (\lambda_0 = x/|x|),$$

satisfying the initial condition $\Lambda^{\lambda_0}(t) = \lambda_0$, may belong to different ergodic components A of the process for different λ_0.

More general, so called Lie algebra condition was proposed in [8], see (B.23) in Appendix B. Below we will consider properties of Lyapunov exponent and the moment Lyapunov exponent under condition C (see (B.22) in Appendix B). Let $\mu_A(d\lambda)$ denote the stationary initial distribution corresponding to the component A.

Applying the strong law of large numbers for the component A and Lemma 6.8, we see as in the proof of Theorem 6.11 that

$$\lim_{t\to\infty} \frac{\ln|X^x(t)|}{t} = a_A = \int Q(\lambda)\mu_A(d\lambda)$$

for μ_A-almost all values of $\lambda = x/|x|$.

Following the proofs of Theorems 6.11 and 6.12, we see that a necessary condition for the system (6.61) to be asymptotically stable is that

$$a_A < 0 \tag{6.68}$$

holds for all ergodic components A. A sufficient condition is that for all $x \in \mathbb{R}^l$,

$$\lim_{t\to\infty} \frac{\ln|X^x(t)|}{t} < 0$$

holds almost surely. The process $\Lambda(t)$ may have infinitely many ergodic components. For instance, the ergodic components of the process $\Lambda(t)$ for the deterministic system $dx/dt = x$, $dy/dt = y$ are all the points of the circle $x^2 + y^2 = 1$.

Therefore it might seem at first sight that (6.68) may involve an infinite set of conditions. However, as we shall show below, the number a_A can take on at most l distinct values. We shall first prove a simple lemma which is usually employed to investigate the properties of Lyapunov's characteristic numbers (see [191, Sect. 77]).

Lemma 6.9 *Let $X_i(t)$ ($i = 1, 2$) be $\mathbb{R}^l$-valued functions such that*

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \ln|X_i(t)| = a_i < \infty.$$

Then

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \ln|x_1 X_1(t) + x_2 X_2(t)| \le \max(a_1, a_2).$$

Proof To be specific, suppose that $a_1 \le a_2$. It follows from the assumptions of the lemma that

$$|X_i(t)| < e^{(a_l+\varepsilon)t}$$

for any $\varepsilon > 0$ and all $t > T(\varepsilon)$. Therefore

$$|x_1 X_1(t) + x_2 X_2(t)| \leq |x_1| e^{(a_1+\varepsilon)t} + |x_2| e^{(a_2+\varepsilon)t} \leq k e^{(a_2+\varepsilon)t}.$$

Hence we have the inequality

$$\varlimsup_{t\to\infty} \frac{1}{t} \ln |x_1 X_1(t) + x_2 X_2(t)| \leq a_2.$$

This proves the lemma.   □

If the process $\Lambda(t)$ is not ergodic, the following lemma may be useful in investigating the stability of the system (6.61).

Lemma 6.10 *Suppose that there are l linearly independent vectors $\lambda_1, \ldots, \lambda_l$ in S_l such that almost surely*

$$\varlimsup_{t\to\infty} \frac{1}{T} \int_0^T Q(\Lambda^{\lambda_i}(t))\, dt < 0. \qquad (6.69)$$

Then the system (6.61) is asymptotically stable.

Proof It follows from (6.65), Lemma 6.8 and (6.69) that

$$\varlimsup_{t\to\infty} \frac{1}{t} \ln |X^{\lambda_i}(t)| < 0$$

almost surely. Hence, by Lemma 6.9, we get

$$\varlimsup_{t\to\infty} \frac{1}{t} \ln |X^x(t)| = \varlimsup_{t\to\infty} \frac{1}{t} \ln \left| \sum_{i=1}^{l} k_i X^{\lambda_i}(t) \right| < 0,$$

where $x = \sum_{i=1}^{l} k_i \lambda_i$. This proves the lemma.   □

Corollary 6.3 *The process $\Lambda(t)$ has at most l ergodic components A_i, corresponding to the different values of*

$$a_i^* = \int Q(\lambda)\mu_i(d\lambda).$$

(Here μ_i is the stationary distribution for the component A_i.) Moreover, if $a_1^* < a_2^* < \cdots < a_l^*$, then $a_l^* < 0$ is a sufficient condition for the system (6.61) to be asymptotically stable.

Proof Let $A_1, \ldots, A_k$ be the ergodic components of the process $\Lambda(t)$ and assume that the corresponding a_i^* are monotone increasing: $a_1^* < \cdots < a_k^*$. Let λ_i be vectors

such that almost surely

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T Q(\Lambda^{\lambda_i}(t)) \, dt = a_i^*.$$

Then the vectors $\lambda_1, \ldots, \lambda_k$ are linearly independent. In fact, otherwise we have $\lambda_i = c_1 \lambda_1 + \cdots + c_{i-1} \lambda_{i-1}$ for some $i \leq k$. Therefore

$$X^{\lambda_i}(t) = \sum_{j=1}^{i-1} c_j X^{\lambda_j}(t).$$

Hence, by Lemma 6.9,

$$a_i^* = \lim_{t \to \infty} \frac{1}{t} \ln |X^{\lambda_i}(t)| \leq \max(a_1^*, \ldots, a_{i-1}^*) = a_{i-1}^*.$$

This is a contradiction, and thus $\lambda_1, \ldots, \lambda_k$ are linearly independent. This, together with Lemma 6.10, implies the assertion. □

Another obvious consequence of Lemma 6.10 is

Corollary 6.4 *If the process $\Lambda(t)$ has an ergodic component whose stationary distribution $\mu_A(d\lambda)$ is not concentrated on any hyperplane $\sum_{i=1}^l k_i \lambda_i + k_0 = 0$, then a sufficient condition for asymptotic stability of the system (6.61) is that $a^* = \int Q(\lambda)\mu_A(d\lambda) < 0$.*

Let us consider the case $l = 2$ in more detail. Let $\lambda(\varphi)$ denote the vector in the plane with components $\lambda_1 = \cos\varphi$; $\lambda_2 = \sin\varphi$, $\hat{\lambda}(\varphi) = -d\lambda(\varphi)/d\varphi$. As we have already stated, the stochastic process $\varphi(t)$ on the circle generated by the system (6.61) is Markovian. One shows easily that

$$d\varphi(t) = \Phi(\varphi(t)) \, dt + \Psi(\varphi(t)) \, d\xi(t), \tag{6.70}$$

where

$$\Psi^2(\varphi) = (A(\lambda(\varphi))\hat{\lambda}(\varphi), \hat{\lambda}(\varphi));$$

$$\Phi(\varphi) = -(B\lambda(\varphi), \hat{\lambda}(\varphi)) + (A(\lambda(\varphi))\lambda(\varphi), \hat{\lambda}(\varphi)),$$

and $\tilde{\xi}(t)$ is a Wiener process with zero mean such that $\mathbf{E}\xi^2(t) = t$.

We first assume that

$$\Psi^2(\varphi) > 0 \quad (0 \leq \varphi < 2\pi). \tag{6.71}$$

(This is implied, for instance, by inequality (6.62).) Then the process $\varphi(t)$ has a unique stationary distribution which is absolutely continuous and has density $\mu(\varphi)$

relative to the uniform distribution on the circle. This density satisfies the Fokker–Planck–Kolmogorov equation, which in our case is

$$\frac{1}{2}\frac{d^2}{d\varphi^2}\left(\Psi^2(\varphi)\mu\right) - \frac{d}{d\varphi}\left(\Phi(\varphi)\mu\right) = 0. \tag{6.72}$$

Equation (6.72) has a unique solution satisfying the normalization condition

$$\int_0^{2\pi} \mu(\varphi)\,d\varphi = 1 \tag{6.73}$$

and the periodicity condition

$$\mu(0) = \mu(2\pi). \tag{6.74}$$

It is easy to see that this solution is given by

$$\mu(\varphi) = k\left[1 + \frac{W(2\pi) - 1}{\int_0^{2\pi} W(s)\,ds}\int_0^{\varphi} W(u)\,du\right][W(\varphi)\Psi^2(\varphi)]^{-1}, \tag{6.75}$$

where

$$W(\varphi) = \exp\left\{-2\int_0^{\varphi} \frac{\Phi(v)dv}{\Psi^2(v)}\right\},$$

and the constant k is determined by the normalization (6.73).

Applying Theorems 6.11 and 6.12 to the case $l = 2$, we obtain that

$$\int_0^{2\pi} Q(\lambda(\varphi))\mu(\varphi)\,d\varphi < 0 \tag{6.76}$$

is a necessary and sufficient condition for asymptotic stability, expressed in terms of quadratures.

We shall now allow that the function $\Psi^2(\varphi)$ may vanish. Since

$$\Psi^2(\varphi) = \sum_{r=1}^{k}(\sigma_r\lambda(\varphi), \hat{\lambda}(\varphi))^2,$$

it follows that apart from the trivial case of a deterministic system ($\sigma_r \equiv 0$), there are two possible cases: either

$$\Psi^2(\varphi) = 0 \tag{6.77}$$

only for $\sin\varphi = 0$ and for $\cos\varphi = 0$, or (6.76) is equivalent to a fourth-degree equation in $\tan\varphi$ or $\cot\varphi$. In both cases (in view of the inequality $\Psi^2(\varphi) \geq 0$), (6.77) is satisfied by at most two values of φ in the interval $0 \leq \varphi < \pi$, each of which pairs off with another value differing from it by π. It follows that the process $\varphi(t)$ has at most four ergodic components if $\sigma_{ir}^k \not\equiv 0$.

Let us discuss the possibilities in greater detail.

1. Let $\varphi_1, \varphi_2, \varphi_1 + \pi, \varphi_2 + \pi$ ($0 \leq \varphi_1 < \varphi_2 < \pi$) be the solutions of (6.77). Then one readily sees that the process $\varphi(t)$ is ergodic for $\text{sign}\,\Phi(\varphi_1) = \text{sign}\,\Phi(\varphi_2)$ and has two ergodic components if $\text{sign}\,\Phi(\varphi_1) = -\text{sign}\,\Phi(\varphi_2)$.

2. Suppose that (6.77) has two solutions φ_1 and $\varphi_1 + \pi$ in the interval $[0, 2\pi]$. Then the process $\varphi(t)$ is always ergodic, provided $\Phi(\varphi_1) \neq 0$.

3. If $\Phi(\varphi_k) = 0$, the process $\varphi(t)$ has stationary points at $\varphi = \varphi_k$ and $\varphi = \varphi_k + \pi$.

Example 6.4 Consider the system

$$\left.\begin{array}{l} dX_1(t) = aX_1\,dt + \sigma_1 X_1\,d\xi_1(t), \\ dX_2(t) = bX_2\,dt + \sigma_2 X_2\,d\xi_2(t). \end{array}\right\} \tag{6.78}$$

Then

$$\Psi^2(\varphi) = (\sigma_1^2 + \sigma_2^2)\cos^2\varphi\sin^2\varphi,$$

$$\Phi(\varphi) = \sin\varphi\cos\varphi(\sigma_1^2\cos^2\varphi - \sigma_2^2\sin^2\varphi - a + b).$$

Thus the points $\varphi_k = k\pi/2$ ($k = 0, 1, 2, 3$) are stationary points of the process $\varphi(t)$. All the invariant measures of the process $\varphi(t)$ are concentrated at these points. Hence it follows that the system (6.78) is stable if and only if both its components are stable. This conclusion may also be derived directly, since the components of the process (6.78) are independent.

More substantial examples follow in the next section.

6.9 Two Examples

Example 6.5 A much discussed question in the literature (see [180, 245] and others) is whether a linear deterministic system can be stabilized with artificially disturbing its parameters by white noise. The problem is understood by different authors in different ways, since the stochastic equation obtained by superimposing noise may be made rigorous in various ways. As mentioned in Chap. 5, a natural approach is to study the problem for "physically feasible" noise in the sense of Sect. 5.5. We have already seen that one-dimensional systems cannot be stabilized by physically feasible noise.

We now consider the system

$$\left.\begin{array}{l} dX_1(t) = b_1 X_1\,dt + \sigma(X_1 d^*\xi_1(t) + X_2 d^*\xi_2(t)), \\ dX_2(t) = b_2 X_2\,dt + \sigma(X_2 d^*\xi_1(t) - X_1 d^*\xi_2(t)), \end{array}\right\} \tag{6.79}$$

in $\mathbb{R}^2$, where the $d^*\xi_i(t)$ are Stratonovich differentials. The generator of this process is obviously (see (5.31))

$$L = \frac{\partial}{\partial t} + \left(b_1 + \frac{\sigma^2}{2}\right)x_1\frac{\partial}{\partial x_1} + \left(b_2 + \frac{\sigma^2}{2}\right)x_2\frac{\partial}{\partial x_2} + \frac{\sigma^2}{2}(x_1^2 + x_2^2)\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right).$$

Hence, using the formulas of Sect. 6.8, we easily see that

$$\mu(\varphi) = c\exp\left\{\frac{b_1 - b_2}{\sigma^2}\cos^2\varphi\right\}; \qquad Q(\lambda(\varphi)) = \frac{\sigma^2}{2} + b_1\cos^2\varphi + b_2\sin^2\varphi.$$

Applying Theorems 6.11 and 6.12, we see that the condition

$$a^* = \int_0^{\pi/2}\left(\frac{\sigma^2}{2} + b_1\cos^2\varphi - b_2\sin^2\varphi\right)\exp\left\{\frac{b_1 - b_2}{\sigma^2}\cos^2\varphi\right\}d\varphi < 0 \quad (6.80)$$

is necessary and sufficient for the system (6.79) to be asymptotically stable.

We may easily give another form to condition (6.80) by using the well-known integral representation

$$I_n(z) = \frac{(-1)^n}{\pi}\int_0^\pi e^{-z\cos\theta}\cos n\theta\,d\theta \qquad (6.81)$$

for the Bessel function $I_n(z)$ of a pure imaginary argument. In fact, if we set $\kappa = (b_1 - b_2)/\sigma^2$, formulas (6.80) and (6.81) yield the condition

$$1 + \frac{2b_1}{\sigma^2} < \kappa\left(1 - \frac{I_1(\kappa/2)}{I_0(\kappa/2)}\right) \qquad (6.82)$$

which is equivalent to (6.80).

The asymptotic representation for the function $I_n(z)$ as $z \to \infty$ shows that the last inequality is valid if, say, $b_2 < 0$ is chosen with sufficiently large absolute value, and $b_1/\sigma^2 < 3/8$.

We have thus proved that for suitable choice of $b_1 > 0$ and $b_2 < 0$ the unstable deterministic system

$$\frac{dx_1}{dt} = b_1 x_1; \qquad \frac{dx_2}{dt} = b_2 x_2 \qquad (6.83)$$

becomes asymptotically stable when its parameters are perturbed by certain physically feasible white noise processes. This result is valid for any deterministic system which is reducible to the canonical form (6.83) by a linear transformation. Indeed, linear transformations do not affect the stability properties of a system.

Remark 6.10 After publication this example in the first edition in 1969 stabilization by noise became very popular theme in Mathematical and Mechanical research, see, e.g., [3] and references therein.

Example 6.6 Consider a linear system with eigenfrequency ω, subject to the action of a damping force proportional to velocity with coefficient k. This system is described by the equation

$$\ddot{x} + k\dot{x} + \omega^2 x = 0. \tag{6.84}$$

It is evident that this system is stable for $k > 0$. However, in many problems it seems natural to assume that k is merely the mean value of the damping coefficient, while its true value is a stochastic process with small correlation interval. It is extremely interesting to study "bifurcation" values of the noise intensity, i.e., values for which the system first becomes unstable.

A limiting case of this problem is to determine a constant σ_0 such that for $\sigma < \sigma_0$ the solution $x = x' \equiv 0$ of the stochastic equation

$$\ddot{x} + (k + \sigma\dot{\xi})\dot{x} + \omega^2 x = 0 \tag{6.85}$$

is asymptotically stable, while for $\sigma > \sigma_0$ it is unstable. The process $\dot{\xi}(t)$ in (6.85) is white noise of unit intensity. We shall interpret this equation in the sense described in Sect. 5.5, as a Stratonovich stochastic equation.

Setting $x_1 = \omega x$, $x_2 = \dot{x}$, we get the system

$$\left.\begin{aligned} dx_1 &= \omega x_2 \, dt, \\ dx_2 &= -(kx_2 + \omega x_1) \, dt - \sigma x_2 \, d^*\xi(t) \end{aligned}\right\} \tag{6.86}$$

with generator

$$L = \omega x_2 \frac{\partial}{\partial x_1} + \left[\left(\frac{\sigma^2}{2} - k\right)x_2 - \omega x_1\right]\frac{\partial}{\partial x_2} + \frac{1}{2}\sigma^2 x_2^2 \frac{\partial^2}{\partial x_2^2}.$$

We shall first find a sufficient condition for stability in mean square, using the algorithm described Sect. 6.3. Our aim is to find conditions under which a quadratic form

$$W = \frac{A}{2} x_1^2 + Bx_1 x_2 + \frac{C}{2} x_2^2,$$

which satisfies equation

$$LW = -x_1^2 - x_2^2$$

is positive definite.

Simple computations lead to the equalities

$$C = \frac{2}{k - \sigma^2}; \qquad B = \frac{1}{\omega}; \qquad A = \frac{2}{k - \sigma^2} + \frac{1}{\omega^2}\left(k - \frac{\sigma^2}{2}\right).$$

It is clear that the form W is positive definite if and only if $\sigma^2 < k$. Hence we get a lower bound for the bifurcation value of the noise: $\sigma_0^2 > k$.

To derive an equation for σ_0^2, we use the results of Sect. 6.8. According to these, we get that $\sigma = \sigma_0$ satisfies, for a process $\Lambda(t)$ which is ergodic on the circle $|x| = 1$, the equation

$$\int_0^{2\pi} Q(\lambda(\varphi))\mu(\varphi)\,d\varphi = 0. \tag{6.87}$$

We easily see from (6.86) and the formulas of Sect. 6.8 that

$$\Psi^2(\varphi) = \sigma^2 \sin^2 \varphi \cos^2 \varphi,$$

$$\Phi(\varphi) = \left(\frac{\sigma^2}{2} - k\right)\sin\varphi\cos\varphi + \sigma^2\sin^3\varphi\cos\varphi - \omega,$$

$$\begin{aligned} Q(\lambda(\varphi)) &= \left(\frac{\sigma^2}{2} - k\right)\sin^2\varphi - \frac{1}{2}\sigma^2\sin^2\varphi - \sigma^2\sin^4\varphi \\ &= \sin^2\varphi(\sigma^2\cos^2\varphi - k). \end{aligned} \tag{6.88}$$

The function $\Psi^2(\varphi)$ vanishes for $\varphi_k = k\pi/2$ ($k = 0, 1, 2, 3$). However, Φ does not vanish at these points ($\Phi(\varphi_k) = -\omega$). Hence the process $\Lambda(t)$ described by (6.70) is ergodic. Let us determine the density $\mu(\varphi)$ of its invariant measure. In this case we cannot use formula (6.75), since (6.72) has singularities at the points $\varphi = \varphi_k$. Nevertheless, the function $\mu(\varphi)$ satisfies (6.72) for $\varphi \neq \varphi_k$. Moreover, it can be shown that μ is bounded and continuous in the neighborhood of the points $\varphi = \varphi_k$ and that it also satisfies condition (6.73). It is readily verified that the unique solution of (6.72) satisfying these additional conditions is the function

$$\mu(\varphi) = \begin{cases} c\int_0^\varphi W(u)\,du[W(\varphi)\Psi^2(\varphi)]^{-1} & \text{for } 0 \le \varphi < \frac{\pi}{2}, \\ c\int_{-\pi/2}^\varphi W(u)\,du[W(\varphi)\Psi^2(\varphi)]^{-1} & \text{for } -\frac{\pi}{2} \le \varphi < 0, \\ \mu(\varphi - \pi) & \text{for } \frac{\pi}{2} \le \varphi < \frac{3\pi}{2}, \end{cases} \tag{6.89}$$

where

$$W(\varphi) = (\cos\varphi)^{-2}(\tan\varphi)^{(2k-\sigma^2)/\sigma^2}\exp\left\{-\frac{4\omega}{\sigma^2}\cot 2\varphi\right\},$$

and the constant c is determined by condition (6.73).

It follows from (6.87), (6.88) and (6.89) that the constant σ_0^2 satisfies the equation

$$\begin{aligned} F(\sigma^2, k, \omega) &\equiv \int_{-\pi/2}^0 \frac{\sin^2\varphi(\sigma^2\cos^2\varphi - k)}{W(\varphi)\Psi^2(\varphi)}\int_{-\pi/2}^\varphi W(v)\,dv\,d\varphi \\ &\quad + \int_0^{\pi/2} \frac{\sin^2\varphi(\sigma^2\cos^2\varphi - k)}{W(\varphi)\Psi^2(\varphi)}\int_0^\varphi W(v)\,dv\,d\varphi = 0. \quad (6.90) \end{aligned}$$

From this we easily derive our previous estimate $\sigma_0^2 > k$. The parameters k, σ^2 and ω have the same dimensionality. We can therefore replace them by two dimensionless quantities, say

$$\widehat{\sigma}^2 = \frac{\sigma^2}{k}; \qquad \widehat{\omega} = \frac{\omega}{k}.$$

When written in these new variables, (6.90) becomes

$$F(\widehat{\sigma}^2, 1, \widehat{\omega}) \equiv \widehat{F}(\widehat{\sigma}^2, \widehat{\omega}) = 0. \tag{6.91}$$

This equation is fairly involved. Nevertheless, it enables us to investigate the dependence of the dimensionless critical noise power $\widehat{\sigma}_0^2$ on the dimensionless frequency $\widehat{\omega}$ for the limiting cases $\widehat{\omega} \to 0$ and $\widehat{\omega} \to \infty$.

It is easy to see that $\widehat{\sigma}_0^2 \to \infty$ as $\widehat{\omega} \to 0$; in other words, low-frequency oscillations of the system are stable under very strong perturbations of the damping coefficient.

Let us now investigate in greater detail the other limiting case $\widehat{\omega} \to \infty$. The quantity σ_0^2 may now be determined by Laplace approximation of the integrals. To this end, we transform the variables in (6.91) by the formulas

$$-\cot 2\varphi = z, \qquad -\cot 2v = u.$$

We introduce the notations

$$R_1(z) = \frac{\sqrt{z^2+1}+z}{2\sqrt{z^2+1}}, \qquad R_2(z) = \frac{\sqrt{z^2+1}-z}{2\sqrt{z^2+1}},$$

$$R_3(z) = \frac{R_2(z)}{R_1(z)}, \qquad a = \frac{2-\widehat{\sigma}^2}{2\widehat{\sigma}^2}.$$

Then

$$\widehat{F}(\widehat{\sigma}^2, \widehat{\omega}) = \frac{1}{4\widehat{\sigma}^2}\left[\int_{-\infty}^{\infty} \frac{dz(\widehat{\sigma}^2 R_1(z) - 1)}{(z^2+1)R_3^\alpha(z)} \int_{-\infty}^{z} \frac{du\, R_3^\alpha(u)e^{4\widehat{\omega}\widehat{\sigma}^{-1}(u-z)}}{(u^2+1)R_1(u)}\right.$$

$$\left. + \int_{-\infty}^{\infty} \frac{dz(\widehat{\sigma}^2 R_2(z) - 1)R_3^\alpha(z)}{(z^2+1)} \int_{-\infty}^{z} \frac{du\, e^{4\widehat{\omega}\widehat{\sigma}^{-2}(u-z)}}{(u^2+1)R_3^\alpha(u)R_2(u)}\right].$$

Applying the asymptotic formula

$$\int_{-\infty}^{z} \varphi(u)e^{\lambda(u-z)}\, dz = \frac{1}{\lambda}\varphi(z) + O\left(\frac{1}{\lambda^2}\right) \qquad (\lambda \to \infty)$$

[179, pp. 446–450], we see that

Fig. 6.1
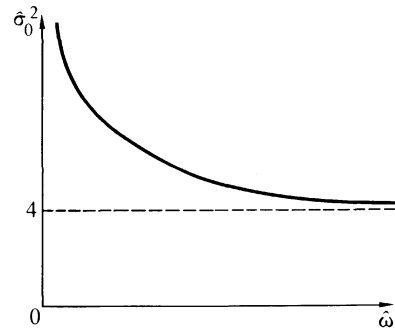
$$\widehat{F}(\widehat{\sigma}^2, \widehat{\omega}) = \frac{1}{16\widehat{\omega}} \int_{-\infty}^{\infty} \frac{dz[\widehat{\sigma}^2(\sqrt{z^2+1}+z) - 2\sqrt{z^2+1}]}{(z^2+1)^2(\sqrt{z^2+1}+z)}$$

$$+ \frac{1}{16\omega} \int_{-\infty}^{\infty} \frac{dz[\widehat{\sigma}^2(\sqrt{z^2+1}-z) - 2\sqrt{z^2+1}]}{(z^2+1)^2(\sqrt{z^2+1}-z)} + O\left(\frac{1}{\widehat{\omega}^2}\right)$$

$$= \frac{1}{8\widehat{\omega}} \int_{-\infty}^{\infty} \frac{dz}{(z^2+1)^2}(\widehat{\sigma}^2 - 2(z^2+1)) + O\left(\frac{1}{\widehat{\omega}^2}\right)$$

$$= \frac{\pi}{8\widehat{\omega}} \left[\frac{\widehat{\sigma}^2}{2} - 2\right] + O\left(\frac{1}{\widehat{\omega}^2}\right)$$

as $\widehat{\omega} \to \infty$. Hence we get the equality

$$\lim_{\widehat{\omega}\to\infty} \widehat{\sigma}_0^2(\widehat{\omega}) = 4.$$

Therefore the critical noise power for high frequencies is close to $4k$ (Fig. 6.1).

Analogous arguments apply to the investigation of random parametric excitations (i.e., random perturbations of the eigenfrequency). If the intensity of the noise is low, the investigation is a simple matter; one can apply then the method of averaging. On this subject, see [147], [148], [257], [254], [119] and others.

6.10 n-th Order Equations[10]

It is well known that the solution $y = 0$ of the equation

$$y^{(n)} + b_1 y^{(n-1)} + \cdots + b_n y = 0 \tag{6.92}$$

is stable if and only if the Routh–Hurwitz conditions

---

[10]See [220]. Some special cases were considered previously in [42, 234].

$$
\Delta_1 = b_1 > 0; \qquad \Delta_2 = \begin{vmatrix} b_1 & b_3 \\ 1 & b_2 \end{vmatrix} > 0;
$$

$$
\Delta_3 = \begin{vmatrix} b_1 & b_3 & b_5 \\ 1 & b_2 & b_4 \\ 0 & b_1 & b_3 \end{vmatrix} > 0; \qquad \ldots; \qquad \Delta_n = \begin{vmatrix} b_1 & b_3 & b_5 & \ldots & 0 \\ 1 & b_2 & b_4 & \ldots & 0 \\ 0 & b_1 & b_3 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & b_n \end{vmatrix} > 0 \Bigg\}
$$

$$
(6.93)
$$

are satisfied (see [84]).

We shall derive analogous necessary and sufficient conditions for mean square stability of the system

$$
y^{(n)} + (b_1 + \dot\eta_1(t)) y^{(n-1)} + \cdots + (b_n + \dot\eta_n(t)) y = 0.
$$

Here $\dot\eta_1(t), \ldots, \dot\eta_n(t)$ are Gaussian white noise processes, generally correlated, so that

$$
\mathbf{E}\dot\eta_i(s)\dot\eta_j(t) = a_{ij}\delta(t-s).
$$

Replacing the processes $\dot\eta_1(t), \ldots, \dot\eta_n(t)$ by independent processes as indicated in Sect. 6.1, and setting

$$
X_1(t) = y(t); \qquad \ldots; \qquad X_n(t) = y^{(n-1)}(t),
$$

we get a system of Itô equations

$$
\begin{aligned}
& dX_1(t) = X_2(t)\,dt, \\
& dX_2(t) = X_3(t)\,dt, \qquad \ldots, \qquad dX_{n-1}(t) = X_n(t)\,dt, \\
& \quad\vdots \\
& dX_n(t) = -\sum_{i=1}^n b_i X_{n-i+1}(t)\,dt - \sum_{i,j=1}^n \sigma_{ij} X_{n-i+1}\,d\xi_j(t),
\end{aligned}
\qquad (6.94)
$$

where $((\sigma_{ij}))((\sigma_{ji})) = ((a_{ij}))$.

It is easy to see that the generator of the process $X(t)$ is

$$
L = \frac{\partial}{\partial t} + \sum_{i=1}^{n-1} x_{i+1}\frac{\partial}{\partial x_i} - \sum_{i=1}^n b_i x_{n-i+1}\frac{\partial}{\partial x_n} + \frac{1}{2}\sum_{i,j=1}^n a_{ij} x_{n-i+1} x_{n-j+1}\frac{\partial^2}{\partial x_n^2}.
$$

Using the methods of Sects. 6.2 and 6.3 we may now determine necessary and sufficient conditions for the system (6.94) to be stable in mean square. However, the resulting conditions will involve determinants of order n^2. We shall therefore adopt another approach; this will give us conditions involving the computation of only $n+1$ determinants, the largest of order n. We shall see that the first n determinants are the same as in (6.93), while the last is obtained from Δ_n by replacing

its first row by a vector whose components are computed by a certain rule from the coefficients a_{ij}.

As we saw in Sect. 6.2 (Remark 6.3), a necessary condition for the system (6.94) to be asymptotically stable in mean square is that the "non-random" system

$$\frac{dx_1}{dt} = x_2, \qquad \ldots, \qquad \frac{dx_{n-1}}{dt} = x_n, \qquad \frac{dx_n}{dt} = -\sum_{i=1}^{n} b_i x_{n-i+1} \qquad (6.95)$$

is asymptotically stable, i.e., the Routh–Hurwitz conditions (6.93) hold. It is known that under these assumptions there exists a positive definite quadratic form $V(x)$ such that

$$L_0 V = \sum_{i=1}^{n-1} x_{i+1} \frac{\partial V}{\partial x_i} - \sum_{i=1}^{n} b_i x_{n-i+1} \frac{\partial V}{\partial x_n}, \qquad (6.96)$$

i.e. the total derivative along the trajectory of the system (6.95) is equal to a pre-scribed negative definite form $W(x)$.

We first assume that the quadratic form

$$a(x) = \sum_{i,j=1}^{n} a_{ij} x_{n-i+1} x_{n-j+1}$$

is positive definite. Then we have

Lemma 6.11 *The trivial solution of the system* (6.94) *is asymptotically stable in mean square if and only if there exists a positive definite quadratic form*

$$V(x) = \sum_{i,j=1}^{n} d_{ij} x_i x_j$$

such that

$$L_0 V = -a(x), \qquad d_{nn} < 1. \qquad (6.97)$$

Proof Suppose that there exists a form $V(x) = \sum d_{ij} x_i x_j$ satisfying the conditions of the lemma. Then, by (6.96) and (6.97), we have

$$LV = L_0 V + \frac{a(x)}{2} \frac{\partial^2 V}{\partial x_n^2} = (d_{nn} - 1) a(x) < 0.$$

Hence, by Theorem 6.3, it follows that the system (6.94) is asymptotically stable in mean square.

Conversely, if the system (6.94) is asymptotically stable, it follows from the same theorem that there exists a positive definite quadratic form

$$V_1(x) = \sum_{i,j=1}^{n} v_{ij} x_i x_j,$$

such that $LV_1 = -a(x)$, i.e.,

$$L_0 V_1 = LV_1 - \frac{a(x)}{2} \frac{\partial^2 V_1}{\partial x_n^2} = -(v_{nn} + 1) a(x).$$

Thus $V(x) = V_1(x)/(v_{nn} + 1)$. Consequently $d_{nn} = v_{nn}/(v_{nn} + 1) < 1$, which we wished to show.   □

Thus in order to obtain the desired conditions we must express the coefficient d_{nn} of the form $V(x)$ defined by (6.97) in terms of the parameters b_i, a_{ij} of the system (6.94).

It follows from Theorem 6.3 that any function $V(x)$ satisfying (6.97) can be written as

$$V(x) = \int_0^\infty a(X^x(u)) \, du.$$

This equality makes it possible to express the coefficients of $V(x)$ including d_{nn}, in terms of a fundamental system of solutions of (6.95). Next, as shown in [22], we may express them in terms of the coefficients a_i, b_{ij}. Indeed, according to [22],

$$d_{nn} = \frac{1}{2\Delta_n} \sum_{r=0}^{n-1} q_{nn}^{(r)} \Delta_{1,r+1}, \tag{6.98}$$

where $\Delta_{1,r+1}$ is the cofactor of the element in the first row and $(r+1)$-th column of the last Hurwitz determinant Δ_n, and the numbers $q_{nn}^{(r)}$ are related to the coefficients a_{ij} of the form $a(x)$ by

$$(-1)^{n-1} \sum_{i,j=1}^{n} a_{n-i+1,n-j+1} D_{ni}(\lambda) D_{nj}(-\lambda) = \sum_{r=0}^{n-1} q_{nn}^{(r)} \lambda^{2(n-r-1)}. \tag{6.99}$$

Here $D_{nj}(\lambda)$ is the cofactor of the element in the *n*-th row and *j*-th column of the determinant

$$D(\lambda) = \begin{vmatrix} -\lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & -\lambda & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -\lambda & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \ldots & -a_1 & -\lambda \end{vmatrix}$$

of the system (6.94).

It is easy to see that

$$D_{ni}(\lambda) D_{nj}(-\lambda) = \lambda^{i+j-2} (-1)^{j-1}.$$

Hence, using (6.99), we infer that

$$\sum_{k=0}^{n-1} \lambda^{2k} \sum_{p+q=2(n-k)} a_{pq}(-1)^{q+1} = \sum_{k=0}^{n-1} q_{nn}^{(n-k-1)} \lambda^{2k}$$

and consequently

$$q_{nn}^{(n-k-1)} = \sum_{p+q=2(n-k)} a_{pq}(-1)^{q+1}. \tag{6.100}$$

It follows from Lemma 6.11, by (6.98) and (6.100), that if $a(x)$ is a positive definite quadratic form, then the system (6.94) is stable in mean square if and only if

$$\Delta_1 > 0; \qquad \Delta_2 > 0; \qquad \ldots; \qquad \Delta_n > 0; \qquad \Delta_n > \Delta/2. \tag{6.101}$$

Here Δ is the determinant

$$\Delta = \begin{vmatrix} q_{nn}^{(0)} & q_{nn}^{(1)} & \ldots & q_{nn}^{(n-1)} \\ 1 & b_2 & \ldots & 0 \\ 0 & b_1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & b_n \end{vmatrix}, \tag{6.102}$$

which differs from the last Hurwitz determinant Δ_n only in its first row. The numbers $q_{nn}^{(r)}$ ($r = 0, \ldots, n-1$) are related to the elements a_{ij} of the correlation matrix by formulas (6.100).

We shall now show that the assumption $a(x) > 0$ ($x \neq 0$) is not essential. To do this we consider along with the system (6.94), another system

$$dX_1 = X_2\, dt; \qquad dX_2 = X_3\, dt; \qquad \ldots; \qquad dX_{n-1} = X_n\, dt,$$

$$dX_n = -\sum_{i=1}^{n} b_i X_{n-i+1}\, dt - \sum_{i,j=1}^{n} X_{n-i+1}\sigma_{ij}\, d\xi_j(t) + \varepsilon \sum_{j=1}^{n} X_j\, d\eta_j(t). \tag{6.103}$$

Here $\eta_1(t), \ldots, \eta_n(t)$ are assumed to be Wiener processes, independent of each other and of the processes $\xi_1(t), \ldots, \xi_n(t)$, and $\varepsilon > 0$ is a small parameter.

It is easy to see that the generator associated with the system (6.103) is

$$L_\varepsilon = L + \frac{1}{2}\varepsilon|x|^2 \frac{\partial^2}{\partial x_n^2}.$$

Since the quadratic form

$$a_\varepsilon(x) = a(x) + \varepsilon^2|x|^2$$

is positive definite for any $\varepsilon > 0$, it follows that the system (6.103) is asymptotically stable in mean square if and only if we have that

$$\Delta_1 > 0; \qquad \Delta_2 > 0; \qquad \ldots; \qquad \Delta_n > 0;$$

$$\Delta_n > \frac{\Delta_\varepsilon}{2} = \frac{\Delta + \varepsilon^2 \Delta_1}{2},$$

$$\Delta_1 = \begin{vmatrix} 1 & -1 & 1 & \ldots & (-1)^{n-1} \\ 1 & b_2 & b_1 & \ldots & 0 \\ 0 & b_1 & b_3 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & b_n \end{vmatrix}. \tag{6.104}$$

If conditions (6.93) are satisfied, the determinant Δ_1 is positive. This follows from the fact that the coefficient $\tilde{d}_{nn}$ of x_n^2 in a positive definite form $\widetilde{V}(x)$ satisfying the equation $L\widetilde{V} = -|x|^2$ is

$$\tilde{d}_{nn} = \frac{\Delta_1}{2\Delta_n}.$$

Consequently

$$\Delta_n > \frac{\Delta}{2}. \tag{6.105}$$

We now assume that the system (6.94) is asymptotically stable in mean square. Then, by Remark 6.5, the system (6.103) is also stable for all sufficiently small $\varepsilon > 0$, whence (6.104) holds. From (6.104) and (6.105) we get (6.101).

Now suppose that the inequalities (6.101) are satisfied. Then there exists a sufficiently small $\varepsilon > 0$ for which the inequalities (6.104) hold, i.e., the system (6.103) is asymptotically stable in mean square for this ε. By Theorem 6.3, there exists a positive definite quadratic form $W(x) = \sum_{i,j=1}^{n} w_{ij} x_i x_j$ such that the form $L_\varepsilon W$ is negative definite. But then the function $LW = L_\varepsilon W - \varepsilon^2 |x|^2 w_{nn}$ is also negative definite. Another application of Theorem 3.2 shows that the system (6.94) is stable in mean square.

We have thus proved

Theorem 6.13 *The system* (6.94) *is asymptotically stable in mean square if and only if conditions* (6.101) *are satisfied, where the determinant* Δ *is given by* (6.102) *and the numbers* q_{nn}^r $(r = 0, \ldots, n-1)$ *in the first row of* Δ *are expressed in terms of the coefficients* a_{ij} *by formulas* (6.100).

It is interesting to observe that the coefficients a_{ij} of the correlation matrix figuring in conditions (6.101) are those for which the sum $i + j$ is even. For example, for

second- and third-order systems our necessary and sufficient conditions for asymptotic stability in mean square are

$$n = 2: \quad b_1 > 0; \qquad b_2 > 0; \qquad 2b_1 b_2 > a_{11}b_2 + a_{22},$$

$$n = 3: \quad \begin{array}{l} b_1 > 0; \qquad b_3 > 0; \qquad b_1 b_2 > b_3; \\ 2(b_1 b_2 - b_3)b_3 > a_{11}b_2 b_3 + a_{33}b_1 + b_3(a_{22} - 2a_{13}). \end{array}$$

If the white noise processes $\dot{\eta}_1, \ldots, \dot{\eta}_n$ superimposed on the coefficients i of (6.92) are independent, i.e., $a_{ij} = 0$ for $i \neq j$, the determinant Δ assumes the particularly simple form

$$\Delta = \begin{vmatrix} a_{11} & -a_{22} & \ldots & (-1)^{n-2}a_{n-1,n-1} & (-1)^{n-1}a_{nn} \\ 1 & b_2 & \ldots & 0 & 0 \\ 0 & b_1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & b_n \end{vmatrix}.$$

Conditions (6.100), (6.101) and (6.102) are sufficient for the system (6.94) to be asymptotically p-stable when $p \leq 2$. Let us now determine sufficient conditions for p-stability when $p > 2$. We first assume that the quadratic form $a(x)$ is positive definite.

A necessary condition for asymptotic p-stability when $p = 2$, hence also when $p > 2$, is that there exists a positive definite quadratic form

$$V(x) = \sum_{i,j=1}^{n} d_{ij}x_i x_j,$$

satisfying (6.97). We set

$$W(x) = [V(x)]^{p/2}.$$

It is readily seen that

$$LW = \frac{p}{2}V^{p/2-2}\left\{ VL_0V + a(x)\left[d_{nn}V + (p-2)\left(\sum_{j=1}^{n} d_{nj}x_j \right)^2 \right] \right\}$$

$$= \frac{p}{2}V^{p/2-2}a(x)\left[(d_{nn} - 1)V + (p-2)\left(\sum_{j=1}^{n} d_{nj}x_j \right)^2 \right]. \tag{6.106}$$

By the well known inequality for positive definite self-adjoint matrices D (see, e.g., [85])

$$(Dx, y)^2 \leq (Dx, x)(Dy, y)$$

we get, by taking $y = (0, \ldots, 0, 1)$, that

$$\left(\sum_{j=1}^{n} d_{nj} x_j\right)^2 \leq d_{nn} V(x).$$

Using this relation, we conclude from (6.106) that

$$LW \leq \frac{p}{2} V^{p/2-1} a(x)[d_{nn}(p-1) - 1].$$

If $d_{nn}(p-1) > 1$, then it follows by Theorem 6.2 that the system (6.94) is asymptotically p-stable.

Thus, a sufficient condition for the system (6.94) to be p-stable ($p \geq 2$) is the occurrence of inequalities

$$\Delta_1 > 0; \qquad \ldots; \qquad \Delta_n > 0; \qquad \Delta_n > \frac{p-1}{2}. \qquad (6.107)$$

The first n of these are also necessary.

That the condition $\Delta_n > (p-1)/2$ is not necessary can be shown on examples. It is also readily seen that already conditions (6.107) are sufficient for p-stability ($p \geq 2$); there is then no need to assume that the quadratic form $a(x)$ is nonsingular.

6.11 Stochastic Stability in the Strong and Weak Senses

As before, let $X(t)$ be a solution of the linear system (6.61) with constant coefficients. We shall show that in a broad range of cases the, suitably normalized, limiting distribution of $\rho(t) = \ln|X(t)|$ as $t \to \infty$ is Gaussian. To do this we make use of the formula

$$\rho(t) = \rho_0 + \int_0^t Q(\Lambda^{\lambda_0}(s)) \, ds + \sum_{r=1}^{k} (\sigma_r \Lambda^{\lambda_0}(s), \Lambda^{\lambda_0}(s)) \, d\xi_r(s) \qquad (6.108)$$

proved in Sect. 6.7, and we assume for simplicity that the nondegeneracy condition (6.62) (or (6.67)) is satisfied. We then have

Theorem 6.14 *If condition* (6.62) *or* (6.67) *is satisfied and* $\mathrm{var}\, \rho(t) \to \infty$ *as* $t \to \infty$, *then*

$$\mathbf{P}\left\{\frac{\rho(t) - a^* t}{\sqrt{\mathrm{var}\, \rho(t)}} < x\right\} \xrightarrow[t \to \infty]{} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} \, dy.$$

Here $a^* = \int_{S_l} Q(\lambda) v(d\lambda)$.

Proof We have already seen (Sect. 6.7) that if condition (6.62) holds, then the transition probability of the process $\Lambda^{\lambda_0}(t)$ has a positive density. Hence, by the compactness of the phase space, the process $\Lambda^{\lambda_0}(t)$ satisfies Doeblin's condition. It is shown in [56, Chap. V], that for the transition probability function $P(\lambda, t, A)$ of the process $\Lambda(t)$ and any bounded measurable function $f(\lambda)$ on S_l

$$|P(\lambda, t, A) - \nu(A)| < ke^{-\alpha t}, \tag{6.109}$$

$$\left| \mathbf{E}\{f(\Lambda(t)) \mid \mathcal{N}_s\} - \int_{S_l} f(\lambda)\nu(d\lambda) \right| < k\|f\|e^{-\alpha(t-s)}, \tag{6.110}$$
$$\left(\|f\| = \max_{S_l} |f(\lambda)| \right)$$

(where $k > 0$ and $\alpha > 0$ are constants). Using (6.108) and (6.109), we easily see that for certain constants $c_i > 0$

$$|\mathbf{E}\rho(t) - a^*t| < |\rho_0| + c_1 \int_0^t e^{-\alpha s}\,ds < c_2. \tag{6.111}$$

It follows from (6.111) and the assumptions of the theorem that

$$\frac{\rho(t) - \mathbf{E}\rho(t)}{\sqrt{\operatorname{var}\rho(t)}} = \frac{\rho(t) - a^*t}{\sqrt{\operatorname{var}\rho(t)}} + o(1) \tag{6.112}$$

as $t \to \infty$. These relations, together with well known limit theorems for additive random functions (see [268]), imply the assertion of the theorem.   □

Remark 6.11 Later P. Baxendale [19] proved the stronger version of central limit theorem for $\rho(t)$, see details below, Sect. B.2 (Remark B.3).

In Sect. 1.5 we presented conditions for weak stochastic stability of systems whose right-hand side is perturbed by stochastic processes of a relatively general form. In fact, for equations with random right-hand side we studied conditions under which the solution $X^{s,x_0}(t)$ which satisfies the initial condition $X^{s,x_0}(s) = x_0$ fulfills the condition

$$\lim_{x_0 \to 0} \sup_{t>s} \mathbf{P}\{|X^{s,x_0}(t)| > \varepsilon\} = 0. \tag{6.113}$$

It is clear that this definition is in general weaker than the definition

$$\lim_{x_0 \to 0} \mathbf{P}\left\{ \sup_{t>s} |X^{s,x}(t)| > \varepsilon \right\} = 0 \tag{6.114}$$

we gave in Sect. 5.3. We claim that nevertheless a weakly stochastically stable linear system with constant coefficients satisfying the assumptions of Theorem 6.14 is also strongly stable. Indeed, if $a^* < 0$, the system is stable in both the strong and the weak sense. But if $a^* = 0$, it follows from Theorem 6.14 that the system is

unstable in the weak sense, since the probability of the event $\{|X(t)| < 1\}$ converges by Theorem 6.14 to $1/2$ as $t \to \infty$.

These arguments would seem to justify the conjecture that, in general, strong and weak stochastic stability are equivalent for autonomous stochastic systems perturbed by white noise. However, the following example refutes this conjecture.[11]

Example 6.7 Let φ be the angle-coordinate of a point on the circle. We consider a process on the circle, governed by the Itô equation

$$d\varphi(t) = \left[-2\sin^2\frac{\varphi}{2} + \sin^3\frac{\varphi}{2}\cos\frac{\varphi}{2} \right]dt - 2\sin^2\frac{\varphi}{2}\,d\xi(t).$$

This system has a unique equilibrium position $\varphi = 0$. It is readily seen by means of Itô's formula (3.30) that the solution of this system satisfying the initial condition $\varphi(0) = \varphi_0$ is the function

$$\varphi(t) = 2\operatorname{arccot}\left(t + \xi(t) + \cot\frac{\varphi_0}{2} \right). \tag{6.115}$$

It follows from (6.115) that the point $t = 0$ is unstable in the sense of (6.114). Indeed, $t + \xi(t) \to \infty$ almost surely as $t \to \infty$, and hence any path for which $\varphi_0 < 0$ is sufficiently small in absolute value almost surely describes an almost complete circle.

On the other hand, we claim that for any $\varepsilon > 0$ there exist sufficiently large numbers $C(\varepsilon)$ and $R(\varepsilon)$ such that for all $t \geq 0$

$$\mathbf{P}\{|t + \xi(t) - C(\varepsilon)| < R(\varepsilon)\} < \varepsilon. \tag{6.116}$$

Indeed,

$$\mathbf{P}\{|t + \xi(t) - C| < R\} = \Phi\left(\frac{C + R - t}{\sqrt{t}} \right) - \Phi\left(\frac{C - R - t}{\sqrt{t}} \right) \tag{6.117}$$

(where Φ is the normal distribution function with parameters $(0, 1)$). We set, for example, $C = T^3$. Then for $t \geq R^3/2$ the difference $2R/\sqrt{t}$ between the arguments in (6.117) satisfies the inequality

$$\frac{2R}{\sqrt{t}} \leq R^{-1/2} < \varepsilon$$

for sufficiently large R. But if $t < R^3/2$, then

$$\frac{C - R - t}{\sqrt{t}} > \frac{R^3/2 - R}{\sqrt{R^3/2}} \to \infty \quad \text{as } R \to \infty.$$

---

[11]The author's attention was drawn to the problem of the connection between strong and weak stochastic stability by N.N. Krasovskii.

These relations imply (6.116). From (6.115) and (6.116) it follows that the system is stable in the sense of definition (6.113); indeed, for sufficiently large $C(\varepsilon)$ we have the obvious inequality

$$\mathbf{P}\{|t + \xi(t) + C(\varepsilon)| < [C(\varepsilon)]^{1/3}\} < \varepsilon.$$

Thus a system which is weakly stable in the sense of (6.113) may be almost surely unstable. It would be interesting to construct an analogous example on the plane.

Chapter 7
Some Special Problems in the Theory of Stability of SDE's

7.1 Stability in the First Approximation[1]

Many problems concerning the stability of a nonlinear stochastic system can be reduced to problems about a linear system, obtained from the original system by dropping terms of higher than first order in x. This circumstance makes the study of stability for linear SDE's especially important.

The first theorem on stability of stochastic systems in the linear approximation was proved in [111], for the case of ODE with depending on jump Markov processes coefficients (see Sect. 5.1). They proved that the full system is stable in probability if the linearized system is exponentially stable in mean square. A similar result was established in [89, 90] for the diffusion-type processes considered in Chaps. 3–6. However, this result leaves unanswered the question of whether the linearization method is applicable to the stability theory of a broad range of systems; there exist linear systems which are almost surely asymptotically stable but not stable in mean square (see Sect. 6.3). This leads to the question of whether the full system is always stable whenever the corresponding linearized system has constant coefficients and is almost surely asymptotically stable. We intend to show that the answer to this question is positive.

We first consider the linear system

$$dX(t) = B(t)X \, dt + \sum_{r=1}^{k} \sigma_r(t) X \, d\xi_r(t) \tag{7.1}$$

with constant coefficients, i.e., $B, \sigma_1, \dots, \sigma_k$ are constant matrices.

---

[1] See [221], [128].

Theorem 7.1 *If the linear system* (7.1) *with constant coefficients is asymptotically stable almost surely* (*or in probability*), *and the coefficients of the system*

$$dX(t) = b(t, X) \, dt + \sum_{r=1}^{k} \sigma_r(t, X) \, d\xi_r(t) \tag{7.2}$$

satisfy an inequality

$$|b(t, x) - Bx| + \sum_{r=1}^{k} |\sigma_r(t, x) - \sigma_r x| < \gamma |x| \tag{7.3}$$

in a sufficiently small neighborhood of the point $x = 0$ *and with a sufficiently small constant* γ, *then the solution* $X = 0$ *of the system* (7.1) *is asymptotically stable in probability.*

Proof By Theorem 6.5, it will suffice to prove that if the system (7.1) is exponentially p-stable for some $p > 0$ and condition (7.3) is satisfied, then the system (7.2) is asymptotically stable in probability. We let

$$
\begin{aligned}
L_0 &= \frac{\partial}{\partial t} + \left(Bx, \frac{\partial}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r x, \frac{\partial}{\partial x} \right)^2, \\
L &= \frac{\partial}{\partial t} + \left(b(t, x), \frac{\partial}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t, x), \frac{\partial}{\partial x} \right)^2
\end{aligned}
\tag{7.4}
$$

denote the generators of the systems (7.1) and (7.2), respectively. By Theorem 6.2, there exists a function $V(t, x)$ such that for some $k_i > 0$

$$
k_1 |x|^p \le V(t, x) \le k_2 |x|^p; \qquad L_0 V(t, x) \le -k_3 |x|^p,
$$
$$
\left| \frac{\partial V}{\partial x} \right| < k_4 |x|^{p-1}, \qquad \left| \frac{\partial^2 V}{\partial x_i \partial x_j} \right| \le k_4 |x|^{p-2}.
\tag{7.5}
$$

By (7.3), (7.4) and (7.5), we have

$$
LV = L_0 V + \left(b(t, x) - Bx, \frac{\partial}{\partial x} \right) V
$$
$$
+ \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t, x) - \sigma_r x, \frac{\partial}{\partial x} \right) \left(\sigma_r(t, x) + \sigma_r x, \frac{\partial}{\partial x} \right) V
$$
$$
\le -k_3 |x|^p + \gamma k_4 |x|^p + \gamma k_5 |x|^p
\tag{7.6}
$$

in a sufficiently small neighborhood of $x = 0$.

The constant k_5 in this inequality depends only on k_4 and on the supremum of absolute values of the coefficients in (7.1).

It follows from (7.6) that the function LV is negative definite in a sufficiently small neighborhood of $x = 0$, provided $\gamma < k_3/(k_4 + k_5)$. Moreover, according to (7.5), V is positive definite and has an infinitesimal upper limit. Applying Corollary 5.1 to the function V, we get the assertion of the theorem.   □

If the coefficients of the system (7.1) are time-dependent, we have the analogous

Theorem 7.2 *Suppose that the coefficients of the system* (7.1) *are bounded functions of time, the trivial solution of* (7.1) *is stable in the large uniformly in t, and condition* (7.3) *is satisfied with a sufficiently small constant γ. Then the trivial solution of* (7.2) *is asymptotically stable in probability.*

Proof By Theorem 6.7, it suffices to prove that if the system (7.1) is exponentially p-stable and condition (7.3) holds, then the assertion of the theorem is valid. This we did in the proof of Theorem 7.1.   □

Remark 7.1 It is clear from the proofs of Theorems 7.1 and 7.2 that the constant γ in condition (7.3) depends only on k_3, k_4 and $\sup_{t>0} \|\sigma_r(t)\|$.

Remark 7.2 It follows from Theorems 7.1 and 7.2 that the system (7.2) is asymptotically stable in probability if the linearized system

$$dX(t) = \frac{\partial b(t,0)}{\partial x} X \, dt + \sum_{r=1}^{k} \frac{\partial \sigma_r(t,0)}{\partial x} X \, d\xi_r(t)$$

is stable in the large uniformly in t, and the derivatives $\partial b/\partial x$ and $\partial \sigma_r/\partial x$ are bounded and uniformly (in t) continuous in x for $x = 0$.

7.2 Instability in the First Approximation

We first recall some well known results for the deterministic case (see [188], [191]), confining ourselves to systems with constant coefficients.

Theorem 7.3 (Lyapunov's Theorem) *Suppose that at least one of the roots of the characteristic equation of the system*

$$\frac{dX}{dt} = BX \tag{7.7}$$

has positive real part. Let the vector $\varphi(x)$ be such that $|\varphi(x)| < A|x|^2$.
Then the solution $X \equiv 0$ of the equation

$$\frac{dX}{dt} = BX + \varphi(X) \tag{7.8}$$

is unstable.

Malkin observed in [191] that Lyapunov's arguments in fact yield a more general result:

Theorem 7.4 (Malkin's Theorem) *If at least one of the roots of (7.7) has positive real part, then the trivial solution of the system (7.8) is unstable if $|\varphi(x)| < \gamma|x|$, where γ is a sufficiently small constant which depends only on the coefficients of a positive definite quadratic form satisfying the assumptions of Lyapunov's second instability theorem.*

It will be clear from the sequel that the situation is far more complicated in regard to stochastic systems. In particular, the examples given in Sect. 7.3 will show that the analog of Malkin's theorem fails to hold.

We shall first prove that the analog of Theorems 7.1 and 7.2 for instability is valid, provided the linearized system is unstable in a sufficiently strong sense. Theorem 6.4 implies that if the system (7.1) is exponentially q-unstable, then there exists a function $V(t, x)$ such that

$$k_1|x|^{-q} \le V(t, x) \le k_2|x|^{-q}; \qquad L_0V \le -k_3|x|^{-q},$$

$$\left|\frac{\partial V}{\partial x}\right| \le k_4|x|^{-q-1}; \qquad \left|\frac{\partial^2 V}{\partial x_i \partial x_j}\right| < k_4|x|^{-q-2} \quad (i, j = 1, \ldots, l) \tag{7.9}$$

for certain constants $k_i > 0$.

Theorem 7.5 *Let the coefficients of the linear system (7.1) be bounded functions of time, and let the trivial solution of this system be exponentially q-unstable for some $q > 0$. Moreover, suppose that inequality (7.3) holds with a sufficiently small constant γ, depending only on $\sup_{t>0} \|\sigma_r(t)\|$ and the constants k_1, k_2, k_3, k_4 figuring in (7.9). Then the solution $X(t) \equiv 0$ of the system (7.2) is unstable in probability.*

Proof It follows from Theorem 6.4 that under the above assumptions there exists for the system (7.1) a function satisfying inequalities (7.9). Hence, as in the case of (7.6), we see that

$$LV \le -k_3|x|^{-q} + \gamma k_4|x|^{-q} + \gamma k_5|x|^{-q},$$

in a sufficiently small neighborhood of $x = 0$, where k_5 depends on k_4 and $\sup_{t>0} \|\sigma_r(t)\|$. It follows now that for sufficiently small γ, the function V satisfies all the assumptions of Theorem 5.6 (see the remark following that theorem). This completes the proof. $\qquad\square$

From this theorem and Theorems 6.6 and 6.8, we get the following result.

Theorem 7.6 *If for any $x \ne 0$, $A > 0$, the solutions of the linear system (7.1) satisfy the identity*

$$\lim_{T \to \infty} \sup_{s>0} \mathbf{P}\left\{ \inf_{u>s+T} |X^{s,x}(y)| < A \right\} \equiv 0, \tag{7.10}$$

and the elements of the matrices B, $\sigma_1, \ldots, \sigma_k$ are bounded, then the solution $X(t) \equiv 0$ is unstable in probability for all systems of type (7.2) whose coefficients satisfy condition (7.3) with sufficiently small γ.

Theorem 7.7 *If the system (7.1) has constant coefficients, the assertion of Theorem 7.6 remains valid if assumption (7.10) is replaced by the requirement that for all $x \neq 0$*

$$\mathbf{P}\{|X^{s,x}(t)| \to \infty \ as \ t \to \infty\} = 1. \tag{7.11}$$

A comparison of the theorems of Lyapunov and Malkin with Theorem 7.7 shows that for deterministic systems the latter furnishes a very poor result. Whereas according to the Lyapunov–Malkin theorems it is sufficient that at least one root of the characteristic equation of the linear system has a positive real part, Theorem 7.7 is valid for a deterministic linear system only when the real parts of *all* roots of the characteristic equation are positive (see Sect. 6.3). Nonetheless, if the system (7.1) with constant coefficients is nondegenerate in the sense that

$$\sum_{r=1}^{k} (\sigma_r x, \lambda)^2 > 0 \tag{7.12}$$

for all non-zero vectors x and λ, then Theorem 7.7 implies the

Corollary 7.1 *If inequality (7.12) is satisfied, then the assertion of Theorem 7.7 holds if conditions (7.11) are satisfied for at least one value of x.*

This corollary is obvious if we observe that, by Theorems 6.11 and 6.12, condition (7.11) holds for one x if and only if it holds for all $x \neq 0$.

There are two questions arising naturally in connection with the last theorem and its corollary.

1. Can we replace assumption (7.11) in Theorem 7.7 by the weaker assumption

$$\sup_{t>s} |X^{s,x}(t)| = \infty \quad \text{a.s.?}$$

2. Can the assertion of Theorem 7.7 be proved under the assumption that (7.11) holds for at least one $x \neq 0$, but without the nondegeneracy condition (7.12)?

We shall see in Sect. 7.3 that the answers to both these questions are in general negative.

7.3 Two Examples

Example 7.1 We again consider the one-dimensional system

$$dX(t) = b(t, X)\, dt + \sigma(t, X)\, d\xi(t), \tag{7.13}$$

such that the linearized system

$$dX(t) = b_0 X \, dt + \sigma_0 X \, d\xi(t) \tag{7.14}$$

has constant coefficients. If $b_0 < \sigma_0^2/2$, we can apply Theorem 7.1, and if $b_0 > \sigma_0^2/2$, we can use Theorem 7.7. If $b_0 = \sigma_0^2/2$, the linear system is unstable, but not asymptotically q-unstable for any $q > 0$ (see Sect. 6.1) and, moreover, we have

$$\mathbf{P}\left\{ \sup_{t > s} |X^{s,x}(t)| = \infty \right\} = 1 \tag{7.15}$$

for $x \neq 0$. It follows also from the results of Sect. 6.1 that if $b_0 = \sigma_0^2/2$, then the system

$$dX(t) = (b_0 - \gamma)X \, dt + \sigma_0 X \, d\xi(t)$$

is asymptotically stable for any $\gamma > 0$.

This implies that the answer to the first of the questions posed at the end of Sect. 7.2 is negative. Thus the analog of Malkin's theorem is false here. However, the solution $X(t) \equiv 0$ of the system (7.13) is nevertheless unstable, if we assume that the differences $b(t, x) - b_0 x$ and $\sigma(t, x) - \sigma_0 x$ tend to zero sufficiently rapidly as $x \to 0$. Indeed, suppose that $b_0 = \sigma_0^2/2$ and that for some $k > 0$, $\alpha > 0$

$$|b(t, x) - b_0 x| + |\sigma(t, x) - \sigma_0 x| < k|x|^{1+\alpha}. \tag{7.16}$$

Consider the auxiliary function $V(x) = \ln \ln(1/|x|)$. The reader will easily verify that in this case $V \to \infty$ as $x \to 0$ and $\inf_{\varepsilon < |x| < \delta} LV < 0$ for sufficiently small fixed $\delta > 0$ and any $\varepsilon < \delta$. The instability of (7.13) under the assumptions $b_0 = \sigma_0^2/2$ and (7.16) now follows from Theorem 5.6.

We have thus shown in the one-dimensional case that if the linear system with constant coefficients satisfies (7.15), and the full system is nearly linear in the sense of (7.16), then the latter is unstable in probability.

It would be interesting to know whether this remains true in the multi-dimensional case.

Example 7.2 Let $\Psi(z)$ denote a differentiable function of the real variable z, with compact support and bounded together with its first derivative. Suppose further that

$$\Psi(0) = 0; \qquad \Psi'(0) = -3; \qquad |\Psi(z)| < 1. \tag{7.17}$$

Using this function, we construct a Markov process on the plane, which is the solution of the system of Itô equations

$$dX_1(t) = \left[X_1 + \varepsilon X_2 \Psi\left(\frac{X_1}{\varepsilon X_2} \right) \right] dt + \sigma X_1 \, d\xi_1(t),$$

$$dX_2(t) = -X_2 \, dt + \delta X_1 \, d\xi_2(t). \tag{7.18}$$

We first observe that for any $\varepsilon > 0$, $\delta > 0$ the coefficients of (7.18) have bounded derivatives with respect to x_1, x_2, and consequently they satisfy the existence conditions (Theorem 3.4). Further, for small ε and δ the coefficients of (7.18) are close to those of the deterministic system

$$\frac{dx_1}{dt} = x_1; \qquad \frac{dx_2}{dt} = -x_2 \tag{7.19}$$

in the sense of (7.3), where the constant γ in (7.3) may be assumed equal to $\min(\varepsilon, \delta)$. Finally, it is clear that the solutions of the system (7.19), except those for which $x_1(0) = 0$, have absolute values diverging to infinity as $t \to \infty$. Nevertheless, we can prove that for any $\varepsilon > 0$ and $\delta > 0$ the solution $X(t) \equiv 0$ of (7.18) is asymptotically stable in the large. This will furnish a negative answer to the second question at the end of Sect. 7.2.[2]

We shall use in the proof the fact that all the coefficients of (7.18) are homogeneous functions of degree 1 and therefore the projection of the process $X(t)$ on the circle $|x| = 1$ is also a Markov process (see Sect. 6.7).[3]

As in Sect. 6.8, we introduce the new variables

$$r(t) = \frac{1}{2}\ln(X_1^2(t) + X_2^2(t)) = \ln|X(t)|;$$

$$\varphi(t) = \arctan\frac{X_2(t)}{X_1(t)}$$

and apply Itô's formula (3.30). The result is

$$d\varphi(t) = -\left[2\sin\varphi\cos\varphi + \varepsilon\sin^2\varphi\,\Psi\left(\frac{\cot\varphi}{\varepsilon}\right)\right]dt$$

$$+ \delta(\cos^2\varphi\,d\xi_2(t) - \sin\varphi\cos\varphi\,d\xi_1(t)), \tag{7.20}$$

$$dr(t) = \left[\cos^2\varphi - \sin^2\varphi + \varepsilon\sin\varphi\cos\varphi\,\Psi\left(\frac{\cot\varphi}{\varepsilon}\right)\right]dt$$

$$+ \delta(\cos^2\varphi\,d\xi_1(t) + \sin\varphi\cos\varphi\,d\xi_2(t)). \tag{7.21}$$

The diffusion coefficient of the Markov process $\varphi(t)$ on the circle $0 \le \varphi < 2\pi$ vanishes only at the points $\varphi_1 = \pi/2$, $\varphi_2 = 3\pi/2$. In view of (7.17), this means that $\varphi = \pi/2$ and $\varphi = 3\pi/2$ are solutions of (7.20). We claim that these solutions are

---

[2]In connection with this Nevelson [217] has proved that an unstable linear system perturbed by small linear random perturbations is also unstable.

[3]This observation, made in the first edition in 1969, was later used by many authors for evaluation of Lyapunov exponent, moment Lyapunov exponent, stability index for SDE with homogeneous of order 1 coefficients, see e.g., [13], [226], [1], [258].

stable. To prove this, we investigate the first-approximation equation in the neigh-
borhood of the point $\varphi = \pi/2$. By (7.17), this equation is

$$d(\varphi - \pi/2) = -(\varphi - \pi/2)\,dt + \delta(\varphi - \pi/2)\,d\xi_1(t).$$

Since the first-approximation equation is asymptotically stable, it follows from The-
orem 7.1 that the solution $\varphi = \pi/2$ of (7.20) is stable in probability. The stability of
the solution $\varphi = 3\pi/2$ is proved in a similar fashion. The diffusion coefficient of the
process $\varphi(t)$ is positive for $\varphi \neq \varphi_i$. Hence, reasoning as in the proof of Theorem 5.7,
we see that $\varphi(t)$ has a limit as $t \to \infty$ for any initial condition. This limit is either
$\pi/2$ or $3\pi/2$.

   Thus, by (7.21) and Lemma 6.8, we see that

$$\lim_{t\to\infty} \frac{\ln|X(t)|}{t} = \lim_{t\to\infty} \frac{1}{t} \int_0^t \Bigg[\cos^2\varphi(s) - \sin^2\varphi(s)$$

$$+ \varepsilon \sin\varphi(s)\cos\varphi(s)\,\Psi\left(\frac{\cot\varphi(s)}{\varepsilon}\right)\Bigg]ds = -1.$$

Thus, $\mathbf{P}\{|X(t)| \to 0 \text{ as } t \to \infty\} = 1$ and the solution of (7.18) is stable in the large,
as required.

   To conclude this section, we note that in this example condition (7.16) does not
hold for any $\alpha > 0$. It is quite probable that the theorem on instability in the first
approximation can be proved on the assumption that (7.11) holds for at least one
value of x and the full system is nearly linear in the sense of (7.16). This would
be a natural generalization of the theorem of Lyapunov quoted at the beginning of
Sect. 7.2.

7.4 Stability Under Damped Random Perturbations

Consider the one-dimensional SDE

$$dX(t) = -X\,dt + \sigma(t)\,d\xi(t); \qquad X(0) = x_0. \tag{7.22}$$

It is easy to see that the solution of this SDE is

$$X(t) = x_0 e^{-t} + \int_0^t e^{s-t}\sigma(s)\,d\xi(s).$$

Hence

$$\mathbf{E}|X(t) - x_0 e^{-t}|^2 = \int_0^t e^{2(s-t)}\sigma^2(s)\,ds \le e^{-t}\int_0^{t/2}\sigma^2(s)\,ds + \int_{t/2}^t \sigma^2(s)\,ds.$$

It follows from this inequality that, although $X(t) \equiv 0$ is not a solution of the system
(7.22), it is nonetheless true that any solution of (7.22) tends to zero as $t \to \infty$,

provided that

$$\int_0^\infty \sigma^2(s)\, ds < \infty.$$

It is natural to expect a similar situation to obtain in a broader range of cases: Sufficiently rapidly damped persistent random perturbations do not change the stability of an asymptotically stable system. Let us prove a result going in that direction.

Theorem 7.8 *Consider the system*

$$\frac{dx}{dt} = F(t, x) \tag{7.23}$$

in $\mathbb{R}^l$. Suppose that there exists for this system a positive definite and admitting an infinitesimally small upper limit function $V(t, x)$ satisfying

$$\frac{d^0 V}{dt} = \frac{\partial V}{\partial t} + \left(F, \frac{\partial V}{\partial x} \right) < -\alpha(t)\varphi(t, x) < 0, \tag{7.24}$$

$$\inf_{t>0} V(t, x) \to \infty \quad as \ |x| \to \infty, \tag{7.25}$$

and such that for every $0 < \varepsilon < R < \infty$,

$$\inf_{\varepsilon < |x| < R} \varphi(t, x) > 0 \quad for \ T > T_{\varepsilon, R} \tag{7.26}$$

and

$$\int_0^\infty \alpha(t)\, dt = \infty \tag{7.27}$$

hold. Suppose further that the coefficients $\sigma_r(t, x)$ of the SDE

$$dX(t) = F(t, X(t))\, dt + \sum_{r=1}^k \sigma_r(t, X(t))\, d\xi_r(t) \tag{7.28}$$

satisfy for a certain constant K_1 and a certain positive and integrable on $[0, \infty)$ function $g(t)$

$$\sum_{r=1}^k \sum_{i,j=1}^l \sigma_r^i \sigma_r^j \frac{\partial^2 V}{\partial x_i \partial x_j} = \sum_{r=1}^k \left(\sigma_r(t, x), \frac{\partial}{\partial x} \right)^2 V(t, x) \le (V(t, x) + K_1)g(t). \tag{7.29}$$

Then every solution of (7.28) has almost surely limit 0 as $t \to \infty$.

Proof We let the generator L of the process (7.28) act on the function

$$W(t, x) = (V(t, x) + K_1) \exp\left\{ \int_t^\infty g(s)\, ds \right\}.$$

Evidently

$$LW = (LV - gV - K_1 g) \exp\left\{ \int_t^\infty g(s)\,ds \right\}$$

$$\leq -\alpha(t)\varphi(t,x) \exp\left\{ \int_t^\infty g(s)\,ds \right\}. \tag{7.30}$$

(7.25), (7.30) and the integrability of the function $g(t)$ imply that

$$LW \leq 0; \qquad \inf_{t\geq 0} W(t,x) \to \infty \quad \text{as } |x| \to \infty.$$

Therefore (see Sect. 5.2 and Theorem 3.5) $W(t, X(t))$ is a bounded from below supermartingale. Theorem 5.1 implies now that almost surely there exists the limit

$$\eta = \lim_{t\to\infty} W(t, X(t)) = \lim_{t\to\infty} V(t, X(t)) + K_1. \tag{7.31}$$

Since $V(t,x)$ is positive definite and (7.31) holds, the theorem will be proved provided we show that $\eta = K_1$ holds almost surely.

For this purpose let us consider the domain $U_T(\varepsilon, R) = \{\varepsilon < |x| < R\} \times (t > T)$. Then (7.30) and the assumptions of the theorem imply that for a suitable choice of $T = T_{\varepsilon, R}$ the inequality $LW \leq -\alpha(t)\delta_{\varepsilon, R}$ holds in the domain $U_T(\varepsilon, R)$. Thus we may conclude by Theorem 3.9 that for every point $(s, x) \in U_T(\varepsilon, R)$ the moment $\tau(\varepsilon, R)$ at which the path of the process $X(t)$ exits from the domain $\varepsilon < |x| < R$ is almost surely finite. (More exactly, we should apply Theorem 3.9 only in the case when $\tau(\varepsilon, R) \geq T$, considering for $t \geq T$ the paths which leave the point $X(T)$ at the moment T.) Thus we may now apply Lemma 3.9 whose assumptions are satisfied for the function $W(t, x)$ in order to conclude that the process $X(t)$ is recurrent in the domain $|x| < \varepsilon$ for every $\varepsilon > 0$ and $x \in \mathbb{R}^l$.

It is now easy to establish the equality $\mathbf{P}\{\underline{\lim}_{t\to\infty} |X(t)| = 0\} = 1$. Indeed, suppose that with positive probability we have $\underline{\lim}_{t\to\infty} |x^{s_0, x_0}(t)| > \delta$ for some $s_0 > 0$ and $x_0 \in \mathbb{R}^l$. Then there exists a number $T > 0$ such that

$$\mathbf{P}\left\{ \inf_{t > T} |X^{s,x}(t)| > \delta \right\} > p/2.$$

The last inequality contradicts the recurrence property of the process $X^{T,x}(t)$ as related to the domain $\{|x| < \delta\}$ and for all $|x| > \delta$.

Since $V(t, x)$ has an infinitesimally small upper limit, (7.31) implies that

$$\mathbf{P}\{\eta = K_1\} = 1.$$

As we have noted above, this suffices to complete the proof of the theorem. $\qquad\square$

Remark 7.3 The assumption that there exists a function V satisfying conditions (7.24), (7.25) and (7.27) is not very restrictive. For a wide class of stable in the large systems (7.23) it is possible to demonstrate the existence of a function V with the

above properties (see [155]). Condition (7.29) is satisfied for instance if V grows not slower than $|x|^2$ as $|x| \to \infty$, it has bounded second derivatives with respect to the space variables, and moreover

$$\sum_{r=1}^{k} |\sigma_r(t, x)|^2 < (|x|^2 + 1)g(t). \tag{7.32}$$

Remark 7.4 Theorem 7.8 throws some light upon the effect of damped random perturbations on a deterministic system which is stable in the large. If the system is dissipative in the sense of Sect. 1.2, one can prove in an analogous fashion that the constant R figuring in the definition of dissipativity satisfies the equality

$$\mathbf{P}\left\{ \varlimsup_{t \to \infty} |X(t, \omega)| < R \right\} = 1.$$

A more delicate analysis shows that in this case for sufficiently "well behaved" systems (7.23) the solution of (7.28) almost surely converges to one of the solutions of the deterministic system (7.23) as $t \to \infty$. We shall not go into the details here.

7.5 Application to Stochastic Approximation[4]

In [237] proposed an iterative procedure for the determination of the roots of the regression equation; they called this method stochastic approximation. Let us briefly describe it here.

Let $Y(x, \omega)$ be a family of random variables depending on a parameter x such that $\mathbf{E}Y(x, \omega) = R(x)$. Suppose that the distribution of $Y(x, \omega)$ is unknown to an "observer", who can only carry out certain "measurements" $y_n(x, \omega)$ of the variable for arbitrary values of the parameter. The problem is to determine a value of the parameter $x = x^0$ for which the function $R(x)$ assumes a prescribed value α.

The procedure proposed by Robbins and Monro to determine x^0 is as follows. Choose a sequence of positive numbers a_n such that

$$\sum_{n=1}^{\infty} a_n = \infty; \qquad \sum_{n=1}^{\infty} a_n^2 < \infty. \tag{7.33}$$

Fixing x_1 arbitrarily, define a sequence of numbers $x_2, \ldots, x_n, \ldots$ by the recursive relation

$$x_{n+1} = x_n - a_n(y_n(x_n, \omega) - \alpha). \tag{7.34}$$

Under certain assumptions on the distribution of $Y(x, \omega)$, it can be shown that $x_n(\omega) \to x^0$ almost surely as $n \to \infty$. This result has since been generalized in

---

[4]A more detailed discussion of these problems is given in Chaps. IV and V of the author's joint book with Nevelson [138].

various directions. A detailed bibliography of the literature up to 1965 can be found in [68].

A natural continuous analog of the procedure (7.34) is

$$\frac{dX}{dt} = -a(t)[Y(X(t), t, \omega) - \alpha],$$ (7.35)

where $Y(x, t, \omega)$ is a "stochastic process" with independent values and expectation $R(x)$. If we assume, as before, that the solution $X(t)$ is a continuous stochastic process,[5] then this equation can be interpreted as the Itô equation

$$dX(t) = -a(t)[(R(X(t)) - \alpha)\,dt + \sigma(X(t))\,d\xi(t)]$$ (7.36)

in $\mathbb{R}^l$.

We shall show that under certain assumptions concerning the functions $a(t)$, $R(x)$ and $\sigma(x)$ this process converges almost surely to a root of the equation $R(x) = \alpha$.

Theorem 7.9 *Given two differentiable functions $R(x)$ and $\sigma(x)$, suppose that there exist a twice continuously differentiable function $V(x)$ and a constant $k > 0$ such that*

$$\left.\begin{aligned} V(x) \to \infty \quad &as\ |x| \to \infty; \\ V(x^0) = 0; \qquad &V(x) > 0 \quad for\ x \neq x^0 \end{aligned}\right\}$$ (7.37)

and

$$(R(x) - \alpha)\frac{dV}{dx} > 0 \quad for\ x \neq x^0,$$

$$\sigma^2(x)\frac{d^2V}{dx^2} < k(V(x) + 1).$$

Then the process (7.36) satisfies the equality

$$\mathbf{P}\left\{\lim_{t \to \infty} X(t) = x^0\right\} = 1$$

for all continuous positive functions $a(t)$ such that

$$\int_0^\infty a(t)\,dt = \infty; \qquad \int_0^\infty a^2(t)\,dt < \infty.$$ (7.38)

Proof The proof follows at once from Theorem 7.8, if we set

$$(R(x) - \alpha)a(t) = -F(t, x - x^0).$$ $\qquad\square$

---

[5]This restriction is not natural, and it can be eliminated by recourse to the stochastic differential equation of a Markov jump process (see, e.g., [251]). However, jump processes are beyond the scope of this book.

Corollary 7.2 *The assumptions of the theorem evidently hold for the function* $V(x) = (x - x^0)^2$, *if* $R(x)$ *and* $\sigma(x)$ *are such that*

$$(x - x^0)(R(x) - \alpha) > 0 \quad \text{for } x \neq x^0,$$

$$\sigma^2(x) < k(x^2 + 1).$$

For the discrete-time case analogous sufficient conditions for convergence of stochastic approximations were given in [94].

Theorem 7.8 also yields convergence conditions for stochastic approximations in many dimensions. We state the result.

Theorem 7.10 *The multi-dimensional stochastic approximation process*

$$dX(t) = -a(t) \left[(R(X(t)) - \alpha) \, dt + \sum_{r=1}^{k} \sigma_r(X(t)) \, d\xi_r(t) \right] \tag{7.39}$$

converges almost surely to a solution of the equation

$$R(x) = \alpha, \tag{7.40}$$

if conditions (7.38) *are satisfied and there exists a function* $V(x)$ *satisfying condition* (7.37) *and the conditions*

$$\left(R(x) - \alpha, \frac{\partial V}{\partial x} \right) > 0 \quad \text{for } x \neq x^0, \tag{7.41}$$

$$\sum_{r=1}^{k} \sum_{i,j=1}^{l} \sigma_r^i(x) \sigma_r^j(x) \frac{\partial^2 V}{\partial x_i \partial x_j} < k_1(V(x) + 1) \tag{7.42}$$

for some constant k_1.

Remark 7.5 If condition (7.41) is satisfied only in a neighborhood of infinity, it can be shown that the process (7.39) converges to an invariant set of the system $dx/dt = -R(x) + \alpha$.

7.6 Stochastic Approximations when the Regression Equation Has Several Roots

Condition (7.41) guarantees the uniqueness of the solution of the equation $R(x) = \alpha$. However, it is interesting to study the properties of the Robbins–Monro procedure when this equation has several roots. We shall show that Lyapunov functions are also applicable in this case.

Fig. 7.1
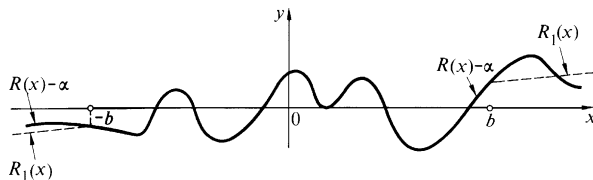
Not striving for maximal generality, we shall limit the discussion in this section to the case of a point x on the line $(-\infty < x < \infty)$. We shall assume that the coefficient $\sigma(x)$ of (7.36) satisfies

$$\sigma^2(x) < k(x^2 + 1) \tag{7.43}$$

for some constant $k > 0$.

Theorem 7.11 *Suppose that the set $A = \{x : R(x) = \alpha\}$ consists of finitely many points $x_1^{(0)}, \ldots, x_n^{(0)}$, the derivative $R'(x)$ is continuous, and condition (7.43) is satisfied. Suppose moreover that*

$$(R(x) - \alpha)x > 0 \quad \text{for } |x| > b \tag{7.44}$$

holds for some number $b > 0$.

Then for any function $a(t) > 0$ satisfying the conditions (7.38) the Robbins–Monro process $X(t)$ defined by (7.36) converges almost surely to a point of the set A.

Proof We set

$$R_1(x) = \begin{cases} R(x) - \alpha & \text{for } |x| \le b, \\ \frac{R(b) - \alpha}{b} x & \text{for } x > b, \\ \frac{R(-b) - \alpha}{-b} x & \text{for } x < -b, \end{cases}$$

$$W(x) = \int_0^x R_1(y)\, dy;$$

$$V(t, x) = (W(x) + k_1) \exp\left\{ \gamma \int_t^\infty a^2(s)\, ds \right\},$$

where the constants $k_1 > 0$ and $\gamma > 0$ will be determined later. The functions $R_1(x)$ and $R(x)$ are illustrated in Fig. 7.1 and the function $W(x)$ in Fig. 7.2.

Applying to the function V the generator

$$L = \frac{\partial}{\partial t} - a(t)(R(x) - \alpha)\frac{\partial}{\partial x} + \frac{1}{2} a^2(t)\sigma^2(x)\frac{\partial^2}{\partial x^2}$$

Fig. 7.2
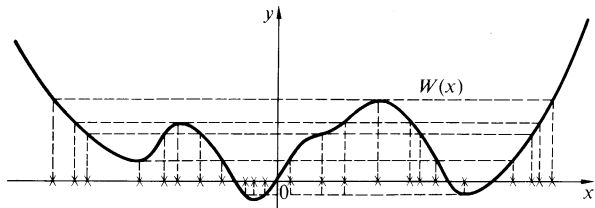
of the process $X(t)$, we get

$$LV = -(W(x) + k_1)\gamma a^2(t) - a(t)(R(x) - \alpha)R_1(x)\exp\left\{\gamma\int_t^\infty a^2(s)\,ds\right\}$$

$$+ \frac{1}{2}a^2(t)\sigma^2(x)R_1'(x)\exp\left\{\gamma\int_t^\infty a^2(s)\,ds\right\}.$$

By the above construction of the function $R_1(x)$, the derivative $R_1'(x)$ is bounded and the product $(R(x) - \alpha)R_1(x)$ is non-negative (see (7.44)). In addition, the function $W(x)$ increases like a parabola when $|x| \to \infty$. Hence, by (7.43), it follows that for a suitable choice of k_1 and γ we have $LV \le 0$ for $t \ge t_0$; in fact, we have the even stronger inequality

$$LV \le -a(t)(R(x) - \alpha)R_1(x). \tag{7.45}$$

Thus the function $V(t, X(t))$ is a supermartingale for $t \ge t_0$ (see Sect. 5.2). Hence the limit

$$\lim_{t\to\infty} V(t, X(t)) = \xi \tag{7.46}$$

is almost surely finite. Moreover, it follows from (7.45) and (3.37) that

$$\int_{t_0}^t a(s)\mathbf{E}[R(X(s) - \alpha)R_1(X(s))]\,ds \le \mathbf{E}V(t_0, X(t_0)) \tag{7.47}$$

for $t > t_0$. The integrand on the left of (7.47) is non-negative, and the function $a(s)$ is not integrable on $[t_0, \infty]$. Thus, there exists a sequence $t_n \to \infty$ such that

$$\lim_{n\to\infty} \mathbf{E}[(R(X(t_n)) - \alpha)R_1(X(t_n))] = 0.$$

It follows from this equality and from Chebyshev's inequality that the sequence $R(X(t_n)) - \alpha$ converges to zero in probability. Then, as is well known, there exists a subsequence t_{n_k} such that

$$\lim_{k\to\infty} R(X(t_{n_k})) = a \quad \text{(a.s.)}.$$

Hence, it follows by (7.46) that the random variable ξ in (7.46) almost surely takes on only values from the finite set $W(x_1^{(0)}) + k_1, \ldots, W(x_n^{(0)}) + k_1$. Let A_1 be the set of points x such that $W(x) + k_1$ takes on one of these values. The assumptions of the

theorem evidently imply that A_1 is a finite set (see Fig. 7.2), in which the elements of A_1 are marked by crosses, and $A \subset A_1$. It follows from (7.46) that the process $X(t)$ converges to the set A_1 as $t \to \infty$. Moreover $X(t)$ almost surely converges to only one element of A_1, for otherwise it would follow from the continuity of $X(t)$ that (7.46) does not hold. Finally let us prove that the probability of the event $B = \{X(t) \to x_0 \text{ as } t \to \infty\}$ is zero if $x_0 \in A_1 \setminus A$. In fact, B implies the event

$$\lim_{t \to \infty} (R(X(t)) - \alpha) R_1(X(t)) = (R(x_0) - \alpha)^2 > 0.$$

Hence, by (7.38), we get

$$B \subset \left\{ \int_{t_0}^{\infty} a(s)(R(X(s)) - \alpha) R_1(X(s)) \, ds = \infty \right\}.$$

This, together with (7.47), implies that $P(B) = 0$. Theorem 7.11 is proved.   □

Let A_2 denote the set of points $x \in A$ at which the function $R(x) - \alpha$ changes sign, from positive to negative. The next theorem shows that under certain additional assumptions the points of A_2 cannot be limits of the stochastic approximation process (7.36). For the discrete-time case this was conjectured by Fabian [69, 70]. Similar questions have been considered by Krasulina [162].

Theorem 7.12 *Under the assumptions of Theorem 7.11, let x^* be a point of the set A_2 such that $R'(x^*) < 0$ and $\sigma^2(x^*) > 0$. Assume moreover that the function $a(t)$ satisfies $|a'(t)| < ka^2(t)$ ($k > 0$ constant). Then x^* cannot be a limit point of the stochastic approximation process (7.36).*

For the proof we need the following two lemmas.

Lemma 7.1 *Let $X(t)$ be a diffusion process with generator L and let D be a bounded domain. Then $X(t)$ almost surely exits from D in a finite time, if there exists a function $V(t, x)$ such that in $I \times D$*

$$LV \leq 0, \tag{7.48}$$

$$\inf_{x \in D} V(t, x) \to \infty \quad \text{as } t \to \infty \tag{7.49}$$

Proof Let τ denote the first exit time from D of the path of the process and let $\tau(t) = \min(\tau, t)$. As mentioned in Sect. 5.2, the process $V(\tau(t), X(\tau(t)))$ is a supermartingale. By Theorem 5.1

$$\lim_{t \to \infty} V(\tau(t), X(\tau(t))) = \xi$$

exists and is finite. The finiteness of ξ and relation (7.49) imply the assertion.   □

Lemma 7.2 *The function*

$$W(z) = \int_0^z dv \int_0^v \frac{e^{u-v}}{\sqrt{uv}}\, du \tag{7.50}$$

is a solution of the equation

$$zw'' + \left(z + \frac{1}{2}\right)w' = 1. \tag{7.51}$$

Moreover, W satisfies

$$W(z) = \ln z + O(1) \quad (z \to \infty), \tag{7.52}$$

$$0 < zW'(z) < c \quad \text{for } z \to 0. \tag{7.53}$$

Proof The relation (7.51) is verified directly, while (7.52) and (7.53) are consequences of the estimates

$$W'(z) = \frac{e^{-z}}{\sqrt{u}} \int_0^z \frac{e^u}{\sqrt{u}}\, du = \frac{1}{z} + \frac{1}{2z^2} + O\left(\frac{1}{z^3}\right) \quad (z \to \infty),$$

$$W'(z) = 2 + O(z) \quad (z \to 0),$$

which are proved by integration by parts. □

Proof of Theorem 7.12 (1) Without loss of generality, we may assume that $x^* = 0$. Let us prove that, if $\varepsilon > 0$ is sufficiently small, then the process $X(t)$, beginning at an arbitrary time $t > 0$ at any point of the interval $(-\varepsilon, \varepsilon)$, reaches the ends of this interval in a finite time. It follows from Theorem 7.11 and the strong Markov property of the process $X(t)$ that this assertion is equivalent to Theorem 7.12.

Using Lemma 7.1, we can thus reduce the proof of Theorem 7.12 to the construction of a function $V(t, x)$ in $\{t > 0\} \times (-\varepsilon, \varepsilon)$ satisfying conditions (7.48), (7.49). Applying the usual methods of stability theory, we shall first construct a function $V(t, x)$ satisfying the conditions (7.48), (7.49) for

$$\widetilde{L} = \frac{\partial}{\partial t} + a(t)\beta x \frac{\partial}{\partial x} + \frac{1}{2} a^2(t)\sigma_0^2 \frac{\partial^2}{\partial x^2},$$

where $\beta = -R'(0)$, $\sigma_0^2 = \sigma^2(0)$. It will then be an easy matter to show that after some modification the function will satisfy these conditions for the "full" operator

$$L = \frac{\partial}{\partial t} - a(t)(R(x) - \alpha) \frac{\partial}{\partial x} + \frac{1}{2} a^2(t)\sigma^2(x) \frac{\partial^2}{\partial x^2}. \tag{7.54}$$

We shall look for the function $V(t, x)$ in the form

$$V(t, x) = \Phi(t) - W(z), \qquad z = \frac{x^2}{\varphi(t)},$$

where W is the function defined by (7.50) and the functions $\Phi(t)$ and $\varphi(t)$ will be specified below. Simple computations lead to the equality

$$\widetilde{L}V(t,x) = \Phi'(t) + zW(z)\left(\frac{\varphi'}{\varphi} - 2\beta a(t)\right) - \frac{a^2(t)\sigma_0^2}{\varphi(t)}[2zW'' + W'].$$

Since $2zW'' + W' = 2 - 2zW'$ by (7.51), it follows now that

$$\widetilde{L}V(t,x) = \Phi'(t) - \frac{2a^2(t)\sigma_0^2}{\varphi(t)} + zW'(z)\left[\frac{\varphi'}{\varphi} + \frac{2\sigma_0^2 a^2(t)}{\varphi} - 2\beta a(t)\right].$$

We now define $\varphi(t)$ by

$$\varphi(t) = 2\sigma^2 \exp\left\{2\beta\int_0^t a(s)\,ds\right\}\int_t^\infty a^2(s)\exp\left\{-2\beta\int_0^s a(u)\,du\right\}ds$$

(the convergence of the integral follows from (7.38)). Thus, we obtain the equality

$$LV(t,x) = \Phi'(t) - \frac{2a^2(t)\sigma_0^2}{\varphi(t)}.$$

Setting

$$\widetilde{\Phi}(t) = \ln f(t); \quad f(t) = \left[\int_t^\infty a^2(s)\exp\left\{-2\beta\int_0^s a(u)\,du\right\}ds\right]^{-1},$$

we see now that for this choice of $\varphi(t)$ and $\Phi(t)$ the function $V(t,x)$ satisfies the condition $\widetilde{L}V(t,x) = 0$. Since $W(z)$ is an increasing function for $z > 0$ (see (7.53)), it follows that $V(t,x) \geq V(t,\varepsilon)$ for $|x| \leq \varepsilon$ and therefore (7.49) will follow if we can show that $V(t,\varepsilon) \to \infty$ as $t \to \infty$. When $t \to \infty$ it follows from (7.52) that

$$V(t,\varepsilon) = \ln f(t) - \ln\left[\frac{\varepsilon^2}{2\sigma_0^2}\exp\left\{-2\beta\int_0^t a(s)\,ds\right\}f(t)\right] + O(1)$$

$$\geq \beta\int_0^t a(s)\,ds + O(1). \tag{7.55}$$

We have thus verified (7.48) and (7.49) for the function $V(x,t)$ and the operator $\widetilde{L}$.

(2) We now prove that if $\gamma < 2\beta$, then the function

$$V_1(t,x) = V(t,x) - \gamma\int_0^t a(s)\,ds = \Phi(t) - W\left(\frac{x^2}{\varphi(t)}\right) - \gamma\int_0^t a(s)\,ds$$

will satisfy conditions (7.48) and (7.49) in the domain $(t > 0) \times (-\varepsilon, \varepsilon)$ for the operator (7.54), provided $\varepsilon > 0$ is sufficiently small.

(7.49) follows for $V_1(t,x)$ from (7.55). We now apply the operator L to the function V_1. Since $\widetilde{L}V = 0$, it follows that

$$LV_1 = LV - \gamma a(t) = (L - \widetilde{L})V - \gamma a(t)$$

$$= -a(t)(R(x) - \alpha - R'(0)x)\frac{\partial V}{\partial x} + \frac{1}{2}a^2(t)(\sigma^2(x) - \sigma_0^2)\frac{\partial^2 V}{\partial x^2} - \gamma a(t)$$

$$\leq \delta\left(a(t)\left|x\frac{\partial V}{\partial x}\right| + a^2(t)\left|\frac{\partial^2 V}{\partial x^2}\right|\right) - \gamma a(t) \tag{7.56}$$

for $x \in (-\varepsilon, \varepsilon)$. The constant $\delta > 0$ in this inequality can be made arbitrarily small by a suitable choice of ε. It is clear from the form of the function V and from (7.51), (7.53) that

$$\left.\begin{aligned}\left|x\frac{\partial V}{\partial x}\right| &= |zW'(z)| < k_1, \\[2mm] \left|\frac{\partial^2 V}{\partial x^2}\right| &= \frac{4}{\varphi(t)}|1 - zW'(z)| < \frac{k_1}{\varphi(t)}\end{aligned}\right\} \tag{7.57}$$

for some constant k_1. Now, applying de l'Hôpital's rule and using the inequality $|a'(t)| < ka^2(t)$, we easily see that

$$a(t)f(t)\exp\left\{-2\beta\int_0^t a(s)\,ds\right\} = \frac{2\sigma_0^2 a(t)}{\varphi(t)}$$

is bounded. Hence, for a suitable constant k_2,

$$\frac{a^2(t)}{\varphi(t)} < k_2 a(t). \tag{7.58}$$

The relations (7.56), (7.57) and (7.58) now imply the required inequality $LV_1 \leq 0$. This we wished to show.   □

7.7 Some Generalizations

In this section we shall survey a few stability problems not yet discussed in this book. The discussion will inevitably be quite sketchy. In some cases we shall only make reference to the literature, not touching upon the problem proper.

7.7.1 Stability and Excessive Functions[6]

Let $X = (X(t), \mathbf{P}^x)$ be a time-homogeneous right-continuous strong Markov process in a Banach space E. Here $\mathbf{P}^x$ is the measure generated by the "initial condition" $X(0) = x$ (for more details, see [64]). We denote by $\|x\|$ the norm of an element x, and by $\mathfrak{B}$ the σ-algebra of measurable sets in E.

---

[6]See [37, 122].

An extremely useful tool for investigating the properties of Markov processes is provided by the excessive functions (see [64]). An *excessive function* is a function $V(x)$ with the properties

$$0 \le T_t V(x) = \int_E P(x, t, dy) V(y) \le V(x) \quad (t \ge 0, \ x \in E),$$

$$T_t V(x) \to V(x) \quad \text{as } t \downarrow 0.$$

It is known [64, Sect. 12.2] that an excessive function V satisfies the inequality

$$\mathbf{E}^x V(X(\zeta)) \le V(x) \tag{7.59}$$

for any Markov time ζ. The function V is said to be excessive for the process X in an open set U if inequality (7.59) is satisfied for all $\zeta \le \tau_U$, where τ_U is the first exit time of a path of the process from U.

Recall that a nonempty set $D \in \mathfrak{B}$ is said to be invariant for the process X if $P(x, t, D) = 1$ for $x \in D$, $t \ge 0$.

Definition 7.1 An invariant point $x_0 \in E$ for the process X is said to be *stable in probability* for the process X if

$$\inf_{\|y - x_0\| \to 0} \mathbf{P}^y \left\{ \sup_{t>0} \|X(t) - x_0\| > \varepsilon \right\} = 0.$$

Lemma 7.3 *A sufficient condition for a point x_0 to be stable in probability for the process X is that there exists a function V, which is excessive for the process X in a neighborhood of x_0, and which satisfies $V(x_0) = 0$ and $\inf_{\|x - x_0\| > \varepsilon} V(x) = V_\varepsilon > 0$ for $\varepsilon > 0$.*

Proof The proof follows from (7.59) and Chebyshev's inequality, since

$$V_\varepsilon \mathbf{P}^x \left\{ \sup_{t>0} \|X(t) - x_0\| > \varepsilon \right\} \le \mathbf{E}^x V(X(\tau_{U_\varepsilon(x_0)}(t))) \le V(x). \qquad \square$$

It is evident that the above stability condition is too general to be of much interest. In certain special cases, however, one can derive more specific conditions. For example, the proof of Theorem 5.3 essentially reduces to verifying that a non-negative function V which is defined in some neighborhood U of the origin is of class $\mathbf{C}_2^0(U)$ and satisfies the inequality $LV \le 0$, is excessive in U.

For the right-continuous strong Markov processes, the fact that non-negative functions for which the weak generator is non-positive, are excessive, can be established by means of a well-known theorem of [64, Sect. 5.5.1]. Together with Lemma 7.3, this enables one to derive stability conditions for such processes (see [166, 167]). A generalization of Theorem 5.3 to jump processes has been considered by Gikhman and Dorogovtsev [91]. See also the recent article [140] and references therein.

7.7.2 Stability of the Invariant Set

Let $\rho(x, U) = \inf_{x_0 \in U} \|x - x_0\|$ denote the distance of a point x from the set U. An invariant set U of the process X is said to be stable in probability if

$$\lim_{\rho(y,U) \to 0} \mathbf{P}^y \left\{ \sup_{t>0} \rho(X(t), U) > \varepsilon \right\} = 0.$$

The proof of the following lemma is similar to that of Lemma 7.1.

Lemma 7.4 *An invariant set D is stable in probability for the process X if there exists a function $V(x)$, excessive for X in a neighborhood of the set D, such that*:

1. $V(x) = 0$ *for $x \in D$*;
2. $\inf_{\rho(x,D)>\epsilon} V(x) = V_\varepsilon > 0$ *for $\varepsilon > 0$*.

The following theorem is the analog of Theorem 5.3 for stability of the invariant set. We prove it by using Lemma 7.4 and proceeding as in the proof of Theorem 5.3.

Theorem 7.13 *Let $X(t)$ be the process described by the generator*

$$L = \frac{\partial}{\partial t} + \left(b(t, x), \frac{\partial}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t, x), \frac{\partial}{\partial x} \right)^2. \tag{7.60}$$

Assume that there exists a function $V(t, x)$, twice continuously differentiable with respect to x and continuously differentiable with respect to t, vanishing for $x \in D$ and satisfying the conditions

$$LV \le 0; \qquad \inf_{\rho(x,D)>\varepsilon;\ t>0} V(t, x) = V_\varepsilon > 0 \quad \text{for } \varepsilon > 0$$

in a neighborhood of the set D.

Then the invariant set D of the process X is stable in probability.

Remark 7.6 If the set D is inaccessible (see Sect. 5.2) to the process, the smoothness conditions imposed on V can be weakened. To be precise, the function V need not be smooth at the points x of the set $\Gamma = \{x : \rho(x, D) = 0\}$. Moreover, using Lemma 7.4 one can establish analogous stability conditions for the invariant set of a jump process.

7.7.3 Equations Whose Coefficients Are Markov Processes

Several authors ([111], [81] and others) have considered the properties of systems described by equations of type

$$\frac{dY}{dt} = F(Y, t, X(t)), \tag{7.61}$$

where Y, F are vectors in $\mathbb{R}^m$ and $X(t)$ is a Markov process with values in $\mathbb{R}^l$. If the process $X(t)$ is governed by the generator (7.60), one can investigate the stability of the system (7.61) with the use of Theorem 7.13. In fact, it is clear that the pair $(X(t), Y(t))$ is also a Markov process, whose generator is defined on sufficiently smooth functions by

$$L_1 V = \frac{\partial V}{\partial t} + \left(F(y, t, x), \frac{\partial V}{\partial y} \right) + \left(b(t, x), \frac{\partial V}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t, x), \frac{\partial}{\partial x} \right)^2 V.$$

Thus, investigation of the stability of the path $Y(t) \equiv 0$ for the process (7.61) (on the assumption that $F(0, t, x) \equiv 0$) reduces to investigation of the stability of the m-dimensional hyperplane $y = 0$ for the $(l + m)$-dimensional Markov process $(X(t), Y(t))$. Hence, using Theorem 7.13, we get

Theorem 7.14 *Suppose that for some $\varepsilon_0 > 0$ and all $t > 0$, $|y| < \varepsilon_0$, there exists a function $V(t, x, y)$, continuously differentiable with respect to $t, y \in \mathbb{R}^m$, twice continuously differentiable with respect to $x \in \mathbb{R}^l$ everywhere except perhaps for the set $\{y = 0\}$, and such that*

$$L_1 V \leq 0; \qquad V(t, x, 0) = 0; \qquad \inf_{t > 0, \, |y| > \varepsilon} V(t, x, y) = V_\varepsilon > 0$$

for $0 < \varepsilon < \varepsilon_0$.

Then the solution $Y \equiv 0$ of the system (7.61) is stable in probability, in the sense that

$$\lim_{|y| \to 0} \mathbf{P} \left\{ \sup_{t > 0} |Y(t)| > \varepsilon \;\middle|\; X(0) = x; \; Y(0) = y \right\} = 0.$$

Many of the problems considered hitherto in simpler situations are of interest for equations of type (7.61). For example, we might study stability in probability of a linear system

$$\frac{dY}{dt} = F(X(t))Y. \tag{7.62}$$

This problem has an extremely simple solution if $X(t)$ is a time-homogeneous ergodic process and $m = 1$. Then

$$Y(t) = Y(0) \exp \left\{ \int_0^t F(X(s)) \, ds \right\}. \tag{7.63}$$

Hence, by the strong law of large numbers (Sect. 4.5), it follows at once that if

$$\overline{F} = \int F(x) \mu(dx) < 0$$

holds, where μ is the stationary distribution of the process $X(t)$, then the process $Y(t)$ is asymptotically stable. The same arguments show that the system (7.62) is unstable if $\overline{F} > 0$.

Conditions for p-stability of systems of type (7.62) are quite complicated, even in the case $m = 1$. In fact, it follows from (7.63) that

$$\mathbf{E}\{|Y(t)|^p | X(0) = x; \ Y(0) = y\} = |y|^p \mathbf{E}^x \exp\left\{p \int_0^t F(X(s))\,ds\right\}.$$

If $X(t)$ is, say a time-homogeneous diffusion process in $\mathbb{R}^l$ with local characteristics $b(x)$ and $\sigma_1(x), \ldots, \sigma_k(x)$, then it follows from the Feynman–Kac formula, that the function

$$u(t, x) = \mathbf{E}^x \exp\left\{p \int_0^t F(X(s))\,ds\right\}$$

is a solution of the equation

$$\frac{\partial u}{\partial t} = \left(b(x), \frac{\partial u}{\partial x}\right) + \frac{1}{2}\left(\sigma(x), \frac{\partial}{\partial x}\right)^2 u + pF(x)u, \qquad (7.64)$$

which satisfies the initial condition

$$u(0, x) = 1. \qquad (7.65)$$

Thus the problem of p-stability of the system (7.62) for $m = 1$ is reduced to investigation of the limiting behavior of the solution of problem (7.64), (7.65) as $t \to \infty$.

If $m > 1$, complications arise. Frisch [81] introduces a linear system of partial differential equations solving this system, one can use quadratures to determine the moments of a process $Y(t)$ which satisfies (7.62) for arbitrary m.

The case of a time-homogeneous Markov process $X(t)$ with finitely many states has been studied by Kac and Krasovskii [111]. In particular, their paper presents algebraic criteria for the stability of systems in the mean square.

7.7.4 Stability Under Persistent Perturbation by White Noise

We have already studied the problem of stability under persistent random perturbations (Sect. 1.6). Our attention was then centered upon random perturbations with finite expectation. It is to be expected that if we narrow down the class of admissible perturbations, we shall be able to derive broader stability criteria. The author has considered in [123] the stability of deterministic systems under perturbation by white noise. We now briefly present some results of that paper.

We consider the equation

$$\frac{dx}{dt} = F(t, x) \qquad (F(t, 0) \equiv 0) \qquad (7.66)$$

in $\mathbb{R}^l$.

We call the solution $x = 0$ of (7.66) *stable under persistent perturbation by white noise* if the solutions $X(t)$ of the Itô equation

$$dX(t) = F(t, X)\, dt + \sum_{r=1}^{k} \sigma_r(t, X)\, d\xi_r(t) + b(t, X)\, dt \qquad (7.67)$$

have the following property: For any $\varepsilon > 0$, there exists a $\gamma > 0$ such that for all processes $X(t)$ satisfying (7.67), with initial condition $|X(0)| < \gamma$ and coefficients σ_r and b such that

$$\sum_{r=1}^{k} |\sigma_r(t, x)| + |b(t, x)| < \gamma, \qquad (7.68)$$

the inequality

$$\mathbf{P}\{|X(t)| > \varepsilon\} < \varepsilon$$

holds for all $t > 0$.

Theorem 7.15 [123] *The solution $x = 0$ is stable under persistent perturbation by white noise if there exists a positive definite function $V(t, x) \in \mathbf{C}_2$ in the domain $(t > 0) \times \mathbb{R}^l$ which has an infinitesimal upper limit and satisfies the conditions*:

1. $\inf_{t>0} V(t, x) \to \infty$ *as* $|x| \to \infty$.
2. *For every $\varepsilon > 0$, there exist positive constants $\alpha_1(\varepsilon)$, $\alpha_2(\varepsilon)$ and $\gamma(\varepsilon)$ such that for all $\sigma_r(t, x)$ and $b(t, x)$ satisfying condition (7.68) with $\gamma = \gamma(\varepsilon)$ the inequality*

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t, x), \frac{\partial}{\partial x} \right)^2 V + \left(F(t, x) + b(t, x), \frac{\partial V}{\partial x} \right)$$

$$\leq -\alpha_1(\varepsilon) - \alpha_2(\varepsilon) \frac{|\partial V/\partial x|^2}{V} \qquad (7.69)$$

holds in the domain $|x| > \varepsilon$.

One can easily infer from Theorem 7.15 simpler stability conditions for persistent perturbations. For example, in [123] the theorem is applied to investigate the effect of random perturbation by white noise on the performance of absolutely stable controlled systems with a single final control element. It is shown that the system is stable under such perturbations if the response curve of the final control element has a bounded derivative.

It is also readily inferred from Theorem 7.15 that an exponentially p-stable stochastic system is stable under persistent perturbation by white noise (see [221]).

7.7.5 Boundedness in Probability of the Output Process of a Nonlinear Stochastic System

Suppose that the solution $X(t) \equiv 0$ of the system of Itô equations

$$dX(t) = b(t, X(t)) \, dt + \sum_{r=1}^{k} \sigma_r(t, X(t)) \, d\xi_r(t) \qquad (7.70)$$

in $\mathbb{R}^l$ is exponentially p-stable for $p = 1$. Assume that the system (7.70) is "driven" by a continuous stochastic process $\zeta(t)$, whose absolute value has bounded expectation, and which is independent of the Wiener processes $\xi_1(t), \ldots, \xi_k(t)$. The "output process" $Y(t)$ of the resulting system is described by the equation

$$dY(t) = b(t, Y(t)) \, dt + \sum_{r=1}^{k} \sigma_r(t, Y(t)) \, d\xi_r(t) + \zeta(t) \, dt. \qquad (7.71)$$

It is shown in [221] that the expectation of the process $|Y(t)|$ is bounded uniformly in t. It follows that the "output process" $Y(t)$ is bounded in probability.

The case of a system described by a linear n-th order equation driven by white noise has been considered in detail by Dym [61]. See also [218].

Chapter 8
Stabilization of Controlled Stochastic Systems

8.1 Preliminary Remarks

As mentioned in the preface, the stability theory of SDEs was developed mainly to meet the needs of stabilization of moving systems subjected to random perturbations. In this chapter we shall consider some problems concerning the stabilization of controlled stochastic systems. The results achieved to date in this field are rather sparse, despite the fact that the basic formulations of the problems and the fundamental equations have been known for some time ([112], [159–161], [279] etc.). The only results of any significance are those pertaining to linear systems and employing quadratic control criteria. We devote to them the exposition which now follows, based on the material of Chaps. 5 through 7.

In conformity with the assumptions adopted hitherto, we shall consider controlled systems of the type

$$dX(t) = b(t, X, u) \, dt + \sum_{r=1}^{k} \sigma_r(t, X, u) \, d\xi_r(t). \tag{8.1}$$

Here $b(t, x, u)$ and $\sigma_r(t, x, u)$ are vector-valued functions, jointly continuous in all arguments and such that $b(t, 0, 0) = \sigma_r(t, 0, 0) = 0$, u is a scalar control parameter and $\xi_r(t)$, $r = 1, \ldots, k$, are independent Wiener processes such that $\mathbf{E}\xi_r(t) = 0$, $\mathbf{E}\xi_r^2(t) = t$. We shall assume that the control u in the system (8.1) is a function of t and $X(t)$, $u = u(t, X(t))$. Then the process described by (8.1) is Markovian.[1] A function $u = u(t, x)$ is said to be admissible control if the coefficients $b(t, x, u(t, x))$, $\sigma_r(t, x, u(t, x))$ are continuous, have continuous derivatives with respect to x which are bounded uniformly in $t > 0$, and moreover $u(t, 0) = 0$.

This chapter was written jointly with M.B. Nevelson.

[1] In the literature this type of control is known as Markov control, or control employing the feedback principle.

The class of admissible controls is denoted by $\mathcal{U}$. Each function (admissible control) $u \in \mathcal{U}$ can be associated with a Markov process $X_u^{s,x}(t)$, which is the solution of (8.1) with initial condition $X_u^{s,x}(s) = x$.

By analogy with the deterministic case (see [158]), we shall consider the following two stabilization problems:

I. Asymptotic (exponential) p-stabilization: To determine an admissible control $u = u_0(t, x)$ such that the system (8.1) with $u = u_0(t, x)$ is asymptotically (exponentially) p-stable (see Sect. 5.7).
II. Optimal stabilization to minimize a given cost: To determine a control $u = u_0(t, x)$ for which the functional

$$J^{s,x_0}(u) = \int_s^\infty \mathbf{E} K(t, X_u^{s,x_0}(t), u(t, X_u^{s,x_0}(t))) \, dt$$

(i.e. the cost) attains a minimum. Here (s, x_0) is a fixed initial point, and $K(t, x, u) \geq 0$ for $t \geq 0$, $x \in \mathbb{R}^l$, $u \in (-\infty, \infty)$.

Remark 8.1 Under the assumptions of the theorem proved in the next section, the function $u_0(t, x)$ solving the optimal stabilization problem turns out to be the same for all initial points (s, x_0).

Remark 8.2 Problem II is in general not a stabilization problem. For example, if $K(t, x, u) = 0$ for $|x| > R$, the optimal strategy may be to force the path of the process $X_u^{s,x}(t)$ out of the R-neighborhood of $x = 0$. In the sequel, however, we shall confine ourselves to costs satisfying the condition that for any $u \in (-\infty, \infty)$ and certain constants $p > 0$, $c > 0$,

$$K(t, x, u) > c|x|^p \tag{8.2}$$

holds. Under this restriction Problems I and II prove to be intimately connected. In fact, suppose that the control $u_0(t, x)$ is a solution of Problem II for a function $K(t, x, u)$ satisfying condition (8.2). It then follows from Lemma 5.7 that

$$\lim_{t \to \infty} \mathbf{E}|X_{u_0}^{s,x}(t)|^p = 0 \quad \text{as } t \to \infty. \tag{8.3}$$

Under certain additional assumptions, this implies that the system (8.1) is asymptotically and exponentially stable.

8.2 Bellman's Principle

In this section we wish to prove a theorem which is a modification of Bellman's principle regarding problems of optimal stabilization of stochastic systems (see [155], [159–161], [279], [75]).

Let V be a function of class $\mathbf{C}_2^{(0)}(E)$. Substituting $u = u(t, x)$ in

$$L_u = \frac{\partial}{\partial t} + \left(b(t, x, u), \frac{\partial}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t, x, u), \frac{\partial}{\partial x} \right)^2,$$

we get the generator of a certain Markov process $X_u^{s,x}(t)$.

Theorem 8.1 *Suppose that there exist functions $V_0(t, x) \in \mathbf{C}_2^{(0)}(E)$, $u_0(t, x) \in \mathcal{U}$, satisfying for all $t \geq 0$, $x \in \mathbb{R}^l$, $u \in (-\infty, \infty)$ and certain positive constants p, n, k_1, k_2 the conditions*

$$V_0(t, x) \leq k_1 |x|^p, \qquad \left| \frac{\partial V}{\partial x_i} \right| \leq k_1 (|x|^n + 1), \tag{8.4}$$

$$L_{u_0} V_0(t, x) + K(t, x, u_0(t, x)) \equiv 0, \tag{8.5}$$

$$L_u V_0(t, x) + K(t, x, u) \geq 0, \tag{8.6}$$

$$K(t, x, u) \geq k_2 |x|^p. \tag{8.7}$$

Then the function $u_0(t, x)$ is a solution of the optimal stabilization problem for the system (8.1) in the sense of minimizing the cost $J^{s,x_0}(u)$ and moreover

$$J^{s,x_0}(u_0) = \min_{u \in \mathcal{U}} J^{s,x_0}(u) = V_0(s, x_0). \tag{8.8}$$

Further, the control $u_0(t, x)$ makes the system (8.1) exponentially p-stable.

Proof Let $u = u(t, x)$ be any admissible control. Applying Itô's formula (3.30) to the function $V(t, X^{s,x_0}(t))$ and noting that by the second of the inequalities (8.4) and Lemma 5.2 all the stochastic integrals appearing in Itô's formula have zero expectation, we get

$$\mathbf{E} V_0(t, X_u^{s,x_0}(t)) - V_0(s, x) = \mathbf{E} \int_s^t L_u V_0(v, X_u^{s,x_0}(v)) \, dv. \tag{8.9}$$

Setting $u = u_0(t, x)$ in this equality and applying (8.5), we obtain

$$\mathbf{E} \int_s^t K(v, X_{u_0}^{s,x_0}(v), u_0(v, X_{u_0}^{s,x_0}(v))) \, dv = V_0(s, x) - \mathbf{E} V_0(t, X_{u_0}^{s,x_0}(t)). \tag{8.10}$$

Letting $t \to \infty$, we get $J^{s,x_0}(u_0) < \infty$. Using this inequality, (8.7) and Remark 8.2, we see that the process $X_{u_0}^{s,x_0}(t)$ satisfies (8.3). From (8.3) and (8.4) we get

$$\mathbf{E} V_0(t, X_{u_0}^{s,x_0}(t)) \leq k_1 \mathbf{E} |X_{u_0}^{s,x_0}(t)|^p \to 0 \quad \text{as } t \to \infty.$$

Hence, letting $t \to \infty$ in (8.10), we have

$$J^{s,x_0}(u_0) = V_0(s, x_0).$$

Now, if $u(t, x)$ is any admissible control such that $J^{s,x_0}(u) < \infty$, then as before, we readily see that

$$\lim_{t\to\infty} \mathbf{E} V_0(t, X_u^{s,x_0}(t)) = 0.$$

Using this equality and the relation

$$\mathbf{E} V_0(t, X_u^{s,x_0}(t)) \geq V_0(s, x_0) - \mathbf{E} \int_s^t K(v, X_u^{s,x_0}(v), u(v, X^{s,x_0}(v))) \, dv,$$

which follows from (8.9), (8.6), we finally see by letting $t \to \infty$ that

$$\min_{u\in\mathcal{U}} J^{s,x_0}(u) \geq V_0(s, x_0).$$

It remains to prove that the system (8.1) is exponentially p-stable when $u = u_0(t, x)$. By Theorem 5.11, it will suffice to prove that

$$V_0(t, x) \geq k_3 |x|^p$$

for some constant $k_3 > 0$. From (8.7) and (8.8) we get

$$V_0(s, x) = J^{s,x_0}(u_0) \geq k_2 \int_s^\infty \mathbf{E} |X_{u_0}^{s,x_0}(v)|^p \, dv.$$

Thus there exists for any $x, s \geq 0$ a $T = T(s, x)$ such that

$$\mathbf{E} |X_{u_0}^{s,x}(T)|^p < \frac{1}{2} |x|^p.$$

In view of the inequality $L_{u_0}(|x|^p) \geq -k_4 |x|^p$ (see (5.55)), the above and Itô's formula imply that

$$V_0(s, x) \geq k_2 \int_s^\infty \mathbf{E} |X_{u_0}^{s,x}(v)|^p \, dv \geq -k_5 \int_s^T \mathbf{E} L_{u_0}(|X_{u_0}^{s,x}(v)|^p) \, dv$$

$$= k_5(|x|^p - \mathbf{E} |X_{u_0}^{s,x}(T)|^p) \geq \frac{k_5}{2} |x|^p = k_3 |x|^p.$$

This completes the proof. $\square$

For subsequent work, it is useful to combine conditions (8.5) and (8.6) into one equation

$$\min_{u\in(-\infty,\infty)} [L_u V_0(s, x) + K(s, x, u)] = 0 \tag{8.11}$$

(Bellman's equation).

Remark 8.3 Condition (8.7), which imposes a restriction on the function $K(t, s, u)$ for all u, seems at first sight rather stringent. One might expect that (8.7) could

be replaced by the weaker condition $K(t, x, u_0(t, x)) > k_2|x|^p$. However, the following example will show that if condition (8.7) is thus weakened the assertion of Theorem 8.1 is no longer true.

Consider the optimal stabilization problem for the deterministic system

$$\frac{dx_1}{dt} = -x_1, \qquad \frac{dx_2}{dt} = x_2 + u$$

with $K(t, x_1, x_2, u) = x_1^2 + u^2$. In this case

$$L_u = -x_1 \frac{\partial}{\partial x_1} + (x_2 + u) \frac{\partial}{\partial x_2}.$$

Conditions (8.4), (8.5) and (8.6) are easily seen to hold for the function

$$V_0 = \frac{x_1^2}{2} + 2x_2^2, \qquad u_0 = -2x_2.$$

It is also clear that condition (8.7) holds, say, with $u = c_1 x_1 + c_2 x_2$, when $c_2 \neq 0$. Nevertheless, the control u_0 is not optimal in our sense, since the optimal control is evidently $u = 0$.

Remark 8.4 We have proved that the control u_0 is optimal for all controls of class $\mathcal{U}$. One might expect a "higher-quality" control to exist in the class of controls which allow for the past history of the process $X(t)$ from the initial time s to the present t. However, it can be proved that u_0 is also optimal in this broader class of admissible controls (see [222]).

Remark 8.5 We may consider simultaneously with the Problem II of optimal stabilization also the problem of minimizing the functional

$$T_T^{s, x_0}(u) = \int_s^T \mathbf{E} K(t, X_u(t), u(t, X_u^{s, x_0}(t))) \, dt, \qquad (8.12)$$

where T is a constant larger than s. By repeating almost word for word the considerations in the proof of Theorem 8.1, we may easily show that if there exist functions $V_0^T(t, x) \in \mathbf{C}_2^{(0)}(E)$, $u_{0T}(t, x) \in \mathcal{U}$ which satisfy for all $T \geq t \geq s$, $x \in \mathbb{R}^l$, $u \in (-\infty, \infty)$ and certain positive constants p, n_1, k_1, k_2 the conditions (8.4), (8.6) and (8.7) and the identity

$$V_0^T(T, x) \equiv 0, \qquad (8.13)$$

then

$$\min_{u \in \mathcal{U}} T_T^{s, x}(u) = T_T^{s, x}(u_0) = V_0^T(s, x).$$

8.3 Linear Systems

Let us apply Theorem 8.1 to the investigation of the system

$$dX = \left[B(t)\,dt + \sum_{r=1}^{k} \sigma_r(t)\,d\xi_r(t) \right] X + \left[h(t)\,dt + \sum_{r=1}^{k} \varphi_r(t)\,d\xi_r(t) \right] u, \quad (8.14)$$

which is linear in x and u. Here $B(t)$ and $\sigma_r(t)$ are $l \times l$ matrices, $h(t)$ and $\varphi_r(t)$ vectors in $\mathbb{R}^l$. The elements of the matrices $B(t)$, $\sigma_r(t)$ and the components of the vectors $h(t)$, $\varphi_r(t)$ are assumed to be continuous and bounded functions of time.

We consider the optimal stabilization problem for the system (8.14), with the kernel of the cost defined by

$$K(t, x, u) = (\alpha(t)x, x) + \lambda u^2. \quad (8.15)$$

Here $\alpha(t)$ is a bounded symmetric $l \times l$ matrix, which is positive definite uniformly in $t \geq s$; $\lambda > 0$.

We wish to find an optimal Lyapunov function $V_0(t, x)$ satisfying the assumptions of Theorem 8.1 and given by a nonnegative quadratic form

$$V_0(t, x) = (C(t)x, x), \quad (8.16)$$

where $C(t)$ is a symmetric $l \times l$ matrix.

Clearly

$$L_u = \frac{\partial}{\partial t} + \left(B(t)x + h(t)u, \frac{\partial}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t)x + \varphi_r(t)u, \frac{\partial}{\partial x} \right)^2$$

is the generator for the system (8.14). Equation (8.11), relating the optimal Lyapunov function $V_0(t, x)$ and the optimal control $u_0(t, x)$, has the form

$$\frac{\partial V_0}{\partial t} + \left(B(t)x, \frac{\partial}{\partial x} \right) V_0 + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(t)x, \frac{\partial}{\partial x} \right)^2 V_0 + (\alpha(t)x, x)$$

$$= -\min_{u \in U} \left\{ u \left[\left(h(t), \frac{\partial}{\partial x} \right) V_0 + \sum_{r=1}^{k} \left(\sigma_r(t)x, \frac{\partial}{\partial x} \right) \left(\varphi_r(t), \frac{\partial}{\partial x} \right) V_0 \right] \right.$$

$$\left. + u^2 \left[\frac{1}{2} \sum_{r=1}^{k} \left(\varphi_r(t), \frac{\partial}{\partial x} \right)^2 V_0 + \lambda \right] \right\}$$

$$= -u_0 \left[\left(h(t), \frac{\partial}{\partial x} \right) V_0 + \sum_{r=1}^{k} \left(\sigma_r(t)x, \frac{\partial}{\partial x} \right) \left(\varphi_r(t), \frac{\partial}{\partial x} \right) V_0 \right]$$

$$- u_0^2 \left[\frac{1}{2} \sum_{r=1}^{k} \left(\varphi_r(t), \frac{\partial}{\partial x} \right)^2 V_0 + \lambda \right]. \quad (8.17)$$

The function $u_0(t, x)$ in (8.17) obviously has the form

$$u_0(t, x) = -\frac{(h(t), \frac{\partial}{\partial x})V_0 + \sum_{r=1}^{k}(\sigma_r(t)x, \frac{\partial}{\partial x})(\varphi_r(t), \frac{\partial}{\partial x})V_0}{2\lambda + \sum_{r=1}^{k}(\varphi_r(t), \frac{\partial}{\partial x})^2 V_0}. \tag{8.18}$$

Substituting (8.16) into (8.18), we get

$$u_0(t, x) = -\frac{(h(t), C(t)x) + \sum_{r=1}^{k}(C(t)\varphi_r(t), \sigma_r(t)x)}{\lambda + \sum_{r=1}^{k}(C(t)\varphi_r(t), \varphi_r(t))}. \tag{8.19}$$

This implies that when the optimal Lyapunov function is defined by (8.16), the optimal control is linear in x.

Using (8.17) and (8.19), we get the following equation for $V_0(t, x)$;

$$\frac{\partial V_0}{\partial t} + \left(B(t)x, \frac{\partial}{\partial x}\right)V_0 + \frac{1}{2}\sum_{r=1}^{k}\left(\sigma_r(t)x, \frac{\partial}{\partial x}\right)^2 V_0 + (\alpha(t)x, x)$$

$$= \frac{[(h(t), \frac{\partial}{\partial x})V_0 + \sum_{r=1}^{k}(\sigma_r(t)x, \frac{\partial}{\partial x})(\varphi_r(t), \frac{\partial}{\partial x})V_0]^2}{4[\lambda + \frac{1}{2}\sum_{r=1}^{k}(\varphi_r(t), \frac{\partial}{\partial x})^2 V_0]}.$$

Since the matrix $C(t)$ is symmetric, this equation is equivalent to

$$\frac{dC}{dt} + CB(t) + B^*(t)C + \sum_{r=1}^{k}\sigma_r^*(t)C\sigma_r(t) + \alpha(t)$$

$$= \frac{(Ch(t) + \sum_{r=1}^{k}\sigma_r^*(t)C\varphi_r(t))(h^*(t)C + \sum_{r=1}^{k}\varphi_r^*(t)C\sigma_r(t))}{\lambda + \sum_{r=1}^{k}(C\varphi_r(t), \varphi_r(t))}. \tag{8.20}$$

From Theorem 8.1 we now infer

Lemma 8.1 *If* (8.20) *has a solution* $C(t)$ *which is bounded and positive definite for all* $t \geq s$, *then the function* $u_0(t, x)$, *defined by* (8.19), *minimizes the functional*

$$J^{s,x}(u) = \int_{s}^{\infty} \mathbf{E}[(\alpha(t)X_u^{s,x}(t), X_u^{s,x}(t)) + \lambda u^2(t, X_u^{s,x}(t))]\,dt.$$

In conclusion we note that the above remains valid in the deterministic case $\sigma_r(t) \equiv 0$, $\varphi_r(t) \equiv 0$. In particular, (8.20) is then simply a matrix Riccati equation:

$$\frac{dC}{dt} + CB(t) + B^*(t)C - \frac{Ch(t)h^*(t)C}{\lambda} + \alpha = 0. \tag{8.21}$$

8.4 Method of Successive Approximations

In the last section we set up a nonlinear differential equation for the matrix $C(t)$ of the optimal Lyapunov function $V_0(t, x)$ associated with a linear control system

$$dX = (B(t)X + h(t)u)\,dt + \sum_{r=1}^{k}(\sigma_r(t)X + \varphi_r(t)u)\,d\xi_r(t) \qquad (8.22)$$

possessing the cost function

$$K(t, x, u) = (\alpha(t)x, x) + \lambda u^2.$$

This equation has a fairly complicated form. Even in the deterministic case ($\sigma_r \equiv 0$, $\varphi_r \equiv 0$), when it becomes the well-known Riccati equation, it is not easy to investigate. We would therefore like to have easily checkable conditions for the existence of a bounded positive definite solution to the above equation, i.e., existence conditions for an optimal linear control. Such conditions can be given in terms of Lyapunov functions. While deriving them, we shall also describe a convenient method for practical computation of the optimal control: The method of successive approximations whose basic idea is due to Bellman [24]. This method has been applied to other problems of optimal control by Wonham [279, 280], Fleming [75] and others.

Theorem 8.2 *Suppose that there exists an admissible control which stabilizes the system* (8.22) *so that exponential stability in the mean square is obtained. Let further* $\lambda > 0$ *be arbitrary and let* $\alpha(t)$ *be a positive definite uniformly with respect to* $t \geq 0$ *symmetric matrix with continuous bounded coefficients. Then there exists a linear control* $u_0(t, x) \in \mathcal{U}$ *which is optimal in the sense of the quality criterion* $\mathcal{T}^{s,x}(u)$. *Moreover*

$$\mathcal{T}^{s,x}(u_0) = \min_{u \in \mathcal{U}} \mathcal{T}^{s,x}(u) = V_0(s, x) = (C_0(s)x, x), \qquad (8.23)$$

where $C_0(s)$ *is the unique bounded positive definite solution of* (8.20).

Proof Let us consider the minimization problem for the functional (8.12), where $K(t, x, u)$ is given by (8.15) and T is a constant. Remark 8.5 implies that to solve this problem it is enough to find functions $V_0^T(t, x)$, $u_0^T(t, x)$ satisfying the conditions (8.4), (8.6), (8.7) and (8.13).

Let $u_{1T}(t, x)$ be an arbitrary control that is admissible and linear in x, for instance $u_{1T}(t, x) \equiv 0$. We define the function $V_1^T(s, x)$ by the formula

$$V_1^T(s, x) = \int_s^T \mathbf{E}K(v, X_{u_{1T}}^{s,x}(v), u_{1T}(v, X_{u_{1T}}^{s,x}(v)))\,dv.$$

Lemma 5.6 implies that this function is twice continuously differentiable with respect to x and once with respect to s. It is evident from the Markov property that

$$V_1^T(s, x) = \int_s^{s+\Delta} \mathbf{E}K(v, X_{u_{1T}}^{s,x}, u_{1T}(v, X_{u_{1T}}^{s,x}(v)))\,dv + \mathbf{E}V_1^T(s + \Delta, X_{u_{1T}}^{s,x}(s + \Delta)).$$

This and Itô's formula imply

$$L_{u_{1T}} V_1^T(s, x) + (\alpha(s)x, x) + \lambda u_{1T}^2(s, x) = 0,$$
$$V_1^T(T, x) \equiv 0. \tag{8.24}$$

On the other hand the linearity of $u_{1T}(t, x)$ implies that the process $X_{u_{1T}}^{s,x}(t)$ is described by a system of linear stochastic equations. Therefore (see Chap. 6) $V_1^T(s, x)$ is a quadratic form in x

$$V_1^T(s, x) = (C_1^T(s)x, x).$$

Let us define now the function $u_{2T}(s, x)$, i.e. the second approximation to the optimal control, by the equation

$$\min_{u \in (-\infty, \infty)} [L_u V_1^T(s, x) + (\alpha(s)x, x) + \lambda u^2] = L_{u_{2T}} V_1^T(s, x) + (\alpha(s)x, x) + \lambda u_{2T}^2. \tag{8.25}$$

Equations (8.24) and (8.25) imply

$$L_{u_{2T}} V_1^T(s, x) + (\alpha(s)x, x) + \lambda u_{2T}^2 \le 0. \tag{8.26}$$

Moreover we get from (8.25) that

$$u_{2T}(s, x) = -\frac{(h(s), \frac{\partial V_1^T(s,x)}{\partial x}) + \sum_{r=1}^k (\sigma_r(s)x, \frac{\partial}{\partial x})(\varphi_r(s), \frac{\partial}{\partial x})V_1^T(s, x)}{2\lambda + \sum_{r=1}^k (\varphi_r(s), \frac{\partial}{\partial x})^2 V_1^T(s, x)}$$

$$= -\frac{(h(s), C_1^T(s)x) + \sum_{r=1}^k (\sigma_r(s)x, C_1^T(s)\varphi_r(s))}{\lambda + \sum_{r=1}^k (C_1^T(s)\varphi_r(s), \varphi_r(s))},$$

which means that $u_{2T}(s, x)$ is a linear function. Suppose now that the function $V_2^T(s, x) = (C_2^T(s)x, x)$ is given by

$$V_2^T(s, x) = \int_s^T \mathbf{E} K(v, X_{u_{2T}}^{s,x}(v), u_{2T}(v, X_{u_{2T}}^{s,x}(v))) \, dv.$$

In a similar way as (8.24), we obtain

$$L_{u_{2T}} V_2^T(s, x) + (\alpha(s)x, x) + \lambda u_{2T}^2 = 0,$$
$$V_2^T(T, x) = 0. \tag{8.27}$$

This and (8.26) imply that the difference $U^T(s, x) = V_1^T(s, x) - V_2^T(s, x)$ satisfies

$$L_{u_{2T}} U^T(s, x) \le 0, \quad U^T(T, x) = 0.$$

These estimates and the equality

$$\mathbf{E} U^T(T, X_{u_{2T}}^{s,x}(T)) - U^T(s, x) = \int_s^T \mathbf{E} L_{u_{2T}} U^T(v, X_{u_{2T}}^{s,x}(v)) \, dv,$$

which is a consequence of Itô's formula, imply easily that

$$-U^T(s, x) \leq \mathbf{E}U^T(T, X_{u_{2T}}^{s,x}(T)) - U^T(s, x) \leq 0.$$

Therefore $U^T(s, x) \geq 0$ and we have shown that

$$V_1^T(s, x) \geq V_2^T(s, x). \tag{8.28}$$

Proceeding further in the same fashion we find the functions $u_{3T}(s, x), u_{4T}(s, x),$ … from the identities

$$\min_{u \in (-\infty, \infty)} [L_u V_{n-1}^T(s, x) + (\alpha(s)x, x) + \lambda u^2]$$

$$= L_{u_{nT}} V_{n-1}^T(s, x) + (\alpha(s)x, x) + \lambda u_{nT}^2, \tag{8.29}$$

where

$$V_n^T(s, x) = \int_s^T \mathbf{E}K(v, X_{u_{nT}}^{s,x}(v), u_{nT}(v, X_{u_{nT}}^{s,x}(v))) \, dv.$$

Evidently the equalities

$$u_{nT}(s, x) = -\frac{(h(s), \frac{\partial V_{n-1}^T(s,x)}{\partial x}) + \sum_{r=1}^k (\sigma_r(s)x, \frac{\partial}{\partial x})(\varphi_r(s), \frac{\partial}{\partial x}) V_{n-1}^T(s, x)}{2\lambda + \sum_{r=1}^k (\varphi_r(s), \frac{\partial}{\partial x})^2 V_{n-1}^T(s, x)}, \tag{8.30}$$

$$L_{u_{nT}} V_n^T(s, x) + (\alpha(s)x, x) + \lambda u_{nT}^2(s, x) = 0 \tag{8.31}$$

hold. Also precisely as above we show that $V_n^T(s, x) \geq V_{n+1}^T(s, x)$ for every $n = 1, 2, \ldots$. The function $V_n^T(s, x) = (C_n^T(s)x, x)$ is a non-negative quadratic form. It is well known that a monotone decreasing sequence of non-negative quadratic forms converges to a quadratic form. Let in our case this limit be $V_0^T(s, x) = (C_0^T(s)x, x)$. This and (8.30) imply the existence of the limit

$$\lim_{n \to \infty} u_{nT}(s, x) = u_{0T}(s, x) = v_{1T}(s)x_1 + \cdots + v_{lT}(s)x_l. \tag{8.32}$$

Finally, (8.29), (8.31) and (8.32) imply that the functions $V_0^T(s, x), u_{0T}(s, x)$ satisfy

$$\min_{u \in (-\infty, \infty)} [L_u V_0^T(s, x) + (\alpha(s)x, x) + \lambda u^2] = L_{u_{0T}} V_0^T(s, x) + (\alpha(s)x, x) + \lambda u_{0T}^2, \tag{8.33}$$

and moreover $V_0^T(T, x) = 0$, and

$$u_{0T}(s, x) = -\frac{(h(s), \frac{\partial V_0^T(s,x)}{\partial x}) + \sum_{r=1}^k (\sigma_r(s)x, \frac{\partial}{\partial x})(\varphi_r(s), \frac{\partial}{\partial x}) V_0^T(s, x)}{2\lambda + \sum_{r=1}^k (\varphi_r(s), \frac{\partial}{\partial x})^2 V_0^T(s, x)}. \tag{8.34}$$

This, together with Remark 8.5 implies that

$$\min_{u \in \mathcal{U}} \mathcal{T}_T^{s,x}(u) = \mathcal{T}_T^{s,x}(u_{0T}) = (C_0^T(s)x, x).$$

Thus the existence of a control which stabilizes the system (8.22) so that exponential stability in mean square is obtained, implies that

$$(C_0^T(s)x, x) \leq \min_{u \in \mathcal{U}} \mathcal{T}^{s,x}(u) \leq k|x|^2 \tag{8.35}$$

holds for a certain positive constant k. We may deduce now from (8.35) and (8.34) the existence of the limits

$$\lim_{T \to \infty} V_0^T(s, x) = V_0(s, x), \tag{8.36}$$

$$\lim_{T \to \infty} u_{0T}(s, x) = u_0(s, x), \tag{8.37}$$

where $V_0(s, x) = (C_0(s)x, x)$ is a quadratic form and $u_0(s, x)$ is an admissible control.

The expression in (8.33) which is preceded by the min sign is a parabola with respect to u and (8.36), (8.37) imply that the coefficients of this parabola have limits as $T \to \infty$, for any fixed s, x. Moreover the coefficient $A(s, x, t)$ at u^2 satisfies the inequality $A(s, x, T) \geq \lambda$. It follows that the functions $V_0(s, x)$, $u_0(s, x)$ are also related to each other by the Bellman equation

$$\min_{u \in (-\infty, \infty)} [L_u V_0(s, x) + (\alpha(s)x, x) + \lambda u^2] = L_{u_0} V_0(s, x) + (\alpha(s)x, x) + \lambda u_0^2 = 0.$$

The assertion which we wished to prove follows now by Theorem 8.1.   □

Note in the conclusion that the application of Theorem 8.2 to the stabilization problem for solutions of n-th order linear ODE with constant coefficients, perturbed by the Gaussian white noises, is considered by Nevelson [219].

Appendix A
Appendix to the First English Edition

We shall be concerned here with some results about stability of stochastic systems which were obtained since the time of the Russian edition of this book. We shall give detailed proofs only for a few of the results. Some theorems will be only stated and others merely mentioned. This varying degree of detail in our exposition is certainly not motivated by our feelings concerning the importance of the material. Rather, we have given throughout the priority to those results which are essentially connected with the main part of the book.

More information concerning other new interesting results can be found in the books [211] and [41], in the paper entitled "Stability of stochastic dynamical systems" and printed in the Lecture Notes in Mathematics, vol. 294 (1972), in the survey papers [277] and [150] and possibly other recent works.

A.1 Moment Stability and Almost Sure Stability for Linear Systems of Equations Whose Coefficients are Markov Processes

1. Consider the equation

$$\dot{Y}(t) = F(X(t))Y(t), \tag{A.1}$$

where $Y(t)$ is an m-dimensional vector, $F(x)$ is an $m \times m$ matrix and $X(t)$ is a random Markov process in $\mathbb{R}^l$ with a corresponding generator L_0 (see [64]).

Let us describe a general method due to Benderskii [27] of obtaining equations for the moments of positive integral degree of the process $Y(t)$. It is evident that $(X(t), Y(t))$ is a Markov process in $\mathbb{R}^{m+l}$. Its generator L acts on sufficiently smooth functions $V(x, y)$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^l$ by

$$LV(x, y) = L_0 V(x, y) + \left(F(x), \frac{\partial V(x, y)}{\partial y} \right).$$

We denote by $Y^{x,y}(t)$ the solution of (A.1) satisfying the initial conditions $X(0) = x$, $Y(0) = y$. The expectation of $Y^{x,y}(t)$ will be denoted by $u(x, y, t)$. Under quite general assumptions the coordinates $u_i = u_i(x, y, t)$ of the vector $u(x, y, t)$ satisfy the equation (see [64])

$$\frac{\partial u_i}{\partial t} = L u_i \tag{A.2}$$

and the initial conditions $u_i(x, y, 0) = y_i$. In the particular case when $X(t)$ is a diffusion process, the above is a consequence of Lemma 3.3.

Moreover (A.1) implies that

$$Y^{x,y}(t) = A(x, t, \omega)y,$$

where $A(x, t, \omega)$ is the fundamental matrix of the system (A.1) corresponding to the initial condition $X(0) = x$ for the process $X(t)$. Consequently the function

$$u(x, y, t) = \mathbf{E} A(x, t, \omega)y = B_1(x, t)y \tag{A.3}$$

depends linearly on y. Substituting (A.3) in (A.2) we obtain the equation

$$\frac{\partial B_1(x, t)}{\partial t} = L_0 B_1(x, t) + B_1(x, t) F(x), \quad B_1(x, 0) = J \tag{A.4}$$

for the matrix $B_1(x, t)$. (Here J stands for the $m \times m$ identity matrix.) After having found $B_1(x, t)$ from (A.4), we may calculate the first moment of $Y(t)$ from (A.3).

The above method allows us to find moments of arbitrary degree of the process $Y(t)$. Indeed, we may derive from (A.1) the equation

$$\frac{d(Y(t) \times Y(t))}{dt} = [F(X(t)) \times J + J \times F(X(t))](Y(t) \times Y(t)).$$

This is an equation of the form (A.1) for the product $Y(t) \times Y(t)$. Let $B_2(x, t)$ be the matrix which satisfies

$$\mathbf{E}[Y^{x,y}(t) \times Y^{x,y}(t)] = B_2(x, t)(y \times y).$$

Then, by the above, this matrix satisfies also

$$\frac{\partial B_2(x, t)}{\partial t} = L B_2(x, t) + B_2(x, t)[F(X(t)) \times J + J \times F(X(t))], \tag{A.5}$$

$$B_2(x, 0) = J \times J.$$

Also in the same way one can obtain equations for the moments of arbitrary high degree of the process $Y(t)$. Let us consider the following particular case.

Suppose that $X(t)$ is a diffusion process with generator

$$L = \left(b(x), \frac{\partial}{\partial x} \right) + \frac{1}{2} \sum_{r=1}^{k} \left(\sigma_r(x), \frac{\partial}{\partial x} \right)^2,$$

where $b(x)$ and $\sigma_r(x)$ are vectors in $\mathbb{R}^l$.

Then (A.4) and (A.5) are second order partial differential equations for the matrices $B_1(x, t)$ and $B_2(x, t)$. An analogous method for obtaining in this particular case equations for the moments was proposed by Frisch [81].

Suppose now that $X(t)$ is a stationary Markov process with finitely many states $x_1, \ldots, x_p$. Let us assume that the probability of transition from the i-th state to the j-th state during a time interval Δt is given for $\Delta t \to 0$ by $P_{ij}(\Delta t) = a_{ij}\Delta t + o(\Delta t)$, $i, j = 1, \ldots, p$, where a_{ij} are some constants. The action of the generator L of the process $X(t)$ on a function $V(x)$ defined at the points $x_1, \ldots, x_p$ is given by

$$LV(x_j) = \sum_{k=1,\, k \neq j}^{p} a_{jk}[V(x_k) - V(x_j)].$$

Hence we get for the matrices $B_1(x_j, t)$, $B_2(x_j, t)$ the following system of linear equations

$$\frac{dB_1(x_j, t)}{dt} = \sum_{k=1}^{p} a_{jk}[B_1(x_k, t) - B_1(x_j, t)] + B_1(x_j, t)F(x_j),$$

$$B_1(x_j, 0) = J,$$

$$\frac{dB_2(x_j, t)}{dt} = \sum_{k=1}^{p} a_{jk}[B_2(x_k, t) - B_2(x_j, t)] + B_2(x_j, t)[F(x_j) \times J + J \times F(x_j)],$$

$$B_2(x_j, 0) = J \times J, \quad j = 1, \ldots, p.$$

We see thus that for a time-homogeneous Markov process $X(t)$ with p states the stability problem for the solution of (A.1) is reduced to the investigation of the stability of the solution of a system of ODEs. Benderskii [26, 27], McKenna and Morrison [195] and others have discussed various partial cases of the above. For a system with discrete time analogous results were obtained by Benderskii in [28]. Darkovskii and Leibovich [50] considered the system (A.1) in the "mixed" case when the process $X(t)$ undergoes changes of value only at discrete moments of time.

2. Benderskii and Pastur [29, 30] have also investigated almost sure stability of the system (A.1). They applied ideas which are contained in Sects. 6.6, 6.7 and 6.8. Let us give a short presentation of their approach.

Just as in Sect. 6.7, let us introduce new variables defined by

$$\lambda = \frac{x}{|x|}, \qquad \rho = \ln |x|.$$

Then we obtain from (A.1) for every Markov process $X(t)$

$$\frac{d\lambda}{dt} = (F(X(t)) - (F(X(t))\lambda, \lambda)J)\lambda \tag{A.6}$$

(as before J is the identity matrix). This is a system of differential equations for the process $\lambda(t)$ on the sphere $|\lambda| = 1$ in $\mathbb{R}^n$. Here

$$\frac{d\rho}{dt} = (F(X(t))\lambda(t), \lambda(t)).$$

This implies the following formula, analogous to (6.65)

$$\frac{\rho(T) - \rho(0)}{T} = \frac{1}{T} \int_0^T (F(X(t))\lambda(t), \lambda(t)) \, dt. \tag{A.7}$$

We conclude that the stability problem, and also the problem of determining the exact rate of growth of the solution of (A.1) reduce to calculating the limit of the right side of (A.7) as $T \to \infty$. Let us note here that if $X(t)$ is a stationary process, then Theorem 2.4 implies that (A.6) possesses a stationary and a stationarily related to $X(t)$ solution $\lambda_0(t)$.

For certain concrete cases Benderskii and Pastur gave conditions under which every solution of (A.6) tends to a stationary one and the pair $X(t)$, $\lambda_0(t)$ is an ergodic stationary process. In this situation there exists a non-random limit

$$\lim_{T \to \infty} \frac{\rho(T) - \rho(0)}{T} = \int (F(x)l, l) \, dP(x, l),$$

where $P(x, l)$ is the joint distribution of $X(t)$, $\lambda_0(t)$. It is evident that in this case we have exponential almost sure stability if and only if the Lyapunov exponent

$$a^* = \int (F(x)l, l) \, dP(x, l)$$

is negative.

The above program has been realized to the fullest in case of second order equations (see the papers of Benderskii and Pastur [29, 30])

$$y'' + X(t)y = 0. \tag{A.8}$$

If we introduce the notation $y(t) = y_1$, $y'(t) = y_2$ we obtain a system of type (A.1) with

$$F(x, t) = \begin{pmatrix} 0 & 1 \\ -X(t) & 0 \end{pmatrix}.$$

It is convenient in this case, just as in Sect. 6.8 to replace here the vector $\lambda(t)$ by the variable $\varphi(t) = \arctan(y_1(t)/y_2(t))$ defined on the circle. Evidently we have

$$\frac{d\varphi(t)}{dt} = -(\sin^2 \varphi(t) + X(t)\cos^2 \varphi(t)).$$

Moreover (A.7) implies that

$$\lim_{T \to \infty} \frac{\rho(T) - \rho(0)}{T} = \lim_{T \to \infty} \frac{1}{2T} \int_0^T [1 - X(t)] \sin(2\varphi(t)) \, dt,$$

provided the last limit exists. In case of a stationary ergodic process $X(t)$ Bender-skii and Pastur [29] have given quite general conditions under which the above limit exists. If $X(t)$ is a Markov process, then the pair $X(t)$, $\varphi(t)$ is also a Markov process and we can apply the methods of Chap. 6 to calculate its joint stationary distribution P. In particular, if $X(t)$ is a Markov process which takes on only two values, then P can be calculated in an open form (see [30]). In conclusion let us remark that (A.8) is important because it makes its appearance also in physics; the behavior of the solutions of this equation is closely connected with the behavior of the solutions of Schroedinger's equation with a random potential (see [30]).

A.2 Almost Sure Stability of the Paths of One-Dimensional Diffusion Processes[1]

We shall be concerned in this section with the one-dimensional stochastic Itô equation whose coefficients are independent of time. Thus it is rather a particular model to which we have devoted much consideration in Chap. 5. However, in contrast to what we did in the Chaps. 5–7, we shall not investigate here the stability of the trivial solution, but the stability of an arbitrary path of the corresponding Markov process.

The idea of the basic (and somewhat unexpected) result can be described as follows: If $X(t)$ is a recurrent process in $\mathbb{R}^1$ which satisfies the equation

$$dX(t) = b(X(t))\,dt + \sigma(X(t))\,d\xi(t), \tag{A.9}$$

then, except for the trivial special case when the coefficients b and σ are periodic, the process $X(t)$ is an almost surely stable in the large solution of (A.9). Thus the solutions of (A.9) with various initial conditions are getting asymptotically close to each other as $t \to \infty$, as we might have expected. This is in general not the situation in case of deterministic systems.

To present all this more precisely, let us make the following assumptions:

1. The functions $b(x)$ and $\sigma(x)$ satisfy a Lipschitz condition on every compact set $K \subset \mathbb{R}^1$ and moreover $\sigma(x)$ never vanishes.

2. If we define

$$Q(x) = \int_0^x \exp\left\{-2\int_0^y \frac{b(z)}{\sigma^2(z)}\,dz\right\}dy,$$

then

$$Q(\pm\infty) = \pm\infty. \tag{A.10}$$

Only Assumption 2 seems to be somewhat restrictive. We have shown in Sect. 3.8 (Example 3.10) that Assumption 2 is equivalent to the recurrence property of the Markov process described by (A.9). We have observed in Sect. 4.2 that a recurrent

---

[1]Results of this section were proven in [137].

process $X(t)$ spends an infinite amount of time in every neighborhood of every point x. We shall prove now the stronger result that for every $x, x_0 \in \mathbb{R}^1$, $T > 0$ and $\alpha > 0$

$$\mathbf{P}^x \left\{ \bigcup_{\tau > T} \left\{ \omega : \sup_{\tau < t < \tau + 1} |X(t, \omega) - x_0| < \alpha \right\} \right\} = 1. \tag{A.11}$$

Let τ_α denote the instant at which the path of the process $X(t)$ reaches for the first time the set $|x - x_0| = \alpha$. Then we conclude by Lemma 3.4 that the function $\mathbf{P}^x \{ \tau_\alpha > s \} = u(s, x)$ satisfies the equation $(s > 0, |x - x_0| < \alpha)$

$$\frac{\partial u}{\partial s} = b(x) \frac{\partial u}{\partial x} + \frac{1}{2} \sigma^2(x) \frac{\partial^2 u}{\partial x^2}$$

and the initial and boundary conditions

$$u(0, x) = 1, \qquad u(s, x_0 \pm \alpha) = 0.$$

The strong maximum principle for parabolic equations implies that

$$\inf_{|x - x_0| \leq \alpha/2} \mathbf{P}^x \{ \tau_\alpha > 1 \} = \inf_{|x - x_0| \leq \alpha/2} u(1, x) = \beta > 0.$$

Just as in the proofs of Lemmas 4.1 and 4.5, let us consider now the cycles (parts of paths of the process $X(t)$) contained between the sets $\Gamma = \{ |x - x_0| = \alpha/2 \}$ and $\Gamma_1 = \{ |x - x_0| = \alpha \}$ with the time parameter exceeding T (there are infinitely many such cycles due to the recurrence property of the process). Let A_i denote the event that the second half of the i-th cycle lasts for a time longer than 1. Then

$$\mathbf{P} \left\{ \bigcup_{\tau > T} \left\{ \omega : \sup_{\tau < t < \tau + 1} |X(t, \omega) - x_0| < \alpha \right\} \right\} \geq \mathbf{P} \left\{ \bigcup_{i=1}^{\infty} A_i \right\}.$$

Since

$$\mathbf{P} \{ A_i \mid A_1^c, \ldots, A_{i-1}^c \} = \mathbf{P} \{ A_i \} \geq \inf_{|x - x_0| \leq \alpha/2} \mathbf{P}^x \{ \tau_\alpha > 1 \} = \beta > 0,$$

we have

$$\mathbf{P} \left\{ \bigcup_{i=1}^{\infty} A_i \right\} = 1 - \mathbf{P} \{ A_1^c \} \mathbf{P} \{ A_2^c \mid \overline{A}_1^c \} \mathbf{P} \{ A_3^c \mid A_1^c, A_2^c \} \cdots = 1$$

and therefore (A.11) is proved.

Remark A.1 Using the method of the above proof, we can establish a more general result: Given an arbitrary recurrent diffusion process in $\mathbb{R}^l$ with a non-degenerate diffusion matrix, one can assemble the segments of its paths into a set which is everywhere dense in $C[0, h]$. Here $C[0, h]$ denotes the space of $\mathbb{R}^l$-valued continuous functions defined on the interval $[0, h]$ with the metric topology of uniform convergence.

Let us consider now the stochastic process

$$Y(t) = Q(X(t)).$$

(This is equivalent to considering the process $X(t)$ in the canonical scale, see [64].) By (A.10), this transformation maps $\mathbb{R}^1$ onto the whole real line $\mathbb{R}^1$. Moreover Itô's formula implies

$$dY(t) = \sigma_1(Y(t)) \, d\xi(t), \qquad (A.12)$$

where

$$\sigma_1(y) = \sigma(Q^{-1}(y)) q(Q^{-1}(y)), \qquad q(y) = Q'(y).$$

It follows easily from Assumptions 1 and 2 that $Y(t)$ is a regular and recurrent process. Let us show first that it is regular. The generator of the process (A.12) is given by

$$L = \frac{1}{2} \sigma_1^2(y) \frac{\partial^2}{\partial y^2}.$$

It follows from this that the auxiliary function

$$V(y) = |y|$$

satisfies for $|y| > 1$ the conditions (3.43) and (3.44). This and Theorem 3.5 (see also Remark 3.4) imply the regularity of the process $Y(t)$. Recurrence follows from the Example 3.10.

Lemma A.1 *Let $Y_i(t)$, $i = 1, 2$, denote the solutions of (A.12) which satisfy the initial conditions $Y_i(0) = y_i$, $y_1 < y_2$. Then there exists a non-negative random variable $\zeta < \infty$ such that*

$$\lim_{t \to \infty} [Y_2(t) - Y_1(t)] = \zeta.$$

Proof Let us first observe that $Y_2(t) \geq Y_1(t)$ holds for all $t \geq 0$. This intuitively obvious fact follows from the "comparison theorem" of Skorokhod given in Sect. 5.3. (One can prove this also in another way by observing that the solution of (A.12) with given initial condition is unique in the case when the initial condition is specified at a random and not dependent on the future time moment τ, and next selecting $\tau = \inf\{t : Y_1(t) = Y_2(t)\}$. Let $Z(t) = Y_2(t) - Y_1(t)$.) The above implies that $Z(t) \geq 0$. Let τ_R denote the moment of the first exit from the circle of radius R of the two-dimensional Markov process $(Y_1(t), Y_2(t))$. Properties of the stochastic integral imply now that the process

$$Z(\tau_R \wedge t) = y_2 - y_1 + \int_0^{\tau_R \wedge t} [\sigma_1(Y_2(s)) - \sigma_1(Y_1(s))] \, d\xi(s)$$

is a positive martingale, that is

$$\mathbf{E}(Z(\tau_R \wedge t) \,|\, \mathcal{N}_s) = Z(\tau_R \wedge s), \quad t > s. \qquad (A.13)$$

Here $\mathcal{N}_s$ denotes the σ-algebra of events generated by the run of the process $\xi(t)$ along the time interval $[0, s]$. The regularity of this process implies that $\lim_{R \to \infty} \tau_R = \infty$, whence, passing to the limit in (A.13) as $R \to \infty$, and applying Fatou's lemma, we get

$$\mathbf{E}(Z(t) \mid \mathcal{N}_s) \leq Z(s). \tag{A.14}$$

This inequality means that $Z(t)$ is a positive supermartingale. To complete the proof of the lemma, it suffices to apply Theorem 5.1. $\qquad\square$

Corollary A.1 *For every initial $y \in \mathbb{R}^1$, the solution of (A.12) is stable in the mean and in probability.*

Proof Indeed, (A.14) implies that for all $t \geq 0$

$$\mathbf{E}[Y_2(t) - Y_1(t)] \leq y_2 - y_1,$$

whence

$$\lim_{y_2 - y_1 \to 0} \sup_{t \geq 0} \mathbf{E}[Y_2(t) - Y_1(t)] = 0.$$

The inequality for supermartingales implies also

$$\mathbf{P}\left\{ \sup_{t > 0}[Y_2(t) - Y_1(t)] \geq \varepsilon \right\} \to \frac{y_2 - y_1}{\varepsilon} \to 0 \quad \text{as } y_2 - y_1 \to 0$$

for every positive ε. $\qquad\square$

The following lemma gives stability conditions in the large for any solution of (A.12).

Lemma A.2 *Let $Y_i(t)$, $i = 1, 2$, denote the solutions of (A.12) which satisfy the initial conditions $Y_i(0) = y_i$, $y_1 < y_2$. Then the following assertions hold true.*

1. *If the function $\sigma_1(y)$ is not periodic, then $\zeta = \lim_{t \to \infty}[Y_2(t) - Y_1(t)] = 0$.*
2. *If the function $\sigma_1(y)$ is periodic with period θ and $k = (y_2 - y_1)/\theta$ is an integer (thus, in particular if $\sigma_1(y)$ does not depend on y), then $Y_2(t) - Y_1(t) = y_2 - y_1$.*
3. *If the function $\sigma_1(y)$ is not identically constant and is periodic with period θ, and $k = (y_2 - y_1)/\theta$ is not an integer, then the distribution of ζ is concentrated at the two points $\theta[k]$ and $\theta([k] + 1)$.*

Proof Let us show that almost surely

$$\sigma_1(y + \zeta) = \sigma_1(y) \tag{A.15}$$

for all $y \in \mathbb{R}^1$. Suppose that this is not the case. Then the continuity of $\sigma_1(y)$ implies that

$$\mathbf{P}\left\{ \inf_{|y - y_0 - \zeta| < \delta, \, |z - y_0| < \delta} |\sigma_1(y) - \sigma_1(z)| > \delta \right\} \geq p_1 \tag{A.16}$$

for some y_0 and some positive δ and p_1. Lemma A.1 implies that the stochastic integral

$$\int_0^t [\sigma_1(Y_2(s)) - \sigma_1(Y_1(s))] \, d\xi(s)$$

has a finite limit as $t \to \infty$. Using this and the inequality

$$\mathbf{P}\left\{ \int_T^\infty [\sigma_1(Y_2(s)) - \sigma_1(Y_1(s))]^2 ds > \delta^2 \right\}$$

$$\leq 1 - 2\mathbf{P}\{\xi(\delta^2) > a\} + \mathbf{P}\left\{ \sup_{T \leq t < \infty} \int_T^t [\sigma_1(Y_2(s)) - \sigma_1(Y_1(s))] \, d\xi(s) > a \right\}$$

which is valid for every $a > 0$ (see [93, Corollary 1 to Theorem 1.4.3]), we get that for every $\varepsilon > 0$, there is a $T_1(\varepsilon)$ such that

$$\mathbf{P}\left\{ \int_T^\infty [\sigma_1(Y_2(s)) - \sigma_1(Y_1(t))]^2 \, ds > \delta^2 \right\} < \varepsilon \tag{A.17}$$

for all $T \geq T_1(\varepsilon)$.

Let us select $T_2(\varepsilon)$ so that

$$\mathbf{P}\left\{ \sup_{t > T_2(\varepsilon)} |Y_2(t) - Y_1(t) - \zeta| > \delta/2 \right\} < \varepsilon \tag{A.18}$$

holds. (A.11) and (A.18) imply that there exists a random variable $\tau > \max(T_1(\varepsilon), T_2(\varepsilon))$ such that

$$\sup_{\tau < s < \tau+1} |Y_1(s) - y_0| < \delta/2, \qquad \mathbf{P}\left\{ \sup_{\tau < s < \tau+1} |Y_2(s) - y_0 - \zeta| > \delta \right\} < \varepsilon.$$

We deduce from (A.16), (A.17) and the last two inequalities that

$$\varepsilon > \mathbf{P}\left\{ \int_\tau^\infty [\sigma_1(Y_2(s)) - \sigma_1(Y_1(s))]^2 \, ds > \delta^2 \right\}$$

$$\geq \mathbf{P}\left\{ \int_\tau^{\tau+1} [\sigma_1(Y_2(s)) - \sigma_1(Y_1(s))]^2 \, ds > \delta^2 \right\} \geq p_1 - \varepsilon.$$

Since ε is here arbitrary, the last inequality is in contradiction with $p_1 > 0$. Thus equality (A.15) is proved. $\square$

(A.15) implies immediately the first assertion of the theorem.

Suppose now that the function $\sigma_1(y)$ is periodic, with period θ and $k = (y_2 - y_1)/\theta$ is an integer. Then the function $\widetilde{Y}_1(t) = Y_1(t) + y_2 - y_1 = Y_1(t) + k\theta$ satisfies (A.12) and the initial condition $\widetilde{Y}_1(0) = y_2$. The second assertion of the theorem follows now from the uniqueness of the solution of (A.12).

Finally, assume that $\sigma_1(y)$ is a not identically constant periodic function with period θ and $k = (y_2 - y_1)/\theta$ is not an integer. Let us consider the system

$$
\begin{aligned}
dY(t) &= \sigma_1(Y(t))\,d\xi(t), \\
dZ(t) &= (\sigma_1(Y(t) + Z(t)) - \sigma_1(Y(t)))\,d\xi(t).
\end{aligned}
\tag{A.19}
$$

This system has got two solutions: $Z_1(t) = Y_2(t) - Y_1(t)$ and $\widetilde{Y}_1(t) = Y_1(t)$, $\widetilde{Y}_2(t) = Y_\theta(t)$, $Z_2(t) = Y_\theta(t) - Y_0(t)$, where $Y_\theta(t)$ and $Y_0(t)$ are the solutions of the first equation in (A.19) defined by the initial conditions $Y_\theta(0) = \theta([k] + 1)$ and $Y_0(0) = 0$. From the uniqueness of the solutions of this system and from $y_2 - y_1 = Z_1(0) < Z_2(0) = \theta([k] + 1)$ follows $Z_1(t) \leq Z_2(t)$ for all t. Moreover the second assertion of the theorem, proved above, implies that $Z_2(t) \equiv \theta([k] + 1)$. Thus $Y_2(t) - Y_1(t) \leq \theta([k] + 1)$. We show in the same way that $\theta[k] \leq Y_2(t) - Y_1(t)$ holds for all t. These inequalities and (A.15) imply the third assertion of the theorem. □

Let us return now to the original equation (A.9). We put $r(x_1, x_2) = |Q(x_2) - Q(x_1)|$. It is evident that $r(x_1, x_2)$ defines a metric in $\mathbb{R}^1$. Lemmas A.1 and A.2 yield now the following theorem which gives criteria for stability in the large of an arbitrary solution of (A.9) which describes a recurrent Markov process.

Theorem A.1 *Suppose that Assumptions 1 and 2 are satisfied. Let $X_1(t)$, $X_2(t)$ be solutions of (A.1) which satisfy the initial conditions $X_1(0) = x_1$, $X_2(0) = x_2$. Then there exists a finite limit*

$$
\lim_{t \to \infty} r(X_1(t), X_2(t)) = \zeta.
$$

The identity $\zeta \equiv 0$ holds for every x_1, x_2 if and only if the function $\sigma_1(y) = \sigma(Q^{-1}(y))q(Q^{-1}(y))$ is not periodic. In the case when $\sigma_1(y)$ is a periodic function with period θ, and the number $k = |Q(x_2) - Q(x_1)|$ is not an integer, the distribution of ζ is concentrated in the two points $[k]\theta$ and $([k] + 1)\theta$, and if k is an integer, then $r(X_1(t), X_2(t)) \equiv |Q(x_2) - Q(x_1)|$ for any $t \geq 0$. The last equality holds also in the case when $\sigma_1(y)$ does not depend on y.

Corollary A.2 *Let $X(t)$ be a positive recurrent Markov process. Then any two solutions $X_1(t)$, $X_2(t)$ of (A.9) with initial conditions $X_1(0) = x_1$, $X_2(0) = x_2$ satisfy*

$$
\lim_{t \to \infty} r(X_1(t), X_2(t)) = 0.
$$

Indeed, since (A.9) describes a positive recurrent Markov process, we have (see [64]) that

$$
\int_{-\infty}^{\infty} \frac{dx}{\sigma^2(x)q(x)} < \infty.
$$

Thus

$$\int_{-\infty}^{\infty} \frac{dy}{(\sigma_1(y))^2} = \int_{-\infty}^{\infty} \frac{dQ(x)}{(\sigma(x)q(x))^2} = \int_{-\infty}^{\infty} \frac{dx}{\sigma^2(x)q(x)} < \infty.$$

It follows from this that the function $\sigma_1(y)$ is not periodic.

The theorem which now follows says that every solution of Itô's homogeneous SDE in $\mathbb{R}^1$ describing a recurrent Markov process is stable in the mean and in probability with respect to the metric $r(x_1, x_2)$.

Theorem A.2 *Suppose that Assumptions* 1 *and* 2 *above are satisfied. Then any two solutions* $X_1(t)$, $X_2(t)$ *of* (A.9) *with initial conditions* $X_1(0) = x_1$, $X_2(0) = x_2$ *satisfy*

$$\lim_{x_2 - x_1 \to 0} \sup_{t \geq 0} r(X_1(t), X_2(t)) = 0 \quad (a.s),$$

$$\lim_{x_2 - x_1 \to 0} \mathbf{P}\left\{ \sup_{t \geq 0} r(X_1(t), X_2(t)) > \varepsilon \right\} = 0,$$

for every positive ε.

The theorem is a consequence of Corollary A.1 and of the recurrence property of the process $X(t)$.

A.3 Reduction Principle

There is known in the stability theory of deterministic systems the so-called reduction principle which is basic for the investigation of critical stability situations. This principle permits us to reduce the investigation of the stability of an $(l + m)$-dimensional system $X(t)$, $Y(t)$ to investigating the stability of two systems: the l-dimensional system of the first approximation of the vector $X(t)$ (the coefficients of this approximation are assumed independent of y), and the m-dimensional system obtained by substituting $X = 0$ in the equations for Y (see [190], Sect. 91). Here the reduction principle will be used for the simpler case when both ramified systems are by linear approximation uniformly stable in the large. Although the result cannot be directly applied to the investigation of critical cases, it nevertheless offers a possibility of simplifying stability investigations in some cases of practical importance.

Thus let there be given an $(l + m)$-dimensional Markov process $X(t)$, $Y(t)$ described by the system of SDEs

$$dX(t) = b(t, X(t), Y(t))\, dt + \sum_{r=1}^{k} \sigma_r(t, X(t), Y(t))\, d\xi_r(t),$$

$$dY(t) = \tilde{b}(t, X(t), Y(t))\, dt + \sum_{r=1}^{k} \tilde{\sigma}_r(t, X(t), Y(t))\, d\xi_r(t),$$

$$\text{(A.20)}$$

where the vectors X, b, σ_r are l-dimensional and Y, $\tilde{b}$, $\tilde{\sigma}_r$ are m-dimensional. As usually, let us assume that the coefficients b, σ_r, $\tilde{b}$, $\tilde{\sigma}_r$ satisfy conditions (5.2) and (5.3) so that, in particular, the system (A.20) has the trivial solution

$$X(t) \equiv 0, \qquad Y(t) \equiv 0.$$

Moreover let us assume that the derivatives with respect to x, y of the coefficients of the system (A.20) are uniformly continuous with respect to t, and

$$\frac{\partial b(t,0,0)}{\partial y} \equiv 0, \qquad \frac{\partial \sigma_r(t,0,0)}{\partial y} \equiv 0.$$

Thus in the system of equations for the first approximation

$$dX(t) = \frac{\partial b(t,0,0)}{\partial x} X(t)\,dt + \sum_{r=1}^{k} \frac{\partial \sigma_r(t,0,0)}{\partial x} X(t)\,d\xi_r(t), \qquad (A.21)$$

$$dY(t) = \left(\frac{\partial b(t,0,0)}{\partial x} X(t) + \frac{\partial b(t,0,0)}{\partial y} Y(t) \right) dt$$

$$+ \sum_{r=1}^{k} \left(\frac{\partial \sigma_r(t,0,0)}{\partial x} X(t) + \frac{\partial \sigma_r(t,0,0)}{\partial y} Y(t) \right) d\xi_r(t) \qquad (A.22)$$

the component of $X(t)$ is also a Markov process. Theorem 7.2 (and the subsequent Remark 7.1) imply that if the trivial solution is uniformly stable in the large for the system of (A.21), (A.22), then it is asymptotically stable in probability for the system (A.20). The theorem which now follows allows us to say somewhat more.

For a one-dimensional process $X(t)$ the theorem was proved by Pinsky [227] who used another method.

Theorem A.3 *Suppose the above assumptions about the coefficients of the system* (A.20) *are valid. Assume further that the trivial solution is uniformly stable in the large for the system* (A.21) *and for the system*

$$dY(t) = \frac{\partial \tilde{b}(t,0,0)}{\partial y} Y(t)\,dt + \sum_{r=1}^{k} \frac{\partial \tilde{\sigma}_r(t,0,0)}{\partial y} Y(t)\,d\xi_r(t). \qquad (A.23)$$

Then the trivial solution is asymptotically stable in probability for the system (A.20).

Proof The assumptions of the theorem and the Theorems 6.7 and 6.2 imply that there exist for a sufficiently small $p > 0$ two homogeneous functions $V_1(t,x)$,

$V_2(t, y)$ of homogeneity degree p such that

$$
\left.
\begin{aligned}
& k_1|x|^p \le V_1(t, x) \le k_2|x|^p, \\
& \left|\frac{\partial V_1(t, x)}{\partial x_i}\right| \le k_3|x|^{p-1}, \quad \left|\frac{\partial^2 V_1(t, x)}{\partial x_i \partial x_j}\right| \le k_3|x|^{p-2}, \quad i, j = 1, \ldots, l, \\
& L_1 V_1(t, x) \le -|x|^p,
\end{aligned}
\right\}
\quad (A.24)
$$

$$
\left.
\begin{aligned}
& k_1|y|^p \le V_2(t, y) \le k_2|y|^p, \\
& \left|\frac{\partial^2 V(t, y)}{\partial y_i}\right| \le k_3|y|^{p-1}, \quad \left|\frac{\partial^2 V_2(t, y)}{\partial y_i \partial y_j}\right| \le k_3|y|^{p-2}, \quad i, j = 1, \ldots, m, \\
& L_2 V_2(t, y) \le -|y|^p,
\end{aligned}
\right\}
$$

$$(A.25)$$

where

$$
L_1 = \frac{\partial}{\partial t} + \left(\frac{\partial b(t, 0, 0)}{\partial x} x, \frac{\partial}{\partial x}\right) + \frac{1}{2}\sum_{r=1}^{k}\left(\frac{\partial \sigma_r(t, 0, 0)}{\partial x} x, \frac{\partial}{\partial x}\right)^2,
$$

$$
L_2 = \frac{\partial}{\partial t} + \left(\frac{\partial \tilde{b}(t, 0, 0)}{\partial y} y, \frac{\partial}{\partial y}\right) + \frac{1}{2}\sum_{r=1}^{k}\left(\frac{\partial \tilde{\sigma}_r(t, 0, 0)}{\partial y} y, \frac{\partial}{\partial y}\right)^2
$$

are the generators of the systems (A.21) and (A.23) respectively. (Here and below we adopt the same notation $(\cdot, \cdot)$ for the inner product in $\mathbb{R}^l$ as well as in $\mathbb{R}^m$.) The generator of the system (A.21), (A.22) is

$$
\begin{aligned}
L = {} & \frac{\partial}{\partial t} + \left(\frac{\partial b(t, 0, 0)}{\partial x} x, \frac{\partial}{\partial x}\right) + \frac{1}{2}\sum_{r=1}^{k}\left(\frac{\partial \sigma_r(t, 0, 0)}{\partial x} x, \frac{\partial}{\partial x}\right)^2 \\
& + \left(\frac{\partial \tilde{b}(t, 0, 0)}{\partial x} x + \frac{\partial \tilde{b}(t, 0, 0)}{\partial y} y, \frac{\partial}{\partial y}\right) \\
& + \frac{1}{2}\sum_{r=1}^{k}\left(\frac{\partial \tilde{\sigma}_r(t, 0, 0)}{\partial x} x + \frac{\partial \tilde{\sigma}_r(t, 0, 0)}{\partial y} y, \frac{\partial}{\partial y}\right)^2 \\
& + \sum_{r=1}^{k}\left(\frac{\partial \sigma_r(t, 0, 0)}{\partial x} x, \frac{\partial}{\partial x}\right)\left(\frac{\partial \tilde{\sigma}_r(t, 0, 0)}{\partial x} x + \frac{\partial \tilde{\sigma}_r(t, 0, 0)}{\partial y} y, \frac{\partial}{\partial y}\right).
\end{aligned}
$$

Let us consider now the auxiliary function

$$
W(t, x, y) = [V_1^{2/p}(t, x) + \varepsilon V_2^{2/p}(t, y)]^{p/2} + A V_1(t, x), \quad (A.26)
$$

where V_1, V_2 are functions satisfying the condition (A.24), (A.25), and the values of the constants $\varepsilon > 0$, $A > 0$ will be specified below. (A.24) and (A.25) evidently

imply that

$$k_4(|x|^p + |y|^p) \leq W(t, x, y) \leq k_5(|x|^p + |y|^p) \tag{A.27}$$

holds for certain $k_4 > 0$, $k_5 > 0$ and arbitrary $\varepsilon > 0$, $A > 0$. Moreover it is evident that $x = 0$ is an unattainable invariant set for the process $X(t)$, $Y(t)$. Hence in the case when $X(0) \neq 0$, we can apply Itô's formula and consequently also Theorem 5.11 to any function W which is not differentiable on the hyperplane $x = 0$. Considering this and (A.27), we see that it will suffice to prove that outside the set $x = 0$ the function $W(t, x, y)$, defined by (A.26) satisfies

$$LW(t, x, y) \leq -k_6(|x|^p + |y|^p), \quad k_6 > 0 \tag{A.28}$$

for some ε and A.

Indeed, by the remark we made above, (A.27), (A.28) and Theorem 5.11 imply that the system (A.21), (A.22) is exponentially p-stable. The assertion of our theorem follows now by Theorem 7.2.

Let $W_i = V_i^{2/p}$. Then evidently

$$L_1 V_1(t, x) = \frac{p}{2} W_1^{p/2-1} L_1 W_1 + \frac{1}{8} p(p-2) W_1^{p/2-2} \sum_{r=1}^{k} \left(\frac{\partial \sigma_r(t, 0, 0)}{\partial x} x, \frac{\partial W_1}{\partial x} \right)^2.$$

From this and from (A.13) follows easily

$$W_1 L_1 W_1 + \frac{p-2}{4} \sum_{r=1}^{k} \left(\frac{\partial \sigma_r(t, 0, 0)}{\partial x} x, \frac{\partial}{\partial x} \right)^2 W_1 \leq -k_7 |x|^4$$

for some constant $k_7 > 0$. In a similar way we obtain

$$W_2 L_2 W_2 + \frac{p-2}{4} \sum_{r=1}^{k} \left(\frac{\partial \widetilde{\sigma}_r(t, 0, 0)}{\partial y} y, \frac{\partial}{\partial y} \right)^2 W_2 \leq -k_8 |y|^4, \quad k_8 > 0.$$

These inequalities, together with (A.24), (A.25) imply that for some constants $k_9, k_{10} > 0$ which do not depend on ε

$$L[(W_1 + \varepsilon W_2)]^{p/2}$$

$$= \frac{p}{2} (W_1 + \varepsilon W_2)^{p/2-2} \Bigg\{ (W_1 + \varepsilon W_2) L_1 W_1$$

$$+ \frac{p-2}{4} \sum_{r=1}^{k} \left(\frac{\partial \sigma_r(t, 0, 0)}{\partial x} x, \frac{\partial}{\partial x} \right)^2 W_1$$

$$+ \varepsilon (W_1 + \varepsilon W_2) L_2 W_2$$

$$+ \frac{p-2}{4} \varepsilon^2 \sum_{r=1}^{k} \left(\frac{\partial \widetilde{\sigma}_r(t,0,0)}{\partial x} x + \frac{\partial \widetilde{\sigma}_r(t,0,0)}{\partial y} y, \frac{\partial}{\partial y} \right)^2 W_2$$

$$+ \varepsilon (W_1 + \varepsilon W_2) \left(\frac{\partial \widetilde{b}(t,0,0)}{\partial x} x, \frac{\partial W_2}{\partial y} \right)$$

$$+ \varepsilon (W_1 + \varepsilon W_2) \sum_{r=1}^{k} \left(\frac{\partial \widetilde{\sigma}_r(t,0,0)}{\partial x} x, \frac{\partial}{\partial y} \right)^2 W_2$$

$$+ \epsilon (W_1 + \varepsilon W_2) \sum_{r=1}^{k} \left(\frac{\partial \widetilde{\sigma}_r(t,0,0)}{\partial x} x, \frac{\partial}{\partial y} \right) \left(\frac{\partial \widetilde{\sigma}_r(t,0,0)}{\partial y} y, \frac{\partial}{\partial y} \right) W_2$$

$$+ \frac{\varepsilon(p-2)}{4} \left(\frac{\partial \widetilde{\sigma}_r(t,0,0)}{\partial x} x + \frac{\partial \widetilde{\sigma}_r(t,0,0)}{\partial y} y, \frac{\partial W_2}{\partial y} \right) \left(\frac{\partial \sigma_r(t,0,0)}{\partial x} x, \frac{\partial W_1}{\partial x} \right) \Bigg\}$$

$$\leq k_9 (W_1 + \varepsilon W_2)^{p/2-2} [-k_{10}(|x|^4 + \varepsilon^2 |y|^4) + \varepsilon |x|^2 |y|^2 + \varepsilon |y||x|^3 + \varepsilon |y|^3 |x|]. \tag{A.29}$$

Moreover $L_1 V_1 \leq -|x|^p$ by (A.24). From this and from the obvious inequalities

$$|x|^p \geq \frac{|x|^p W_1^{2-p/2}}{(W_1 + \varepsilon W_2)^{2-p/2}} \geq k_{11} (W_1 + \varepsilon W_2)^{p/2-2} |x|^4$$

we find that

$$L V_1 \leq -k_{11} (W_1 + \varepsilon W_2)^{p/2-2} |x|^4. \tag{A.30}$$

The inequalities (A.29) and (A.30) yield an estimate

$$L W \leq k_9 (W_1 + \varepsilon W_2)^{p/2-2} [-A k_{12} |x|^4 - k_{10} \varepsilon^2 |y|^4$$
$$+ \varepsilon |x|^2 |y|^2 + \varepsilon |y||x|^3 + \varepsilon^2 |y|^3 |x|],$$

with independent of ε and A constants k_i.

It is evident that if ε is sufficiently small and A is sufficiently large, then the expression between square brackets is a negative definite form of degree four (it is easiest to see this by putting $y \sqrt{\varepsilon} = z$). This implies (A.28) and hence also the assertion of the theorem. $\qquad \square$

A.4 Some Further Results

Let us conclude this section by mentioning some further investigations related to the questions considered in this book.

To begin with, there is a series of works by Friedman [76], Friedman and Pinsky [78–80] and Pinsky [226]. In these papers there are given stability conditions for

a point and for an invariant set of a similar sort as we gave here in Chap. 5 and in Sect. 7.7.2. Moreover, special considerations are devoted to the angle component $\varphi(t)$ for processes in the plane. An interesting generalization of the invariance theorem of La Salle to stochastic systems was given by Kushner [170, 171]. Morozan [212] clarified the idea of boundedness in various probabilistic meanings for solutions of stochastic dynamical systems of a more general type than the ones considered in the present book.

Much attention has been given to the investigation of stability problems for stochastic systems with delay. Kolmanovskii [143], Kushner [173] and others proved general theorems of the Lyapunov type. Shaihet [244] considered stability in the first approximation, and Tsarkov [262] gave stability criteria in mean square for linear systems.[2]

---

[2]After the first edition of this book the stability problems for SDEs with delay were considered by X. Mao and other authors, see, e.g., [192] and references therein.

Appendix B
Appendix to the Second Edition. Moment Lyapunov Exponents and Stability Index[3]

B.1 Preliminaries

Moment Lyapunov exponents are of great importance for investigating asymptotic behavior of solutions of SDEs. For $p \in \mathbb{R}$ and for a linear system with constant coefficients

$$dX(t) = BX(t)\,dt + \sum_{r=1}^{k} \sigma_r X(t)\,d\xi_r(t), \tag{B.1}$$

the p-th moment Lyapunov exponent of a solution of (B.1) is defined by

$$g(p; x) = \overline{\lim_{t \to \infty}} \frac{1}{t} \ln \mathbf{E}|X^x(t)|^p, \quad x \neq 0. \tag{B.2}$$

It will be shown below that given a certain nondegeneracy condition on (B.1) $g(p; x)$ is independent of x, $g(p; x) = g(p)$ for all $p \in \mathbb{R}$, $x \neq 0$ and

$$g(p) = \lim_{t \to \infty} \frac{1}{t} \ln \mathbf{E}|X^x(t)|^p, \quad x \neq 0. \tag{B.3}$$

For the one-dimensional equation ($d = 1$)

$$dX(t) = bX(t)\,dt + \sum_{r=1}^{k} \sigma_r X(t)\,d\xi_r(t) \tag{B.4}$$

(b and σ_r are some constants), it can be easily calculated

$$g(p) = pb + \frac{1}{2}p(p-1)\sum_{r=1}^{k}\sigma_r^2. \tag{B.5}$$

In what follows, we consider $d \geq 2$.

---

[3]Written jointly with G.N. Milstein.

The process

$$\Lambda^\lambda(t) = X^x(t)/|X^x(t)|, \quad \lambda = x/|x|, \; x \neq 0,$$

defined on the unit sphere $\mathbf{S}^{d-1} = \{|x| = 1\}$ in $\mathbb{R}^d$ with center at the origin (it is introduced in Sect. 6.7), plays a significant role in development of the concept of moment Lyapunov exponents.

With help of Itô formula we obtain that the process $\Lambda^\lambda(t)$, $|\lambda| = 1$, satisfies the SDE

$$d\Lambda = q(\Lambda)\,dt + \sum_{r=1}^{k} h_r(\Lambda)\,d\xi_r(t), \tag{B.6}$$

where the vector fields $q(\lambda)$ and $h_r(\lambda)$, $r = 1, \ldots, k$, on $\mathbf{S}^{d-1}$ are equal to

$$q(\lambda) = B\lambda - (B\lambda, \lambda)\lambda - \frac{1}{2}\sum_{r=1}^{k}(\sigma_r\lambda, \sigma_r\lambda)\lambda$$

$$- \sum_{r=1}^{k}(\sigma_r\lambda, \lambda)\sigma_r\lambda + \frac{3}{2}\sum_{r=1}^{k}(\sigma_r\lambda, \lambda)^2\lambda, \tag{B.7}$$

$$h_r(\lambda) = \sigma_r\lambda - (\sigma_r\lambda, \lambda)\lambda, \quad r = 1, \ldots, k.$$

Let us recall the equation for $\rho(t) = \ln|X^x(t)|$ (see (6.63)):

$$d\rho(t) = Q(\Lambda)dt + \sum_{r=1}^{k}(\sigma_r\Lambda, \Lambda)\,d\xi_r(t), \quad \Lambda = \Lambda^\lambda(t), \; \lambda = x/|x|, \; \rho(0) = \ln|x|, \tag{B.8}$$

where

$$Q(\lambda) = (B\lambda, \lambda) + \frac{1}{2}\sum_{r=1}^{k}(\sigma_r\lambda, \sigma_r\lambda) - \sum_{r=1}^{k}(\sigma_r\lambda, \lambda)^2. \tag{B.9}$$

From (B.8)

$$|X^x(t)|^p = |x|^p \exp\left\{ p\int_0^t Q(\Lambda)\,ds + p\sum_{r=1}^{k}\int_0^t (\sigma_r\Lambda, \Lambda)\,d\xi_r(s) \right\}. \tag{B.10}$$

Hence $Z(t) := |X^x(t)|^p$ satisfies the scalar linear equation

$$dZ = \left(pQ(\Lambda) + \frac{1}{2}p^2\sum_{r=1}^{k}(\sigma_r\Lambda, \Lambda)^2 \right) Z\,dt + pZ\sum_{r=1}^{k}(\sigma_r\Lambda, \Lambda)\,d\xi_r(t),$$

$$Z(0) = |x|^p. \tag{B.11}$$

Let us note that now it is not difficult to get for the two-dimensional system of the form (B.1) with

$$B = \begin{pmatrix} a_0 & b_0 \\ -b_0 & a_0 \end{pmatrix}, \qquad \sigma_r = \begin{pmatrix} a_r & b_r \\ -b_r & a_r \end{pmatrix}, \quad r = 1, \ldots, k,$$

that

$$g(p) = p\left(a_0 + \frac{1}{2}\sum_{r=1}^{k}(b_r^2 - a_r^2)\right) + \frac{1}{2}p^2\sum_{r=1}^{k}a_r^2.$$

SDE (B.6) can be considered on the whole space $\mathbb{R}^d$. In such a case $\Lambda(0) = \lambda$ where $|\lambda|$ is not necessarily equal to 1. We have

$$d(1 - (\Lambda, \Lambda)) = -(1 - (\Lambda, \Lambda))\left[2(B\Lambda, \Lambda)\right.$$

$$\left. + \sum_{r=1}^{k}(\sigma_r\Lambda, \sigma_r\Lambda) - 4\sum_{r=1}^{k}(\sigma_r\Lambda, \Lambda)^2\right]dt$$

$$- 2(1 - (\Lambda, \Lambda))\sum_{r=1}^{k}(\sigma_r\Lambda, \Lambda)\,d\xi_r(t). \tag{B.12}$$

From here we obtain the confirmation of the fact that the unit sphere $\mathbf{S}^{d-1} = \{\lambda : (\lambda, \lambda) = 1\}$ is an invariant manifold for (B.6). The sets $\{\lambda : \lambda = 0\}$, $\{\lambda : 0 < (\lambda, \lambda) < 1\}$, and $\{\lambda : (\lambda, \lambda) > 1\}$ are invariant ones for (B.6) as well.

Sometimes it is more convenient to consider moment Lyapunov exponents for linear SDEs in the Stratonovich form

$$dX(t) = \sigma_0 X(t)\,dt + \sum_{r=1}^{k}\sigma_r X(t)d^*\xi_r(t). \tag{B.13}$$

Recall that the SDE (B.13) is equivalent to the following SDE in the Itô form (see (5.30))

$$dX(t) = \left(\sigma_0 + \frac{1}{2}\sum_{r=1}^{k}\sigma_r^2\right)X(t)\,dt + \sum_{r=1}^{k}\sigma_r X(t)\,d\xi_r(t). \tag{B.14}$$

The process $\Lambda^\lambda(t) = X^x(t)/|X^x(t)|$, $\lambda = x/|x|$, for X from (B.13) is a diffusion process on $\mathbf{S}^{d-1}$ satisfying the Stratonovich system

$$d\Lambda = h_0(\Lambda)dt + \sum_{r=1}^{k}h_r(\Lambda)d^*\xi_r(t), \quad \Lambda(0) = \lambda = \frac{x}{|x|}, \tag{B.15}$$

where

$$h_r(\lambda) = \sigma_r\lambda - (\sigma_r\lambda, \lambda)\lambda, \quad r = 0, 1, \ldots, k. \tag{B.16}$$

The process $|X^x(t)|$ for X from (B.13) satisfies the equation

$$d|X| = (\sigma_0 \Lambda, \Lambda) \cdot |X| \, dt + \sum_{r=1}^{k} \sigma_r(\Lambda, \Lambda) \cdot |X| d^* \xi_r(t), \tag{B.17}$$

$$\Lambda = \Lambda^\lambda(t), \quad \lambda = x/|x|, \ |X(0)| = |x|.$$

We note that we remain the same notation for the Itô and Stratonovich X, Λ, ρ, because this does not cause any confusion. Besides, they are the same if we lay $\sigma_0 = B - \frac{1}{2} \sum_{r=1}^{k} \sigma_r^2$.

It can easily be proved that

$$|X^x(t)| = |x| \exp\left\{ \int_0^t (\sigma_0 \Lambda, \Lambda) \, ds + \sum_{r=1}^{k} \int_0^t (\sigma_r \Lambda, \Lambda) \, d^* \xi_r(s) \right\}$$

$$= |x| \exp\left\{ \int_0^t (\sigma_0 \Lambda, \Lambda) \, ds \right.$$

$$\left. + \sum_{r=1}^{k} \int_0^t \left(\frac{1}{2}(\sigma_r + \sigma_r^\top)(\sigma_r \Lambda, \Lambda) - (\sigma_r \Lambda, \Lambda)^2 \right) ds \right\}$$

$$\times \exp\left\{ \sum_{r=1}^{k} \int_0^t (\sigma_r \Lambda, \Lambda) \, d\xi_r(s) \right\}, \tag{B.18}$$

where the last integral is of the Itô form.

The vector fields $h_r(\lambda)$, $r = 0, 1, \ldots, k$, are tangent to the sphere $\mathbf{S}^{d-1}$. They are the orthogonal projections of the linear vector fields $\sigma_r x$ onto $\mathbf{S}^{d-1}$. The sphere $\mathbf{S}^{d-1}$ remains invariant for the Stratonovich equation (B.15). It is clear from the equation

$$d(1 - (\Lambda, \Lambda)) = -2(1 - (\Lambda, \Lambda)) \left[(\sigma_0 \Lambda, \Lambda) \, dt + \sum_{r=1}^{k} (\sigma_r \Lambda, \Lambda) \, d^* \xi_r(t) \right].$$

Let us consider the Itô equation

$$d\Lambda = h_0(\Lambda) \, dt + \sum_{r=1}^{k} h_r(\Lambda) \, d\xi_r(t), \qquad \Lambda(0) = \lambda = \frac{x}{|x|}, \tag{B.19}$$

with the same vector fields $h_r(\lambda)$, $r = 0, 1, \ldots, k$, being tangent to the sphere $\mathbf{S}^{d-1}$. We have for Λ satisfying (B.19)

$$d(1 - (\Lambda, \Lambda)) = -2(1 - (\Lambda, \Lambda)) \left[(\sigma_0 \Lambda, \Lambda) \, dt + \sum_{r=1}^{k} (\sigma_r \Lambda, \Lambda) \, d\xi_r(t) \right]$$

$$-\sum_{r=1}^{k}(h_r(\Lambda), h_r(\Lambda))\,dt.$$

From here, the sphere $\mathbf{S}^{d-1}$ is as a rule not invariant for (B.19) though the vector fields $h_r(\lambda)$, $r = 0, 1, \ldots, k$, are tangent to $\mathbf{S}^{d-1}$. We see that the behavior of solutions to the Stratonovich equation concerning invariance of a surface more closely corresponds to our intuition.

B.2   Basic Theorems

B.2.1   Nondegeneracy Conditions

The first nondegeneracy condition for (B.1) was introduced in Chap. 6 (see (6.62)). It is not difficult to show that (6.62) is equivalent to the condition

$$\dim L(\sigma_1 x, \ldots, \sigma_k x) = d, \quad x \in \mathbb{R}^d,\ x \neq 0, \tag{B.20}$$

where L denotes the linear hull spanned by the given vector fields. It can be proved that the weaker nondegeneracy condition (6.67) is equivalent to

$$\dim L(h_1(\lambda), \ldots, h_k(\lambda)) = d - 1, \quad \lambda \in \mathbf{S}^{d-1}. \tag{B.21}$$

We shall use the more weak Lie algebra condition (which we shall call condition C), introduced in [8]:

$$\dim LA(h_1(\lambda), \ldots, h_k(\lambda)) = d - 1, \quad \lambda \in \mathbf{S}^{d-1}, \tag{B.22}$$

where LA denotes the Lie algebra generated by the given vector fields.

We recall that the Lie algebra of the vector fields in a domain of $\mathbb{R}^d$ is considered with respect to the commutator:

$$[\alpha, \beta]_i = \sum_{j=1}^{d} \alpha_j \frac{\partial \beta_i}{\partial \lambda_j} - \sum_{j=1}^{d} \beta_j \frac{\partial \alpha_i}{\partial \lambda_j},$$

where $\alpha(\lambda)$ and $\beta(\lambda)$ are some vector fields.

It is known (see [60, Sect. 3.24]) that if two fields $\alpha(\lambda)$ and $\beta(\lambda)$ are tangent to a smooth surface then their commutator is tangent to the surface as well. The commutator $[h_r, h_q]$ of $h_r(\lambda) = \sigma_r \lambda - (\sigma_r \lambda, \lambda)\lambda$ and $h_q(\lambda) = \sigma_q \lambda - (\sigma_q \lambda, \lambda)\lambda$ has the coordinates

$$[h_r, h_q]_i = \sum_{j=1}^{d} (\sigma_r \lambda - (\sigma_r \lambda, \lambda)\lambda)_j \frac{\partial (\sigma_q \lambda - (\sigma_q \lambda, \lambda)\lambda)_i}{\partial \lambda_j}$$

$$- \sum_{j=1}^{d} (\sigma_q \lambda - (\sigma_q \lambda, \lambda)\lambda)_j \frac{\partial(\sigma_r \lambda - (\sigma_r \lambda, \lambda)\lambda)_i}{\partial \lambda_j}$$

$$= ((\sigma_q \sigma_r - \sigma_r \sigma_q)\lambda)_i - ((\sigma_q \sigma_r - \sigma_r \sigma_q)\lambda, \lambda)\lambda_i,$$

i.e., the commutator of two fields from (B.16) which are generated by the matrices σ_r and σ_q is the field determined by the commutator $\sigma_q \sigma_r - \sigma_r \sigma_q$ of the matrices. In other words, the Lie algebra of vector fields in (B.22) is generated by the matrices from the Lie algebra of matrices $LA(\sigma_1, \ldots, \sigma_k)$.

Remark B.1 Let us note that the most results concerning moment Lyapunov exponents are true under the condition

$$\dim LA(h_0(\lambda), h_1(\lambda), \ldots, h_k(\lambda)) = d - 1, \quad \lambda \in \mathbf{S}^{d-1}, \tag{B.23}$$

which is weaker than the condition C (see [10], [8], [19]). But in order to avoid some complications we impose (B.22).

B.2.2 Semigroups of Positive Compact Operators and Moment Lyapunov Exponents

Let us introduce the following semigroup of positive operators $T_t(p)$ on $\mathbf{C}(\mathbf{S}^{d-1})$, $p \in \mathbb{R}$:

$$T_t(p)f(\lambda) = Ef(\Lambda^\lambda(t))|X^\lambda(t)|^p, \quad \lambda \in \mathbf{S}^{d-1}, \ f \in \mathbf{C}(\mathbf{S}^{d-1}), \ p \in \mathbb{R}. \tag{B.24}$$

The semigroup property flows out the calculation (we omit here the parameter p at the notation $T_t(p)$):

$$T_{t+s}f(\lambda) = \mathbf{E}\{f(\Lambda^\lambda(t+s))|X^\lambda(t+s)|^p\}$$

$$= \mathbf{E}\{f(\Lambda^{t,\Lambda^\lambda(t)}(t+s))|X^{t,X^\lambda(t)}(t+s)|^p\}$$

$$= \mathbf{E}\{f(\Lambda^{t,\Lambda^\lambda(t)}(t+s))|X^{t,\Lambda^\lambda(t)}(t+s)|^p\}|X^\lambda(t)|^p\}$$

$$= \mathbf{E}\{\mathbf{E}[f(\Lambda^{t,\Lambda^\lambda(t)}(t+s))|X^{t,\Lambda^\lambda(t)}(t+s)|^p\|\mathcal{N}_t]|X^\lambda(t)|^p\}$$

$$= \mathbf{E}\{T_s f(\Lambda^\lambda(t))|X^\lambda(t)|^p\} = T_t T_s f(\lambda).$$

It is not difficult to show that this semigroup is strongly continuous. Recall that a positive operator Q on $\mathbf{C}(\mathbf{S}^{d-1})$ ($T_t(p)$, $t > 0$, is evidently positive) is called strongly positive if $Qf(\lambda) > 0$, $\lambda \in \mathbf{S}^{d-1}$, for any nontrivial $f \geq 0$.

Theorem B.1 *Let the condition* (C):

$$\dim LA(h_1(\lambda), \ldots, h_k(\lambda)) = d - 1, \quad \lambda \in \mathbf{S}^{d-1}, \tag{B.25}$$

be fulfilled. Then any operator $T_t(p)$, $t > 0$, $p \in \mathbb{R}$, is strongly positive and compact. Further, for each $p \in \mathbb{R}$ and $t > 0$ the generator $L(p)$ of $T_t(p)$ has a strictly positive eigenfunction $e(\lambda; p) > 0$ corresponding to the principal eigenvalue $g(p)$ which is real, simple and strictly dominates the real part of any other point of the spectrum of $L(p)$:

$$L(p)e(\lambda; p) = g(p)e(\lambda; p). \tag{B.26}$$

The eigenfunction for $T_t(p)$ is the same $e(\lambda; p)$ with the eigenvalue $e^{g(p)t}$:

$$T_t(p)e(\lambda; p) = e^{g(p)t}e(\lambda; p). \tag{B.27}$$

Proof At the beginning let us suppose that the condition

$$\dim LA(\sigma_1 x, \ldots, \sigma_k x) = d, \quad x \neq 0, \tag{B.28}$$

is fulfilled. $LA(\sigma_1 x, \ldots, \sigma_k x)$ consists of the linear vector fields which matrices belong to $LA(\sigma_1, \ldots, \sigma_k)$. The vector fields from $LA(h_1(\lambda), \ldots, h_k(\lambda))$ are the orthogonal projections of the linear vector fields from $LA(\sigma_1 x, \ldots, \sigma_k x)$ onto the plane which is tangent to $\mathbf{S}^{d-1}$ at the point $\lambda = x/|x|$. From here it is not difficult to prove that (B.28) implies (B.25), i.e., (B.28) is stronger than (B.25). It follows from (B.28) due to Hörmander's theorem (see [238, p. 253]) that the transition probability function $P(t, x, A)$ of the Markov process defined by (B.13) possesses a density $p(t, x, y)$ which is strictly positive and continuous with respect to $t > 0$, $x \neq 0$, $y \neq 0$. Let us note in passing that under condition (B.20), which is equivalent to (6.62) and which is stronger than (B.28), this fact follows from the simplest variant of the Hörmander theorem, from the Weyl lemma (see [238, p. 254]). We have

$$T_t(p)f(\lambda) = \mathbf{E}\{f(\Lambda^\lambda(t))|X^\lambda(t)|^p\} = \mathbf{E}\left\{f\left(\frac{X^\lambda(t)}{|X^\lambda(t)|}\right)|X^\lambda(t)|^p\right\}$$

$$= \int p(\lambda, t, y)f(y/|y|)|y|^p \, dy =: \varphi(\lambda), \quad \lambda \in \mathbf{S}^{d-1}.$$

We see from here that $T_t(p)$ is strongly positive. Let us prove compactness of $T_t(p)$: $\mathbf{C}(\mathbf{S}^{d-1}) \to \mathbf{C}(\mathbf{S}^{d-1})$. To this aim we should prove that the image of the unit ball from $\mathbf{C}(\mathbf{S}^{d-1})$ is uniformly bounded (this is evident) and equicontinuous. For arbitrary $\varepsilon > 0$, let $c > 0$ and $C > 0$ be such that for fixed $t > 0$ and $p \in \mathbb{R}$ for all $f \in \mathbf{C}(\mathbf{S}^{d-1})$ with $\|f\| \leq 1$ and for all $\lambda \in \mathbf{S}^{d-1}$

$$\int_{\{y:|y|<c\}\cup\{y:|y|>C\}} p(\lambda, t, y)f(y/|y|)|y|^p dy \leq \frac{\varepsilon}{3}.$$

For $p \geq 0$, let $\delta > 0$ be such that

$$|p(\lambda'', t, y) - p(\lambda', t, y)| \leq \frac{\varepsilon}{3C^p} \quad \text{if } |\lambda'' - \lambda'| \leq \delta, \ c \leq |y| \leq C.$$

Then for any f with $\| f \| \leq 1$ we obtain

$$|\varphi(\lambda'') - \varphi(\lambda')| \leq \frac{2\varepsilon}{3} + \int_{c \leq |y| \leq C} |p(\lambda'', t, y) - p(\lambda', t, y)| |f(y/|y|)| |y|^p \, dy \leq \varepsilon.$$

If $p < 0$ then we choose $\delta > 0$ such that

$$|p(\lambda'', t, y) - p(\lambda', t, y)| \leq \frac{\varepsilon c^p}{3} \quad \text{if } |\lambda'' - \lambda'| \leq \delta, \ c \leq |y| \leq C.$$

Thus, the equicontinuity is proved, hence in the case (B.28) any operator $T_t(p)$, $t > 0$, is strongly positive and compact.

Let us return to the condition (B.25). Consider the vector field $\alpha\lambda$, $\alpha \neq 0$. It is orthogonal to any field $h_j(\lambda)$, $j = 1, \ldots, k$. Therefore (B.25) implies

$$\dim LA(h_1(\lambda), \ldots, h_k(\lambda), \alpha\lambda) = d, \quad \alpha \neq 0, \ \lambda \in \mathbf{S}^{d-1}. \tag{B.29}$$

Since αx commutates with every $\sigma_j x$, $j = 1, \ldots, k$, we get

$$\dim LA(\sigma_1 x, \ldots, \sigma_k x, \alpha x) = \dim L(LA(\sigma_1 x, \ldots, \sigma_k x), \alpha x) = d, \quad x \neq 0.$$

Thus from (B.25) it follows

$$\dim LA(\sigma_1 x, \ldots, \sigma_k x, \sigma_{k+1} x) = d, \quad \sigma_{k+1} = \alpha J, \ x \neq 0, \tag{B.30}$$

where J is identity matrix.

Together with (B.13) consider the system

$$d\widetilde{X}(t) = \sigma_0 \widetilde{X}(t) \, dt + \sum_{r=1}^{k+1} \sigma_r \widetilde{X}(t) \, d^*\xi_r(t). \tag{B.31}$$

Due to (B.30) every $\widetilde{T}_t(p)$, $t > 0$, is strongly positive and compact. We have $h_{k+1}(\lambda) = \sigma_{k+1}\lambda - (\sigma_{k+1}\lambda, \lambda)\lambda = 0$. From here $\widetilde{\Lambda} = \Lambda$. Therefore (see (B.18))

$$|\widetilde{X}^\lambda(t)|^p = |X^\lambda(t)|^p \exp\left\{ p \int_0^t (\sigma_{k+1}\lambda, \lambda) d^*\xi_{k+1} \right\} = |X^\lambda(t)|^p \exp\{\alpha p \xi_{k+1}(t)\},$$

$$\widetilde{T}_t(p) f(\lambda) = \mathbf{E}\{f(\widetilde{\Lambda}^\lambda(t))|\widetilde{X}^\lambda(t)|^p\} = \mathbf{E}\{f(\Lambda^\lambda(t))|X^\lambda(t)|^p \exp\{\alpha p \xi_{k+1}(t)\}\}$$

$$= T_t(p) f(\lambda) \exp\left\{ \frac{1}{2}\alpha^2 p^2 t \right\},$$

i.e.,

$$T_t(p) f(\lambda) = \exp\left\{ -\frac{1}{2}\alpha^2 p^2 t \right\} \widetilde{T}_t(p) f(\lambda),$$

hence $T_t(p)$ for any $t > 0$ is strongly positive and compact.

The other claims of theorem follow from the Perron-Frobenius theory for positive semigroups (see [214], [49]). Theorem is proved.   □

Remark B.2 Under the condition (C) (see (B.25)) the process $\Lambda(t)$ is ergodic with the unique invariant measure $\nu(d\lambda)$ and the formula

$$\lim_{t\to\infty} \frac{1}{t} \ln|X^x(t)| = a^* = \int_{\mathbf{S}^{d-1}} Q(\lambda)\nu(d\lambda) \tag{B.32}$$

is true.

This assertion follows from the fact that the condition (B.25) just as the more strong condition (6.62) ensures positiveness of the transition probability.

Theorem B.2 *Let the condition (C) be fulfilled. Then for any $x \neq 0$*

$$\lim_{t\to\infty} \frac{1}{t} \ln\mathbf{E}|X^x(t)|^p = g(p). \tag{B.33}$$

Proof Let $L(p)f(\lambda) = g(p)f(\lambda)$, $f(\lambda) > 0$, $\lambda \in \mathbf{S}^{d-1}$. Hence $T_t(p)f = e^{tg(p)}f$ and there exist positive constants $0 < k < K < \infty$ such that $0 < k \leq f(\lambda) \leq K < \infty$. We have

$$\lim_{t\to\infty} \frac{1}{t} \ln\mathbf{E}|X^x(t)|^p = \lim_{t\to\infty} \frac{1}{t} \ln(\mathbf{E}|X^\lambda(t)|^p|x|^p) = \lim_{t\to\infty} \frac{1}{t} \ln\mathbf{E}|X^\lambda(t)|^p,$$

$$\lambda = \frac{x}{|x|}.$$

Further,

$$k\mathbf{E}|X^\lambda(t)|^p \leq \mathbf{E}\{f(\Lambda^\lambda(t))|X^\lambda(t)|^p\} \leq K\mathbf{E}|X^\lambda(t)|^p.$$

Therefore

$$\frac{1}{K}\mathbf{E}\{f(\Lambda^\lambda(t))|X^\lambda(t)|^p\} \leq \mathbf{E}|X^\lambda(t)|^p \leq \frac{1}{k}\mathbf{E}\{f(\Lambda^\lambda(t))|X^\lambda(t)|^p\}.$$

Hence

$$\lim_{t\to\infty} \frac{1}{t} \ln\mathbf{E}|X^x(t)|^p = \lim_{t\to\infty} \frac{1}{t} \ln\mathbf{E}\{f(\Lambda^\lambda(t))|X^x(t)|^p\}$$

$$= \lim_{t\to\infty} \frac{1}{t} \ln T_t(p)f(\lambda) = \lim_{t\to\infty} \frac{1}{t}(tg(p) + \ln f(\lambda)) = g(p).$$

Theorem is proved. □

The mean $\mathbf{E}|X^x(t)|^p$ is connected with the characteristic function ψ of $\xi = \ln|X^x(t)|$ which can be considered as a function of the complex variable z:

$$\psi(z) = \psi(z; t, x) := \mathbf{E}\exp\{iz\xi\} = \mathbf{E}\exp\{iz\ln|X^x(t)|\}.$$

If $z = p$ is real, we get the classical characteristic function of the random variable $\xi = \ln|X^x(t)|$:

$$\psi(p) = \psi(p; t, x) = \mathbf{E}\exp\{ip\ln|X^x(t)|\} = \mathbf{E}|X^x(t)|^{ip}, \quad |x| \neq 0. \tag{B.34}$$

If $z = -ip$ is purely imaginary, we get

$$\psi(-ip) = \psi(-ip; t, x) = \mathbf{E}\exp\{p \ln|X^x(t)|\} = \mathbf{E}|X^x(t)|^p, \quad |x| \neq 0,$$

i.e., the p-th moment of the random variable $|X^x(t)|$. Thus the function $\psi(z)$ is an extension of the function $\mathbf{E}|X^x(t)|^p$.

Clearly, $\psi(-ip)$, $-\infty < p < \infty$, takes positive values and

$$|\psi(q + ip)| \leq \psi(ip). \tag{B.35}$$

Many facts about analytic characteristic functions can be found in [184, 186, 236].

There exists $\psi'(z)$, i.e., for every $t, x \neq 0$ the function $\psi(z)$ is entire. Existence of the derivative $\psi'(z) = \psi'_z(z; t, \lambda)$ with respect to z and the equality $\psi'_z(z; t, \lambda) = i\mathbf{E}(\ln|X^\lambda(t)|\exp\{iz \ln|X^\lambda(t)|\})$ can be proved in the standard way by differentiation under the sign of expectation.

Introduce the strongly continuous semigroup $T_t(z)$ on the Banach space $\mathbf{C}(\mathbf{S}^{d-1})$ of complex-valued functions $f(\lambda)$, $\lambda \in \mathbf{S}^{d-1}$:

$$T_t(z)f(\lambda) = \mathbf{E}[f(\Lambda^\lambda(t))\exp\{z \ln|X^\lambda(t)|\}].$$

Let $L(z)$ be the generator of the semigroup $T_t(z)$. For $z = p$ real, some properties of $T_t(p)$ were mentioned above. Not all of them are fulfilled for arbitrary complex z. For example, the property of positivity is violated. At the same time many of them remain true. In particular, the operator $T_t(z)$ for any z and $t > 0$ is compact as well.

Clearly,

$$\psi(z; t, \lambda) = T_t(iz)\mathbf{1}(\lambda) = \mathbf{E}\exp\{iz \ln|X^\lambda(t)|\},$$

where the function $\mathbf{1}(\lambda)$ is identically equal to 1.

Let us fix $t = 1$ and consider the family $T_1(z)$ for z belonging to a sufficiently small neighborhood of the point $z = p$. This family analytically depends on z (see [115]). The operator $T_1(p)$ has $\exp(g(p))$ as an eigenvalue with the eigenfunction $e(\lambda; p)$: $T_1(p)e(\lambda; p) = \exp(g(p))e(\lambda; p)$. From Theorem B.1 we have that the eigenvalue $g(p)$ of $L(p)$ is simple and $g(p)$ strictly dominates the real part of any other point of the spectrum of $L(p)$. Since $T_1(z)$ and $L(z)$ analytically depend on z, the spectrum of $L(z)$ for sufficiently small $(z - p)$ contains an isolated eigenvalue, which is close to $g(p)$. And both the eigenvalue and a corresponding eigenvector $e(\lambda; z)$ depend on z analytically (see [115]).

Thus, the following result is obtained.

Theorem B.3 *Let the condition* (C) *(see* (B.25)) *be fulfilled. Then the function* $g: \mathbb{R} \to \mathbb{R}$ *is analytic.*

Theorem B.4 *Let the condition* (C) *be fulfilled. Then the function* $g: \mathbb{R} \to \mathbb{R}$ *is convex,* $g(p)/p$ *is increasing,* $g(0) = 0$, *and* $a^* = g'(0)$.

Proof Consider the function $g(p; t, x) := \frac{1}{t} \ln \mathbf{E}|X^x(t)|^p$. Due to the Cauchy–Schwarz inequality, we have

$$\mathbf{E}|X^x(t)|^{(p+q)/2} \leq (\mathbf{E}|X^x(t)|^p)^{1/2}(\mathbf{E}|X^x(t)|^q)^{1/2},$$

whence

$$g\left(\frac{p+q}{2}; t, x\right) \leq \frac{1}{2}g(p; t, x) + \frac{1}{2}g(q; t, x),$$

i.e., $g(p; t, x)$ is convex. Taking the lim we prove that $g(p)$ is convex.

For $p \geq 1$, $q \geq p$, the inequality

$$(\mathbf{E}|X^x(t)|^p)^{1/p} \leq (\mathbf{E}|X^x(t)|^q)^{1/q}, \qquad (B.36)$$

i.e., increase of the function $(\mathbf{E}|X^x(t)|^p)^{1/p}$, is well known (it follows from the Hölder inequality). If $0 < p < 1$, $q \geq p$, let us select $r > 0$ such that $pr > 1$. Denote $|Y(t)| = |X^x(t)|^{1/r}$. We get

$$(\mathbf{E}|Y(t)|^{pr})^{1/pr} \leq (\mathbf{E}|Y(t)|^{qr})^{1/qr},$$

whence (B.36) follows for $p > 0$, $q \geq p$. For $p < 0$, we have

$$(\mathbf{E}|X^x(t)|^p)^{1/p} = \frac{1}{(\mathbf{E}(\frac{1}{|X^x(t)|})^{-p})^{-1/p}}.$$

If the positive parameter $-p$ decreases (it is the same as p increases) then the denominator decreases as well, hence the fraction, i.e., $(\mathbf{E}|X^x(t)|^p)^{1/p}$ is increasing for $p < 0$. From this, increase of $g(p)/p$ on $(-\infty, 0) \cup (0, \infty)$ follows. Because $\lim_{p \to 0} g(p)/p = g'(0)$ exists the increase holds on $(-\infty, \infty)$.

Due to Jensen's inequality, for any $-\infty < p < \infty$

$$\frac{p}{t}\mathbf{E}\ln|X^x(t)| \leq \frac{1}{t}\ln\mathbf{E}|X^x(t)|^p.$$

From (B.8)

$$\mathbf{E}\rho(t) = \mathbf{E}\ln|X^x(t)| = \ln|x| + \int_0^t \mathbf{E}Q(\Lambda^\lambda(s))ds, \quad \lambda = \frac{x}{|x|}.$$

These two relations give for $p > 0$

$$\frac{g(-p)}{-p} \leq \lim_{t \to \infty} \frac{1}{t}\int_0^t \mathbf{E}Q(\Lambda^\lambda(s))\,ds = a^* \leq \frac{g(p)}{p},$$

whence $a^* = g'(0)$ follows. $\square$

Together with the characteristic function $\psi(z) = \psi(z; t, x) = \mathbf{E}\exp\{iz \ln|X^x(t)|\}$ let us consider the moments $m_n = m_n(t, x)$ and semi-invariants (cumulants) $\gamma_n =$

$\gamma_n(t, x)$ of $\xi = \ln|X^x(t)|$. The moments can be expressed in terms of the coefficients of the Taylor-series expansion for $\psi(z) = \psi(z; t, x)$:

$$\psi(z) = \sum_{n=0}^{\infty} \frac{\psi^{(n)}(0)}{n!} z^n, \qquad m_n(t, x) = m_n := \mathbf{E}\xi^n = \frac{1}{i^n} \psi^{(n)}(0).$$

The cumulants are equal to (in taking the logarithm $\ln \psi(z)$ we choose a branch such that $\ln \psi(z) = 0$ at $z = 0$):

$$\gamma_n(t, x) := \gamma_n = \frac{1}{i^n} \frac{d^n}{dz^n} \ln \psi(z)\Big|_{z=0}$$

$$= \frac{d^n}{dp^n} \ln \psi(-ip)\Big|_{p=0} = \frac{d^n}{dp^n} \ln \mathbf{E}|X^x(t)|^p\Big|_{p=0}.$$

This notion is well-defined since for every $t, x \neq 0$ there exists a sufficiently small neighborhood $|z| < \delta$ (where δ in general depends on t, x) in which the function $\psi(z) = \psi(z; t, x)$ does not vanish. We have (since $\gamma_0 = 0$)

$$\ln \psi(z) = \sum_{n=1}^{\infty} \frac{i^n \gamma_n}{n!} z^n, \quad |z| < \delta, \tag{B.37}$$

$$\ln \mathbf{E}|X^x(t)|^p = \sum_{n=1}^{\infty} \frac{\gamma_n(t, x)}{n!} p^n, \quad |p| < \delta. \tag{B.38}$$

The relation for the moments and semi-invariants can be obtained in the following well known way. Put $iz = w$ and $\varphi(w) = 1 + \sum_{n=1}^{\infty} (m_n/n!)w^n$. Then $\ln \varphi(w) = \sum_{n=1}^{\infty} (\gamma_n/n!)w^n$. We have $\varphi \cdot (\ln \varphi)' = \varphi'$. This is equivalent to

$$\left(1 + \sum_{n=1}^{\infty} \frac{m_n}{n!} w^n\right) \cdot \sum_{n=1}^{\infty} \frac{\gamma_n}{(n-1)!} w^{n-1} = \sum_{n=1}^{\infty} \frac{m_n}{(n-1)!} w^{n-1}. \tag{B.39}$$

Putting $w = 0$ in (B.39), we find $\gamma_1 = m_1$. Subsequently differentiating (B.39) with respect to w and putting $w = 0$, we obtain $\gamma_2 = m_2 - m_1^2$, $\gamma_3 = m_3 - 3m_1 m_2 + 2m_1^3$, $\gamma_4 = m_4 - 4m_1 m_3 - 3m_2^2 + 12m_1^2 m_2 - 6m_1^4$, and so on.

Let us note a remarkable feature of semi-invariants: any semi-invariant of a sum of independent random variables is equal to the sum of the semi-invariants of these variables. In contrast, the second and higher moments do not possess this property.

Using Theorem B.3 and (B.38), we get (we recall that $g(0) = 0$)

$$\sum_{n=1}^{\infty} \frac{g^{(n)}(0)}{n!} p^n = g(p) = \lim_{t \to \infty} \frac{1}{t} \ln \mathbf{E}|X^x(t)|^p = \lim_{t \to \infty} \sum_{n=1}^{\infty} \frac{\gamma_n(t, x)/t}{n!} p^n,$$

whence the following conjecture arises

$$\lim_{t \to \infty} \frac{\gamma_n(t, x)}{t} = g^{(n)}(0), \quad n = 1, 2, \ldots. \tag{B.40}$$

A rigorous proof of (B.40) is given in [201]. For $n = 1$ the first semi-invariant γ_1 is equal to $\mathbf{E} \ln |X(t, x)|$ and the connection between the Lyapunov exponent a^* and $g'(0)$,

$$a^* = \lim_{t \to \infty} \frac{\mathbf{E} \ln |X^x(t)|}{t} = g'(0), \tag{B.41}$$

confirms this assertion. The second semi-invariant is equal to $\gamma_2(t, x) = \mathbf{E} \ln^2 |X^x(t)| - (\mathbf{E} \ln |X^x(t)|)^2 = \text{var} \ln |X^x(t)|$. The equality (B.40) for $n = 2$ can be proved without any serious difficulties (see [201]). To prove (B.40) in the general case, some properties of the analytic characteristic function $\psi(z; t, x) = \mathbf{E} \exp\{iz \ln |X^x(t)|\}$ are studied. Since $\psi(z; t, x)$ takes, as a rule, zero values if $d > 1$, the cumulant generating function $\ln \psi(z; t, x)$ is not defined everywhere. At the same time for every $t \geq 0$ there exists $\delta_t > 0$ such that for any $x \in \mathbf{R}^d$ the function $\ln \psi(z; t, x)$ is analytic in $C_{\delta_t} := \{z : |z| < \delta_t\}$. Much more complicated assertion consists in the fact that there exists such $\delta > 0$ which is independent of t. Moreover, under the nondegeneracy condition (C) there exists C_δ such that the function $(\ln \psi(z; t, x))/t$ is analytic in C_δ and uniformly bounded with respect to $t > 0$ and x with $|x| = 1$. Due to this fact, it is possible to use the classical Vitali convergence theorem (see [261]) and prove the basic result (B.40). The following theorem is proved in [201]:

Theorem B.5 *Let the condition (C) be fulfilled. Then (B.40) is true. The moment Lyapunov exponent $g(p)$ can be extended to complex z belonging to a circle $C_\delta = \{z : |z| < \delta\}$ in the sense that for such z*

$$\lim_{t \to \infty} \frac{1}{t} \ln \psi(z; t, \lambda) = \lim_{t \to \infty} \frac{1}{t} \ln \mathbf{E} \exp\{iz \ln |X^\lambda(t)|\} = g(iz) \tag{B.42}$$

is fulfilled. The limits in (B.40) and (B.42) do not depend on $x \in \mathbf{R}^d$.

Remark B.3 Let the condition (C) be fulfilled. Theorem 6.14 claims that

$$\mathbf{P}\left\{ \frac{\ln |X^x(t)| - a^* t}{\sqrt{\text{var} \ln |X^x(t)|}} < x \right\} \underset{t \to \infty}{\longrightarrow} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy \tag{B.43}$$

provided that

$$\text{var} \ln |X^x(t)| \to \infty \quad \text{as } t \to \infty. \tag{B.44}$$

From (B.40), if $n = 2$, we get

$$\lim_{t \to \infty} \frac{\text{var} \ln |X^x(t)|}{t} = \lim_{t \to \infty} \frac{\mathbf{E} \ln^2 |X^x(t)| - (\mathbf{E} \ln |X^x(t)|)^2}{t} = g''(0). \tag{B.45}$$

Therefore (B.44) is fulfilled if $g''(0) \neq 0$ (i.e., $g''(0) > 0$). So we have arrived at the central limit theorem at the form given in [19]:

$$\mathbf{P}\left\{ \frac{\ln |X^x(t)| - a^* t}{\sqrt{g''(0)t}} < x \right\} \underset{t \to \infty}{\longrightarrow} \Phi(x). \tag{B.46}$$

It is observed in [19] that $g''(0) = 0$ if and only if $g(p) = pa^*$ for all $p \in \mathbb{R}$ which happens if and only if there exists an invertible matrix G such that all the matrices $G(A - \lambda^* I)G^{-1}, G\sigma_1 G^{-1}, \ldots, G\sigma_k G^{-1}$ are skew-symmetric.

B.2.3 Generator of the Process Λ

A function $\varphi(M)$ on $\mathbf{S}^{d-1}$ is a function of $d - 1$ variables. It can be defined by various ways. Specifically, any function $\varphi(M)$, $M \in \mathbf{S}^{d-1}$, can be considered as a trace of some function of d variables defined in a neighborhood of $\mathbf{S}^{d-1}$. Clearly, the last function of d variables is not defined uniquely. But this nonuniqueness created no problems because we are interested either in values $\varphi(M)$ for $M \in \mathbf{S}^{d-1}$ or in derivatives of φ along vector fields which are tangent to $\mathbf{S}^{d-1}$. In particular, we note that the considered vector fields $h_i(\lambda) = \sigma_i \lambda - (\sigma_i \lambda, \lambda)\lambda$ are the trace functions of the functions $\sigma_i x - (\sigma_i x, x)x$.

We consider the process $\Lambda^\lambda(t)$ governed by the SDE (see (B.6), (B.7))

$$d\Lambda = q(\Lambda)\,dt + \sum_{r=1}^{k} h_r(\Lambda)\,d\xi_r(t). \tag{B.47}$$

The sphere

$$\mathbf{S}^{d-1} : \lambda_1^2 + \cdots + \lambda_d^2 = 1 \tag{B.48}$$

is invariant for the process Λ.

Let $\Lambda_{\mathbf{S}^{d-1}}(t)$ be the process Λ on the sphere $\mathbf{S}^{d-1}$ and $L_{\mathbf{S}^{d-1}}$ be the generator of the process.

Theorem B.6 *Let $f_{\mathbf{S}^{d-1}} \in \mathbf{C}_2(\mathbf{S}^{d-1})$. Let $f \in \mathbf{C}_2(\mathbb{R}^d)$ have a compact support and be an extension of $f_{\mathbf{S}^{d-1}}$ (such a function f is not unique and $f_{\mathbf{S}^{d-1}}$ is the trace of f):*

$$f_{\mathbf{S}^{d-1}}(P) = f(P), \quad P \in \mathbf{S}^{d-1}. \tag{B.49}$$

Then

$$L_{\mathbf{S}^{d-1}} f_{\mathbf{S}^{d-1}}(P) = Lf(P), \quad P \in \mathbf{S}^{d-1}, \tag{B.50}$$

where

$$L = \frac{1}{2} \sum_{i,j=1}^{d} \sum_{r=1}^{k} h_{ri} h_{rj} \frac{\partial^2}{\partial \lambda_i \partial \lambda_j} + \sum_{i=1}^{d} q_i \frac{\partial}{\partial \lambda_i} \tag{B.51}$$

is the generator of the process $\Lambda(t)$.

Proof For definiteness, let (B.48) be resolvable with respect to λ_d in some piece $\mathbf{S}_0^{d-1}$ of $\mathbf{S}^{d-1}$:

$$\lambda_d = \sqrt{1 - \lambda_1^2 - \cdots - \lambda_{d-1}^2} =: \varphi(\lambda_1, \ldots, \lambda_{d-1}). \tag{B.52}$$

We have in $\mathbf{S}_0^{d-1}$:

$$d\Lambda_i = q_i(\Lambda_1, \ldots, \Lambda_{d-1}, \varphi(\Lambda_1, \ldots, \Lambda_{d-1}))\, dt$$
$$+ \sum_{r=1}^{k} h_{ri}(\Lambda_1, \ldots, \Lambda_{d-1}, \varphi(\Lambda_1, \ldots, \Lambda_{d-1}))\, d\xi_r(t), \quad i = 1, \ldots, d.$$

$$\tag{B.53}$$

Itô's formula gives

$$d\Lambda_d = \sum_{i=1}^{d-1} \frac{\partial \varphi}{\partial \lambda_i} \left(q_i\, dt + \sum_{r=1}^{k} h_{ri} d\xi_r(t) \right) + \frac{1}{2} \sum_{i,j=1}^{d-1} \sum_{r=1}^{k} h_{ri} h_{rj} \frac{\partial^2 \varphi}{\partial \lambda_i \partial \lambda_j}\, dt, \tag{B.54}$$

where the arguments are the same as in (B.53). Comparing the last equality from (B.53) with (B.54), we get

$$\sum_{i=1}^{d-1} \frac{\partial \varphi}{\partial \lambda_i}(\lambda_1, \ldots, \lambda_{d-1}) h_{ri}(\lambda_1, \ldots, \lambda_{d-1}, \varphi(\lambda_1, \ldots, \lambda_{d-1}))$$

$$= h_{rd}(\lambda_1, \ldots, \lambda_{d-1}, \varphi(\lambda_1, \ldots, \lambda_{d-1})) \tag{B.55}$$

and (with the same arguments as in (B.55))

$$\sum_{i=1}^{d-1} \frac{\partial \varphi}{\partial \lambda_i} q_i + \frac{1}{2} \sum_{i,j=1}^{d-1} \sum_{r=1}^{k} h_{ri} h_{rj} \frac{\partial^2 \varphi}{\partial \lambda_i \partial \lambda_j} = q_d \tag{B.56}$$

for the points from $\mathbf{S}_0^{d-1}$. Due to (B.52), we get

$$f_{\mathbf{S}^{d-1}}(P) = f(\lambda_1, \ldots, \lambda_{d-1}, \varphi(\lambda_1, \ldots, \lambda_{d-1}))$$

$$= f_{\mathbf{S}^{d-1}}(\lambda_1, \ldots, \lambda_{d-1}), \quad P \in \mathbf{S}_0^{d-1}. \tag{B.57}$$

We have for $P \in \mathbf{S}_0^{d-1}$ (see the first $d-1$ equations from (B.53) and then (B.57), (B.55), and (B.56))

$$L_{\mathbf{S}^{d-1}} f_{\mathbf{S}^{d-1}}(P) = L_{\mathbf{S}^{d-1}} f_{\mathbf{S}^{d-1}}(\lambda_1, \ldots, \lambda_{d-1})$$

$$= \sum_{i=1}^{d-1} q_i \frac{\partial f_{\mathbf{S}^{d-1}}}{\partial \lambda_i} + \frac{1}{2} \sum_{i,j=1}^{d-1} \sum_{r=1}^{k} h_{ri} h_{rj} \frac{\partial^2 f_{\mathbf{S}^{d-1}}}{\partial \lambda_i \partial \lambda_j}$$

$$= \sum_{i=1}^{d-1} q_i \left(\frac{\partial f}{\partial \lambda_i} + \frac{\partial f}{\partial \lambda_d} \frac{\partial \varphi}{\partial \lambda_i} \right)$$

$$+ \frac{1}{2} \sum_{i,j=1}^{d-1} \sum_{r=1}^{k} h_{ri} h_{rj} \left(\frac{\partial^2 f}{\partial \lambda_i \partial \lambda_j} + \frac{\partial^2 f}{\partial \lambda_i \partial \lambda_d} \frac{\partial \varphi}{\partial \lambda_j} \right.$$

$$\left. + \frac{\partial^2 f}{\partial \lambda_d \partial \lambda_j} \frac{\partial \varphi}{\partial \lambda_i} + \frac{\partial^2 f}{\partial \lambda_d^2} \frac{\partial \varphi}{\partial \lambda_j} \frac{\partial \varphi}{\partial \lambda_i} + \frac{\partial f}{\partial \lambda_d} \frac{\partial^2 \varphi}{\partial \lambda_i \partial \lambda_j} \right)$$

$$= \sum_{i=1}^{d} q_i \frac{\partial f}{\partial \lambda_i} + \frac{1}{2} \sum_{i,j=1}^{d-1} \sum_{r=1}^{k} h_{ri} h_{rj} \frac{\partial^2 f}{\partial \lambda_i \partial \lambda_j}$$

$$+ \sum_{j=1}^{d-1} \sum_{r=1}^{k} h_{rj} \frac{\partial^2 f}{\partial \lambda_j \partial \lambda_d} \left(\sum_{i=1}^{d-1} h_{ri} \frac{\partial \varphi}{\partial \lambda_i} \right)$$

$$+ \frac{1}{2} \frac{\partial^2 f}{\partial \lambda_d^2} \sum_{r=1}^{k} \left(\sum_{i=1}^{d-1} h_{ri} \frac{\partial \varphi}{\partial \lambda_i} \right) \left(\sum_{j=1}^{d-1} h_{rj} \frac{\partial \varphi}{\partial \lambda_j} \right)$$

$$= \sum_{i=1}^{d} q_i \frac{\partial f}{\partial \lambda_i} + \frac{1}{2} \sum_{i,j=1}^{d-1} \sum_{r=1}^{k} h_{ri} h_{rj} \frac{\partial^2 f}{\partial \lambda_i \partial \lambda_j}$$

$$+ \sum_{j=1}^{d-1} \sum_{r=1}^{k} h_{rj} \frac{\partial^2 f}{\partial \lambda_j \partial \lambda_d} h_{rd} + \frac{1}{2} \frac{\partial^2 f}{\partial \lambda_d^2} \sum_{r=1}^{k} h_{rd}^2$$

$$= \sum_{i=1}^{d} q_i \frac{\partial f}{\partial \lambda_i} + \frac{1}{2} \sum_{i,j=1}^{d} \sum_{r=1}^{k} h_{ri} h_{rj} \frac{\partial^2 f}{\partial \lambda_i \partial \lambda_j}$$

$$= Lf(\lambda_1, \ldots, \lambda_{d-1}, \varphi(\lambda_1, \ldots, \lambda_{d-1})) = Lf(P).$$

Theorem is proved. $\square$

B.2.4 Generator of Semigroup $T_t(p)f(\lambda)$

Let us derive the formula for generator $L(p)$ of the semigroup $T_t(p)f(\lambda)$. Let $f(\lambda)$ be the trace of function $f(x) \in \mathbf{C}_2(\mathbb{R}^d \setminus \{0\})$ with a compact support. We have

$$L(p)f(\lambda) = \lim_{t \to 0} \frac{1}{t}(T_t(p)f(\lambda) - f(\lambda)) = \lim_{t \to 0} \frac{1}{t}(\mathbf{E}\{f(\Lambda^\lambda(t))|X^\lambda(t)|^p\} - f(\lambda))$$

$$= \lim_{t \to 0} \frac{1}{t} \left(\mathbf{E} \left\{ f\left(\frac{X^\lambda(t)}{|X^\lambda(t)|} \right) |X^\lambda(t)|^p \right\} - f(\lambda) \right) = LF(\lambda), \tag{B.58}$$

where L is the generator of the Markov process governed by system (B.1) and

$$F(x) := f(x/|x|)|x|^p. \tag{B.59}$$

Let us calculate $LF(\lambda)$. To this aim we apply Itô's formula to $F(X^\lambda(t))$:

$$dF(X^\lambda(t)) = LF(X^\lambda(t))\,dt + \sum_{r=1}^{k}\left(\frac{\partial F}{\partial x}(X^\lambda(t)), \sigma_r X^\lambda(t)\right)d\xi_r(t). \tag{B.60}$$

Thus, $LF(\lambda)$ is the drift in (B.60) at the time $t = 0$. We get

$$dF(X^\lambda(t))$$

$$= d[f(\Lambda^\lambda(t))|X^\lambda(t)|^p]$$

$$= |X^\lambda(t)|^p\,df(\Lambda^\lambda(t)) + f(\Lambda^\lambda(t))\,d|X^\lambda(t)|^p + df(\Lambda^\lambda(t))\cdot d|X^\lambda(t)|^p$$

$$= |X^\lambda(t)|^p\left(\sum_{i=1}^{d}\frac{\partial f}{\partial x_i}(\Lambda)q_i(\Lambda)\,dt + \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2 f}{\partial x_i \partial x_j}(\Lambda)\sum_{r=1}^{k}h_{ri}(\Lambda)h_{rj}(\Lambda)dt\right)$$

$$+ f(\Lambda)\left(pQ(\Lambda) + \frac{1}{2}p^2\sum_{r=1}^{k}(\sigma_r\Lambda, \Lambda)^2\right)|X|^p\,dt$$

$$+ \left(\sum_{r=1}^{k}\left(\frac{\partial f}{\partial x}(\Lambda), h_r(\Lambda)d\xi_r\right)\right)\cdot p\sum_{r=1}^{k}(\sigma_r\Lambda, \Lambda)|X|^p\,d\xi_r$$

$$+ \sum_{r=1}^{k}\left(\frac{\partial F}{\partial x}(X^\lambda(t)), \sigma_r X^\lambda(t)\right)d\xi_r(t). \tag{B.61}$$

From here $LF(\lambda) = L(p)f(\lambda)$ as the drift at the time $t = 0$ can be found. As a result we obtain the following formula for the generator $L(p)$ of the semigroup $T_t(p)$.

Theorem B.7 *Let $f(\lambda) \in \mathbf{C}_2(\mathbf{S}^{d-1})$. Let $f(x) \in \mathbf{C}_2(\mathbb{R}^d)$ have a compact support and be an extension of $f(\lambda)$. Then*

$$L(p)f(\lambda) = LF(\lambda) = \sum_{i=1}^{d}\frac{\partial f}{\partial x_i}(\lambda)q_i(\lambda) + \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2 f}{\partial x_i \partial x_j}(\lambda)\sum_{r=1}^{k}h_{ri}(\lambda)h_{rj}(\lambda)$$

$$+ p\sum_{r=1}^{k}\left(\sum_{i=1}^{d}\frac{\partial f}{\partial x_i}(\lambda)h_{ri}(\lambda)\right)\cdot(\sigma_r\lambda, \lambda)$$

$$+ \left(pQ(\lambda) + \frac{1}{2}p^2\sum_{r=1}^{k}(\sigma_r\lambda, \lambda)^2\right)f(\lambda). \tag{B.62}$$

Remark B.4 One can write (B.60) and (B.61) for $dF(X^x(t))$, $x \neq 0$. As a result we get the following formula for $LF(x)$ from (B.59)

$$LF(x) = |x|^p L(p) f(\lambda), \quad \lambda = x/|x|, \ x \neq 0. \tag{B.63}$$

We pay attention that $f(x)$ here is a function which is not necessarily homogeneous of some degree: it is arbitrary function belonging to $\mathbf{C}_2(\mathbb{R}^d \setminus \{0\})$. Clearly, then $f(\lambda)$, $\lambda \in \mathbf{S}^{d-1}$, is trace of $f(x)$ on $\mathbf{S}^{d-1}$. But the function $F(x) = |x|^p f(x/|x|)$ is evidently homogeneous of degree p.

For completeness of presentation let us adduce formulae for the generators in the Stratonovich case.

Proposition B.1 *Let $f(\lambda) \in \mathbf{C}_2(\mathbf{S}^{d-1})$. Let $f(x) \in \mathbf{C}_2(\mathbb{R}^d)$ have a compact support and be an extension of $f(\lambda)$. Then the generator $L^{(s)}$ of the Markov process governed by the Stratonovich SDE (B.15) is given by the formula*

$$L^{(s)} f(\lambda) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\lambda) q_i^{(s)}(\lambda) + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j}(\lambda) \sum_{r=1}^{k} h_{ri}(\lambda) h_{rj}(\lambda) \tag{B.64}$$

and the generator $L^{(s)}(p)$ of the semigroup $T_t^{(s)}(p) f(\lambda) = \mathbf{E}\{f(\Lambda^\lambda(t)) | X^\lambda(t)|^p\}$, where Λ and X satisfy (B.15) and (B.13), is given by

$$L^{(s)}(p) f(\lambda) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\lambda) q_i^{(s)}(\lambda) + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j}(\lambda) \sum_{r=1}^{k} h_{ri}(\lambda) h_{rj}(\lambda)$$

$$+ p \sum_{r=1}^{k} \left(\sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\lambda) h_{ri}(\lambda) \right) \cdot (\sigma_r \lambda, \lambda)$$

$$+ \left(p Q^{(s)}(\lambda) + \frac{1}{2} p^2 \sum_{r=1}^{k} (\sigma_r \lambda, \lambda)^2 \right) f(\lambda), \tag{B.65}$$

where the vector $q^{(s)}(\lambda)$ and scalar $Q^{(s)}(\lambda)$ are equal to

$$q^{(s)}(\lambda) = h_0(\lambda) + \frac{1}{2} \sum_{r=1}^{k} (\sigma_r^2 \lambda - (\sigma_r^2 \lambda, \lambda)\lambda) - \frac{1}{2} \sum_{r=1}^{k} (\sigma_r \lambda, \sigma_r \lambda)\lambda$$

$$- \sum_{r=1}^{k} (\sigma_r \lambda, \lambda)\sigma_r \lambda + \frac{3}{2} \sum_{r=1}^{k} (\sigma_r \lambda, \lambda)^2 \lambda, \tag{B.66}$$

$$Q^{(s)}(\lambda) = \left(\left(\sigma_0 + \frac{1}{2} \sum_{r=1}^{k} \sigma_r^2 \right) \lambda, \lambda \right) + \frac{1}{2} \sum_{r=1}^{k} (\sigma_r \lambda, \sigma_r \lambda) - \sum_{r=1}^{k} (\sigma_r \lambda, \lambda)^2. \tag{B.67}$$

B.2.5 *Various Representations of Semigroup $T_t(p) f(\lambda)$*

Consider the Itô SDE (see (B.6))

$$d\Lambda = q(\Lambda)dt + \sum_{r=1}^{k} h_r(\Lambda)d\xi_r(t), \quad \Lambda(0) = \lambda, \tag{B.68}$$

and the semigroup $T_t(p) f(\lambda) = \mathbf{E}\{f(\Lambda^\lambda(t))Z(t)\}$ where $Z(t) := |X^\lambda(t)|^p$ satisfies the scalar linear stochastic equation

$$dZ = Z\left(pQ(\Lambda) + \frac{1}{2}p^2 \sum_{r=1}^{k}(\sigma_r\Lambda, \Lambda)^2\right)dt$$

$$+ Zp\sum_{r=1}^{k}(\sigma_r\Lambda, \Lambda)d\xi_r(t), \quad Z(0) = 1. \tag{B.69}$$

The expectation $\mathbf{E}\{f(\Lambda^\lambda(t))Z(t)\}$ has various probabilistic representations. Parallel with system (B.68)–(B.69) introduce the SDE

$$d\widetilde{\Lambda} = q(\widetilde{\Lambda})\,dt - \sum_{r=1}^{k}\mu_r(\widetilde{\Lambda})h_r(\widetilde{\Lambda})\,dt + \sum_{r=1}^{k}h_r(\widetilde{\Lambda})d\xi_r(t), \quad \widetilde{\Lambda}(0) = \lambda, \tag{B.70}$$

$$d\widetilde{Z} = \widetilde{Z}\left(pQ(\widetilde{\Lambda}) + \frac{1}{2}p^2\sum_{r=1}^{k}(\sigma_r\widetilde{\Lambda}, \widetilde{\Lambda})^2\right)dt$$

$$+ \widetilde{Z}\sum_{r=1}^{k}(p(\sigma_r\widetilde{\Lambda}, \widetilde{\Lambda}) + \mu_r(\widetilde{\Lambda}))\,d\xi_r(t), \quad \widetilde{Z}(0) = 1, \tag{B.71}$$

where $\mu_r(\lambda)$ are arbitrary scalar smooth functions of $\lambda \in \mathbf{S}^{d-1}$ (they can depend also on p). One can verify that for the process $\widetilde{X}$ governed by the SDE

$$d\widetilde{X}(t) = B\widetilde{X}(t)dt - \sum_{r=1}^{k}\mu_r(\widetilde{X}/|\widetilde{X}|)\sigma_r\widetilde{X}(t)\,dt + \sum_{r=1}^{k}\sigma_r\widetilde{X}(t)\,d\xi_r(t),$$

the SDE for $\widetilde{\Lambda} = \widetilde{X}/|\widetilde{X}|$ is given by (B.70). The fact that $\mathbf{S}^{d-1}$ is invariant for (B.70) can be understood most easily if (B.70) is written in the Stratonovich form

$$d\widetilde{\Lambda} = h_0(\widetilde{\Lambda})\,dt - \sum_{r=1}^{k}\mu_r(\tilde{\Lambda})h_r(\widetilde{\Lambda})\,dt + \sum_{r=1}^{k}h_r(\widetilde{\Lambda})d^*\xi_r(t), \quad \widetilde{\Lambda}(0) = \lambda, \tag{B.72}$$

where

$$h_0(\lambda) = \sigma_0\lambda - (\sigma_0\lambda, \lambda)\lambda = q(\lambda) - \frac{1}{2}\sum_{r=1}^{k}\frac{\partial h_r}{\partial \lambda}\,h_r, \quad \sigma_0 = B - \frac{1}{2}\sum_{r=1}^{k}\sigma_r^2.$$

The vector fields $h_0, h_1, \ldots, h_r$ are tangent to the sphere $\mathbf{S}^{d-1}$. Hence the drift in (B.72) is tangent to the sphere as well. This ensures belonging $\widetilde{\Lambda}^\lambda(t)$ to $\mathbf{S}^{d-1}$.

Due to Girsanov's theorem (see, e.g., [164])

$$T_t(p)f(\lambda) = \mathbf{E}\{f(\Lambda^\lambda(t))Z(t)\} = \mathbf{E}\{f(\widetilde{\Lambda}^\lambda(t))\widetilde{Z}(t)\}. \tag{B.73}$$

Due to (B.73), taking in (B.70)–(B.71)

$$\mu_r(\lambda) = -p(\sigma_r\lambda, \lambda), \tag{B.74}$$

we obtain the next useful Proposition.

Proposition B.2 *The semigroup $T_t(p)$ has the following probabilistic representation*

$$T_t(p)f(\lambda) = \mathbf{E}\left\{f(\Lambda^\lambda(t))\exp\int_0^t\left[pQ(\Lambda^\lambda(t)) + \frac{1}{2}p^2R(\Lambda^\lambda(t))\right]dt\right\}, \tag{B.75}$$

where $R(\lambda) = \sum_{r=1}^{k}(\sigma_r\lambda, \lambda)^2$ and $\Lambda = \Lambda^\lambda(t)$ satisfies the SDE

$$d\Lambda = \left[q(\Lambda) + p\sum_{r=1}^{k}(\sigma_r\Lambda, \Lambda)h_r(\Lambda)\right]dt + \sum_{r=1}^{k}h_r(\Lambda)\,d\xi_r(t), \quad \Lambda(0) = \lambda. \tag{B.76}$$

It turns that for any smooth function $f(\lambda)$ the drifts in the expressions for $d[f(\Lambda^\lambda(t))Z(t)]$ and $d[f(\widetilde{\Lambda}^\lambda(t))\widetilde{Z}(t)]$ at the time $t = 0$ should coincide. In other words, the generators $\mathcal{L}$ and $\widetilde{\mathcal{L}}$ of the processes $(\Lambda(t), Z(t))$ and $(\widetilde{\Lambda}(t), \widetilde{Z}(t))$ should coincide. Indeed,

$$d[f(\widetilde{\Lambda}^\lambda(t))\widetilde{Z}(t)]$$

$$= \widetilde{Z}(t)\sum_{i=1}^{d}\frac{\partial f}{\partial x_i}(\widetilde{\Lambda})\left[q_i(\widetilde{\Lambda}(t)) - \sum_{r=1}^{k}\mu_r(\widetilde{\Lambda})h_{ri}(\widetilde{\Lambda})\right]dt$$

$$+ \frac{1}{2}\widetilde{Z}(t)\sum_{i=1}^{d}\frac{\partial^2 f}{\partial x_i\partial x_j}(\widetilde{\Lambda})\sum_{r=1}^{k}h_{ri}(\widetilde{\Lambda})h_{rj}(\widetilde{\Lambda})dt + df(\widetilde{\Lambda})\cdot d\widetilde{Z}(t)$$

$$+ \widetilde{Z}(t)f(\widetilde{\Lambda})\left[pQ(\widetilde{\Lambda}) + \frac{1}{2}p^2\sum_{r=1}^{k}(\sigma_r\widetilde{\Lambda}, \widetilde{\Lambda})^2\right]dt$$

$$+ \widetilde{Z}(t)\left[\sum_{i=1}^{d}\frac{\partial f}{\partial x_i}(\widetilde{\Lambda})\sum_{r=1}^{k}h_{ri}(\widetilde{\Lambda})d\xi_r(t)\right]$$

$$+ f(\tilde{\Lambda}) \sum_{r=1}^{k} (p(\sigma_r \tilde{\Lambda}, \tilde{\Lambda}) + \mu_r(\tilde{\Lambda})) d\xi_r(t) \bigg], \tag{B.77}$$

where $\tilde{\Lambda} = \tilde{\Lambda}^\lambda(t)$.

Since

$$df(\tilde{\Lambda}^\lambda(t)) \cdot d\tilde{Z}(t) = \tilde{Z}(t) \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\tilde{\Lambda}) \sum_{r=1}^{k} h_{ri}(\tilde{\Lambda})[p(\sigma_r \tilde{\Lambda}, \tilde{\Lambda}) + \mu_r(\tilde{\Lambda})] dt, \tag{B.78}$$

the terms in the drift of (B.77) containing μ_r are annihilated, hence $\mathcal{L}$ and $\tilde{\mathcal{L}}$ coincide. Thus, in the formula (B.73), the systems for $\tilde{\Lambda}$ and $\tilde{Z}$ vary simultaneously in such a way that the semigroup is not changed.

Due to (B.62), from (B.77) and (B.78), we get

$$d[f(\tilde{\Lambda}^\lambda(t))\tilde{Z}(t)] = \tilde{Z}(t)L(p)f(\tilde{\Lambda}) dt + \tilde{Z}(t)\bigg[\sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\tilde{\Lambda}) \sum_{r=1}^{k} h_{ri}(\tilde{\Lambda}) d\xi_r(t)$$

$$+ f(\tilde{\Lambda}) \sum_{r=1}^{k} (p(\sigma_r \tilde{\Lambda}, \tilde{\Lambda}) + \mu_r(\tilde{\Lambda})) d\xi_r(t) \bigg]. \tag{B.79}$$

Proposition B.3 *Let $f(\lambda) \in \mathbf{C}_2(\mathbf{S}^{d-1})$. Let $f(x) \in \mathbf{C}_2(\mathbb{R}^d)$ has a compact support and be an extension of $f(\lambda)$. Let $L(p)f(\lambda) = g(p)f(\lambda)$, $f(\lambda) > 0$, $\lambda \in \mathbf{S}^{d-1}$. Then $e^{-g(p)t}\tilde{Z}(t)f(\tilde{\Lambda}^\lambda(t))$ is a positive continuous martingale relative to the filtration generated by $\xi_r(t)$, $t \geq 0$, $r = 1, \ldots, k$.*

Proof Due to the condition of the Proposition and (B.79), direct calculation gives

$$d[e^{-g(p)t}\tilde{Z}(t)f(\tilde{\Lambda}^\lambda(t))] = e^{-g(p)t}\tilde{Z}(t)\bigg[\sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\tilde{\Lambda}) \sum_{r=1}^{k} h_{ri}(\tilde{\Lambda})d\xi_r(t)$$

$$+ f(\tilde{\Lambda}) \sum_{r=1}^{k} (p(\sigma_r \tilde{\Lambda}, \tilde{\Lambda}) + \mu_r(\tilde{\Lambda}))d\xi_r(t) \bigg],$$

whence the result follows. $\qquad \square$

Generator (B.62) is the sum of two operators $M(p)$ and $N(p)$, where

$$M(p)f(\lambda) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\lambda)\bigg[q_i(\lambda) + p\sum_{r=1}^{k} (\sigma_r\lambda, \lambda)h_{ri}(\lambda) \bigg]$$

$$+ \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j}(\lambda) \sum_{r=1}^{k} h_{ri}(\lambda)h_{rj}(\lambda), \tag{B.80}$$

$$N(p)f(\lambda) = \left[pQ(\lambda) + \frac{1}{2}p^2 \sum_{r=1}^{k} (\sigma_r \lambda, \lambda)^2 \right] f(\lambda). \tag{B.81}$$

The operator $M(p)$ is of the form (B.51) (see Theorem B.6). It is the generator of the Markov process Λ governed by SDE (B.76). Denote by $G_t(p)$ the Markov semigroup generated by $M(p)$. We have for $f \in \mathcal{D}(L(p))$ (see (B.62) for $L(p)$)

$$\frac{dT_t(p)f}{dt} = (M(p) + N(p))T_t(p)f. \tag{B.82}$$

Taking into account that $N(p)$ is a bounded operator and using the perturbation theory of semigroups (see [115, Sect. 9.2]), it is not difficult to obtain

$$T_t(p) = G_t(p) + \int_0^t G_{t-s}(p)N(p)T_s(p)\,ds. \tag{B.83}$$

Just this formula is used in [8] for proving the compactness of $T_t(p)$.

We note that formula (B.83) can be derived using probabilistic approach.

Letting in (B.75) $f(\lambda) = \mathbf{1}(\lambda)$, we get

$$T_t(p)\mathbf{1}(\lambda) = \mathbf{E}|X^\lambda(t)|^p = \mathbf{E}\exp\int_0^t \left[pQ(\Lambda^\lambda(t)) + \frac{1}{2}p^2 R(\Lambda^\lambda(t)) \right] dt. \tag{B.84}$$

Clearly, there exist the constants $Q_{\min} \leq Q_{\max}$ and $0 \leq R_{\min} \leq R_{\max}$ such that for $\lambda \in \mathbf{S}^{d-1}$ the inequalities

$$Q_{\min} \leq Q(\lambda) \leq Q_{\max}, \qquad 0 \leq R_{\min} \leq \sum_{r=1}^{k}(\sigma_r \lambda, \lambda)^2 \leq R_{\max}$$

are valid. From here

$$pQ_{\min} + \frac{1}{2}p^2 R_{\min} \leq pQ(\lambda) + \frac{1}{2}p^2 \sum_{r=1}^{k}(\sigma_r \Lambda^\lambda(t), \Lambda^\lambda(t))^2$$

$$\leq pQ_{\max} + \frac{1}{2}p^2 R_{\max}, \quad p \geq 0.$$

Hence

$$\left(pQ_{\min} + \frac{1}{2}p^2 R_{\min} \right)t \leq \ln \mathbf{E}|X^\lambda(t)|^p \leq \left(pQ_{\max} + \frac{1}{2}p^2 R_{\max} \right)t, \quad p \geq 0,$$

and

$$pQ_{\min} + \frac{1}{2}p^2 R_{\min} \leq g(p) \leq pQ_{\max} + \frac{1}{2}p^2 R_{\max}, \quad p \geq 0. \tag{B.85}$$

For $p < 0$ we have

$$p Q_{\max} + \frac{1}{2} p^2 R_{\min} \le g(p) \le p Q_{\min} + \frac{1}{2} p^2 R_{\max}, \quad p < 0. \tag{B.86}$$

From (B.85) and (B.86) we get

Proposition B.4 *If $R_{\min} > 0$ then the growth rate of $g(p)$ is quadratic with respect to p, as $|p| \to \infty$.*

B.3 Stability Index

Main results of this section are proven in [6], [19]. See also [21] for the analogous results concerning the products of random transformations.

B.3.1 Stability Index for Linear Stochastic Differential Equations

Clearly, if $g(p) > 0$ then $E|X(t)|^p \to \infty$ as $t \to \infty$ and if $g(p) < 0$ then $E|X(t)|^p \to 0$ as $t \to \infty$. If $g(p) = 0$ then there exist two constants $0 < c < C < \infty$ such that $c \le \mathbf{E}|X(t)|^p \le C$, $t \ge 0$. This easily follows from the fact that $|X(t)|^p f(\Lambda(t))$ is a martingale (see Proposition B.3). Let $a^* = g'(0) < 0$. In this case $g(p) < 0$ for sufficiently small positive p. If $g(p) \to \infty$ as $p \to \infty$ (Proposition B.4 gives a sufficient condition for such a behavior of $g(p)$) then the equation

$$g(p) = 0 \tag{B.87}$$

has a unique root $\gamma^* > 0$ (recall that $g(0) = 0$). The uniqueness follows from the convexity of g. We call γ^* the *stability index*.

Theorem B.8 *Assume the condition (C). Let $a^* = g'(0) < 0$ and $\gamma^* > 0$ is the root of (B.87). Then for some $K \ge 1$ and for any $\delta > 0$ under $|x| < \delta$ the inequalities*

$$\frac{1}{K} (|x|/\delta)^{\gamma^*} \le \mathbf{P}\left\{ \sup_{t \ge 0} |X^x(t)| > \delta \right\} \le K (|x|/\delta)^{\gamma^*} \tag{B.88}$$

are fulfilled.

Proof Let $f(\lambda)$ be a positive eigenfunction of $L(\gamma^*)$. Due to Theorem B.1: $L(\gamma^*) f = g(\gamma^*) f$. Without loss of generality we can set $1 \le f(\lambda) \le K$ for some $K > 0$. Since $g(\gamma^*) = 0$, the process $f(\Lambda^\lambda(t)) |X^x(t)|^{\gamma^*} = f(\Lambda^\lambda(t)) |X^\lambda(t)|^{\gamma^*} |x|^{\gamma^*}$, $\lambda = x/|x|$, is a positive continuous martingale. Due to Theorem 5.2 we have for any $T > 0$

$$\mathbf{P}\left\{ \sup_{0 \le t \le T} f(\Lambda^\lambda(t)) |X^x(t)|^{\gamma^*} > \delta^{\gamma^*} \right\}$$

$$\leq \frac{1}{\delta \gamma^*} \mathbf{E}(f(\Lambda^\lambda(T))|X^x(T)|^{\gamma^*})$$

$$= \frac{1}{\delta \gamma^*} f(\Lambda^\lambda(0))|X^x(0)|^{\gamma^*} = \frac{1}{\delta \gamma^*} f(\lambda)|x|^{\gamma^*} \leq K(|x|/\delta)^{\gamma^*}. \tag{B.89}$$

From here

$$\mathbf{P}\left\{\sup_{t \geq 0} f(\Lambda^\lambda(t))|X^x(t)|^{\gamma^*} > \delta^{\gamma^*}\right\} = \lim_{T \to \infty} \mathbf{P}\left\{\sup_{0 \leq t \leq T} f(\Lambda^\lambda(t))|X^x(t)|^{\gamma^*} > \delta^{\gamma^*}\right\}$$

$$\leq K(|x|/\delta)^{\gamma^*}.$$

Therefore

$$\mathbf{P}\left\{\sup_{t \geq 0} |X^x(t)| > \delta\right\} = \mathbf{P}\left\{\sup_{t \geq 0} |X^x(t)|^{\gamma^*} > \delta^{\gamma^*}\right\}$$

$$\leq P\left\{\sup_{t \geq 0} f(\Lambda^\lambda(t))|X^x(t)|^{\gamma^*} > \delta^{\gamma^*}\right\} \leq K(|x|/\delta)^{\gamma^*}, \tag{B.90}$$

hence the right hand inequality of (B.88) is proved.

Along with the martingale $M(t) = f(\Lambda^\lambda(t))|X^x(t)|^{\gamma^*}$ consider the stopped martingale $M(t \wedge \tau)$ where $\tau = \inf\{t : |X^x(t)| = \delta\}$. Note that $\{\sup_{t \geq 0} |X^x(t)| > \delta\} = \{\tau < \infty\}$. We have for any $t > 0$

$$|x|^{\gamma^*} \leq M(0) = EM(t \wedge \tau) = E(M(t \wedge \tau)\mathbb{1}_{\{\tau = \infty\}}) + E(M(t \wedge \tau)\mathbb{1}_{\{\tau < \infty\}})$$

$$\leq E(M(t)\mathbb{1}_{\{\tau = \infty\}}) + K\delta^{\gamma^*} \mathbf{P}\{\tau < \infty\}, \tag{B.91}$$

where $\mathbb{1}_{\{\cdot\}}$ is the indicator of the corresponding set $\{\cdot\}$. Since $a^* < 0$ and $\gamma^* > 0$, $M(t) \to 0$ almost surely as $t \to \infty$ and $M(t)$ is uniformly bounded by $K\delta^{\gamma^*}$ on the set $\{\tau = \infty\}$. Therefore $\mathbf{E}(M(t)\mathbb{1}_{\{\tau = \infty\}}) \to 0$ as $t \to \infty$ and we get from (B.91)

$$\mathbf{P}\left\{\sup_{t \geq 0} |X^x(t)| > \delta\right\} = \mathbf{P}\{\tau < \infty\} \geq \frac{1}{K}(|x|/\delta)^{\gamma^*}, \tag{B.92}$$

i.e., the left hand inequality of (B.88) is proved as well.   □

We note that for the one-dimensional system (B.4) in the case $\gamma^* = 1 - 2b/\sum_{r=1}^k \sigma_r^2 > 0$ we have

$$\mathbf{P}\left\{\sup_{t \geq 0} |X^x(t)| > \delta\right\} = (|x|/\delta)^{\gamma^*} \tag{B.93}$$

if $0 < |x| < \delta$. To prove (B.93), one can use Lemma 3.4 taking (ε, δ) as U^ε and taking the following function f^ε defined on $\Gamma = \{x = \varepsilon\} \cup \{x = \delta\}$ by $f^\varepsilon(\varepsilon) = 0$, $f^\varepsilon(\delta) = 1$. Getting $\mathbf{P}^x\{X(\tau_{U^\varepsilon}) = 1\}$ and letting $\varepsilon \to 0$, we obtain (B.93).

In the case $a^* = g'(0) > 0$, $|X^x(t)| \to \infty$ as $t \to \infty$ almost surely, i.e., $\sup_{t \geq 0} |X^x(t)| = \infty$. It is natural to consider inequalities for $\mathbf{P}\{\inf_{t \geq 0} |X^x(t)| < \delta\}$ (with $|x| > \delta$) instead.

Theorem B.9 *Assume the condition (C). Let $a^* = g'(0) > 0$ and $\gamma^* < 0$ is the root of (B.87). Then for some $K \geq 1$ and for any $\delta > 0$ under $|x| > \delta$ the inequality*

$$\frac{1}{K}(|x|/\delta)^{\gamma^*} \leq \mathbf{P}\left\{\inf_{t \geq 0}|X^x(t)| < \delta\right\} \leq K(|x|/\delta)^{\gamma^*} \tag{B.94}$$

is fulfilled.

Proof We note that the sets $\{\inf_{t\geq 0}|X^x(t)| < \delta\}$ and $\{\sup_{t\geq 0}|X^x(t)|^{\gamma^*} > \delta^{\gamma^*}\}$ are the same. Now the right hand inequality of (B.94) is proved literally as in the previous theorem with the exception of the first equality in (B.90): it should be $\mathbf{P}\{\inf_{t\geq 0}|X^x(t)| > \delta\} = \mathbf{P}\{\sup_{t\geq 0}|X^x(t)|^{\gamma^*} > \delta^{\gamma^*}\}$. Concerning the second part of the proof, it is sufficient to do the following small alterations. Here $M(t)$ and τ are the same but $\{\tau < \infty\} = \{\inf_{t\geq 0}|X^x(t)| < \delta\} = \{\sup_{t\geq 0}|X^x(t)|^{\gamma^*} > \delta^{\gamma^*}\}$. Taking into account that now $a^* > 0$ and $\gamma^* < 0$, we get

$$\mathbf{P}\left\{\inf_{t \geq 0}|X^x(t)| < \delta\right\} = \mathbf{P}\left\{\sup_{t \geq 0}|X^x(t)|^{\gamma^*} > \delta^{\gamma^*}\right\} = \mathbf{P}\{\tau < \infty\} \geq \frac{1}{K}(|x|/\delta)^{\gamma^*}$$

instead of (B.92), i.e., the left hand inequality of (B.94). $\qquad\square$

Analogously to (B.93) in the case $\gamma^* = 1 - 2b/\sum_{r=1}^{k}\sigma_r^2 < 0$ for the one-dimensional system (B.4) we have

$$\mathbf{P}\left\{\inf_{t \geq 0}|X^x(t)| < \delta\right\} = (|x|/\delta)^{\gamma^*}$$

if $|x| > \delta$.

Remark B.5 The theory of moment Lyapunov exponents and of the stability index can be carried over to SDEs with homogeneous coefficients with degree one (see Example 7.2, Sect. 7.3, and [8], [19], [6]).

B.3.2 Stability Index for Nonlinear SDEs

Consider an autonomous system of SDEs in the sense of Itô

$$dX = a_0(X)\,dt + \sum_{r=1}^{k}a_r(X)\,d\xi_r(t), \tag{B.95}$$

where X is an d-dimensional vector, $a_r(x) = [a_r^1, \ldots, a_r^d]^\top$, $r = 0, 1, \ldots, k$, are d-dimensional vector functions which are continuously differentiable in $\mathbb{R}^d$. Let the origin be a stationary point for the SDE (B.95), i.e., $a_r(0) = 0$, $r = 0, 1, \ldots, k$.

The linearized SDE for (B.95) has the form

$$dX = BX\,dt + \sum_{r=1}^{k} \sigma_r X \, d\xi_r(t), \qquad (B.96)$$

where $B = \{b^{ij}\}$ and $\sigma_r = \{\sigma_r^{ij}\}$ are $d \times d$-matrices with the elements $b^{ij} = \frac{\partial a_0^i}{\partial x_j}(0)$ and $\sigma_r^{ij} = \frac{\partial a_r^i}{\partial x_j}(0)$, $i, j = 1, \ldots, d, r = 1, \ldots, k$.

In the deterministic case, the solutions $X^x(t)$ of a nonlinear system and of the corresponding linearized system usually have many common features in their asymptotic behavior if x is sufficiently small. In the stochastic case, theorems on stability and instability of SDEs in the linear approximation (see Chap. 7) show that the most important asymptotic properties of linear SDEs are also robust with respect to perturbations of higher order. Nevertheless, the stochastic case is far intricate, and some asymptotic characteristics for (B.96) do not reflect the behavior of the solutions of (B.95). For example, such an important characteristic for the system (B.96) as the moment Lyapunov exponent

$$g(p) = \lim_{t \to \infty} \frac{1}{t} \ln \mathbf{E}|X^x(t)|^p \qquad (B.97)$$

is usually positive for sufficiently large $p > 0$ even for stable systems because of large deviations. At the same time, a situation may occur such that all the trajectories $X^x(t)$ of the nonlinear system (B.95) are uniformly bounded for $|x| \le \rho$, $\rho > 0$ is some number, $0 \le t < \infty$. In such a case the limit in (B.97) is always non-positive for the SDE (B.95), i.e., it is either zero or even negative. It turns out that stability index is just a characteristic that precisely relates (B.95) and (B.96) in the sense of asymptotic behavior of solutions. Let a^* be a Lyapunov exponent and γ^* be a stability index for the linear system (B.96). For its solutions, we have estimates for the probabilities $\mathbf{P}\{\sup_{t \ge 0} |X^x(t)| > \delta\}$, $|x| < \delta$ (the case $a^* < 0$, $\gamma^* > 0$) and $\mathbf{P}\{\inf_{t \ge 0} |X^x(t)| < \delta\}$, $|x| > \delta$ (the case $a^* > 0$, $\gamma^* < 0$) given by Theorems B.8 and B.9. In applications the first probability is called ruin probability. It represents the risk that the process $X^x(t)$ exceeds some threshold level before the large-time limiting behavior has taken effect (let us recall that $|X^x(t)| \to 0$ almost surely as $t \to \infty$). The second probability for suitable small δ is sometimes called extinction probability.

In the next theorem, we obtain the same estimates (with the same γ^*) for the probabilities with X being a solution of (B.95). It is very natural that the solutions of (B.95) and (B.96) behave so identically because these equations are close to each other so long as both the solutions remain in a neighborhood of $x = 0$.

Theorem B.10 *Assume the condition* (C) *for SDE* (B.96). *Assume that the coefficients of SDE* (B.95) *are close to the coefficients of the linear SDE* (B.96): *there exist* $\alpha > 0$ *(usually* $\alpha = 1$*),* $C > 0$, $\Delta > 0$ *such that*

$$|a_0(x) - Bx| + \sum_{r=1}^{k} |a_r(x) - \sigma_r x| < C|x|^{1+\alpha}, \quad \text{for } |x| \le \Delta. \qquad (B.98)$$

Let a^* and γ^* be the Lyapunov exponent and stability index for SDE (B.96). Then

Case $a^* < 0$, $\gamma^* > 0$: There exist a sufficiently small $\delta > 0$ and positive constants c_1, c_2 such that for all $|x| < \delta$ the solution $X^x(t)$ of the nonlinear SDE (B.95) satisfies the inequalities

$$c_1(|x|/\delta)^{\gamma^*} \leq \mathbf{P}\left\{\sup_{t \geq 0} |X^x(t)| > \delta\right\} \leq c_2(|x|/\delta)^{\gamma^*}. \qquad (B.99)$$

Case $a^* > 0$, $\gamma^* < 0$: There exist positive constants c_3, c_4 such that for any $0 < \delta < \Delta$ and all x with $\delta < |x| < \Delta$

$$c_3(|x|/\delta)^{\gamma^*} \leq \mathbf{P}\left\{\inf_{0 \leq t \leq \tau} |X^x(t)| < \delta\right\} \leq c_4(|x|/\delta)^{\gamma^*}, \qquad (B.100)$$

where $X^x(t)$ is the solution of the nonlinear SDE (B.95) and $\tau := \inf\{t : |X^x(t)| > \Delta\}$.

Proof Denote by $\tilde{L}$ and L the generators of Markov processes corresponding to (B.95) and (B.96), respectively. Thus,

$$\tilde{L}F(x) = \sum_{i=1}^{d} a_0^i(x)\frac{\partial F}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^{d}\sum_{r=1}^{k} a_r^i(x)a_r^j(x)\frac{\partial^2 F}{\partial x_i \partial x_j}(x),$$

$$F \in \mathbf{C}^2(\mathbb{R}^d \setminus \{0\}).$$

It follows from (B.98) after direct (but rather bulky) calculations that for any function F of the form $F(x) = |x|^p f(x/|x|)$ and for $|x| \leq \Delta$

$$|(\tilde{L} - L)F(x)| \leq C|x|^{p+\alpha}, \qquad (B.101)$$

where $C > 0$ is a constant (which may differ from C in (B.98)).

Case $a^* < 0$, $\gamma^* > 0$. Let $0 < \varepsilon < \alpha$. Let $L(p)$ be the generator of the semigroup $T_t(p)$ connected with linear system (B.96) and $f_{\gamma^*}(\lambda)$, $f_{\gamma^*+\varepsilon}(\lambda)$ be strictly positive eigenfunctions corresponding to the eigenvalues $g(\gamma^*) = 0$ and $g(\gamma^* + \varepsilon) > 0$:

$$L(\gamma^*)f_{\gamma^*}(\lambda) = g(\gamma^*)f_{\gamma^*}(\lambda) = 0,$$
$$L(\gamma^* + \varepsilon)f_{\gamma^*+\varepsilon}(\lambda) = g(\gamma^* + \varepsilon)f_{\gamma^*+\varepsilon}(\lambda). \qquad (B.102)$$

Introduce the functions

$$V_{\pm}(x) := |x|^{\gamma^*} f_{\gamma^*}(x/|x|) \pm |x|^{\gamma^*+\varepsilon} f_{\gamma^*+\varepsilon}(x/|x|). \qquad (B.103)$$

Clearly, these functions are strictly positive for $|x| < \delta$ if δ is sufficiently small. Moreover, there exist positive constants a_1, a_2 such that for all $|x| < \delta$

$$a_1|x|^{\gamma^*} \leq V_{\pm}(x) \leq a_2|x|^{\gamma^*}. \qquad (B.104)$$

Due to (B.63), we have

$$LV_-(x) = |x|^{\gamma^*} L(\gamma^*) f_{\gamma^*}(\lambda) - |x|^{\gamma^*+\varepsilon} L(\gamma^* + \varepsilon) f_{\gamma^*+\varepsilon}(\lambda)$$
$$= -|x|^{\gamma^*+\varepsilon} g(\gamma^* + \varepsilon) f_{\gamma^*+\varepsilon}(x/|x|) < 0. \tag{B.105}$$

Now due to (B.101),

$$\tilde{L}V_-(x) = LV_-(x) + (\tilde{L} - L)V_-(x)$$
$$= -|x|^{\gamma^*+\varepsilon} g(\gamma^* + \varepsilon) f_{\gamma^*+\varepsilon}(x/|x|) + O(|x|^{\gamma^*+\alpha}) < 0 \quad \text{for } |x| < \delta. \tag{B.106}$$

Therefore $V_-(X^x(t \wedge \tau))$, where $X^x(t)$ is a solution of (B.95) and $\tau := \inf\{t : |X^x(t)| = \delta\}$, is a supermartingale. Hence

$$EV_-(X^x(t \wedge \tau)) \leq V_-(x) \leq a_2 |x|^{\gamma^*}. \tag{B.107}$$

On the other hand,

$$EV_-(X^x(t \wedge \tau)) \geq a_1 \mathbf{E}|X^x(t \wedge \tau)|^{\gamma^*} \geq a_1 \delta^{\gamma^*} \mathbf{P}\left\{\sup_{t \geq 0} |X^x(t)| > \delta\right\}. \tag{B.108}$$

The inequalities (B.107) and (B.108) imply the right hand inequality of (B.99).

Because

$$LV_+(x) > 0,$$

the process $V_+(X^x(t \wedge \tau))$, where $X^x(t)$ is a solution of the nonlinear SDE (B.95) and $\tau := \inf\{t : |X^x(t)| = \delta\}$, is a submartingale. Hence

$$EV_+(X^x(t \wedge \tau)) \geq V_+(x) \geq a_1 |x|^{\gamma^*}. \tag{B.109}$$

Using the approach similar to one used in proving (B.91) and (B.92), we get

$$\mathbf{P}\left\{\sup_{t \geq 0} |X^x(t)| > \delta\right\} \geq \frac{1}{a_2 \delta^{\gamma^*}} \mathbf{E}V_+(X^x(t \wedge \tau)). \tag{B.110}$$

Inequalities (B.109) and (B.110) imply the left hand inequality of (B.99). Thus, the first case is proved.

The proof of the second case $a^* > 0$, $\gamma^* < 0$ is completely analogous (see also the proof of Theorem B.9). □

Remark B.6 The previous content is devoted to stability analysis of stationary points for linear and non-linear autonomous SDEs. In [199, 200], the concepts of Lyapunov exponents, moment Lyapunov exponents, and stability index are carried over for invariant manifolds of non-linear stochastic systems.

Remark B.7 The example in [134] shows that the inequality with $C|x|^{1+\alpha}$ in (B.98) cannot be replaced by the inequality with $\gamma|x|$ with sufficiently small $\gamma > 0$.

B.4 Moment Lyapunov Exponent and Stability Index for System with Small Noise

In this section, asymptotic expansion series for the moment Lyapunov exponent and stability index are constructed and justified for the two-dimensional linear stochastic system close to a harmonic oscillator.[4] As an example, a one-degree-of-freedom mechanical system parametrically excited in stiffness and damping is considered and several terms of the asymptotic expansion are obtained.

B.4.1 Introduction and Statement of Problem

Consider the Itô two-dimensional linear SDE

$$dX(t) = BX(t)\,dt + \sum_{r=1}^{k} \sigma_r X(t)\,d\xi_r(t). \tag{B.111}$$

Here B, $\sigma_1, \ldots, \sigma_k$ are 2×2 real matrices, $\dim X = 2$, and $\xi_r(t)$ are independent standard scalar Wiener processes on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Define (see Sect. B.1)

$$g(p; x) = \overline{\lim_{t \to \infty}} \frac{1}{t} \ln \mathbf{E}|X^x(t)|^p, \quad x \neq 0, \ p \in \mathbb{R}, \tag{B.112}$$

and suppose that the following Lie algebra condition holds (see Remark B.1)

$$\dim LA(h_0(\lambda), h_1(\lambda), \ldots, h_k(\lambda)) = 1, \quad \lambda \in \mathbf{S}^1, \tag{B.113}$$

where

$$h_0(\lambda) = \sigma_0\lambda - (\sigma_0\lambda, \lambda)\lambda, \quad \sigma_0 := B - \frac{1}{2}\sum_{r=1}^{k}\sigma_r^2,$$

$$h_r(\lambda) = \sigma_r\lambda - (\sigma_r\lambda, \lambda)\lambda, \quad r = 1, \ldots, k.$$

Then (see Sect. B.2) the limit in (B.112) exists and is independent of x. Let us recall that this limit $g(p)$ is a convex analytic function of $p \in \mathbb{R}^1$, $g(p)/p$ is increasing, and

$$a^* = \dot{g}(0) = \lim_{t \to \infty} \frac{1}{t} \ln |X^x(t)|, \quad \mathbf{P}\text{-a.s.}$$

is the Lyapunov exponent (a dot denotes differentiation with respect to p). The function $g(p)$ is the moment Lyapunov exponent.

---

[4]This section is based on the results of [136].

Let L be the generator of the Markov process defined by (B.111). Then (see Sect. B.2.4)

$$L(p)f(\lambda) = |x|^{-p} L(|x|^p f(x/|x|))$$

is a second order differential operator on the unit circle $\mathbf{S}^1$ and $g(p)$ is the principal simple eigenvalue of $L(p)$ (see Theorem B.1),

$$L(p)e(\lambda; p) = g(p)e(\lambda; p), \qquad (B.114)$$

with a strictly positive eigenfunction $e(\lambda; p)$.

If there is a nonzero solution γ^* of the equation $g(p) = 0$, then it is unique and is called the stability index (see Sect. B.3). The stability index controls the probability with which an almost surely (a.s.) stable system exceeds a threshold. It has been shown in [19] (see Sect. B.3.1) that the probability of exit from the ball $|x| < r$ has the order $|x|^{\gamma^*}$ for $x \to 0$ for any $r > 0$ if (B.111) is stable; and the solution $x = 0$ of (B.111) is exponentially p-stable for $p < \gamma^*$ and exponentially p-unstable for $p > \gamma^*$. Analogous results are valid for the a.s. unstable systems.

It has been proved in Sect. B.3.2 that the probability of exit for a nonlinear SDE close in some sense to a homogeneous one also has the order $|x|^{\gamma^*}$, where γ^* is the stability index of the linearized system. An analogous result has been established for the a.s. unstable systems.

Although Lyapunov exponent a^*, moment Lyapunov exponent $g(p)$, and the stability index γ^* are very important characteristics for the analysis of linear SDEs, in the general case it is impossible to find explicit expressions for these characteristics. Therefore, in some papers for stochastic linear systems close to deterministic, asymptotic expansions for a^*, $g(p)$, and γ^* have been proposed.

Asymptotics of the Lyapunov exponent for the second-order linear stochastic systems has been considered in [17], where using the exact formula for a^* and the saddle-point method, the zero- and first-order terms of the expansion for a^* have been obtained in general case and a general expansion has been obtained in the case when the unperturbed system is a simple harmonic oscillator. Also in [17], a general expansion with evaluation of the n-th remainder term is derived for the conservative system. In the case of real noises the asymptotics of the Lyapunov exponent and rotation number have been obtained in [9], [226]. By means of a singular perturbation scheme an asymptotic expansion for the Lyapunov exponent and rotation number has been obtained in [224] for two-dimensional systems with small white noises.

The problem of asymptotic analysis becomes more complicated for the moment Lyapunov exponent. It is known that the Lyapunov exponent $a^* = \dot{g}(0)$ of the two-dimensional stochastic system can be evaluated in quadratures using invariant measure of the corresponding process on $\mathbf{S}^1$ (see Sect. 6.8). The same is true for $g^{(n)}(0)$ (the procedure is described in [7, p. 121–122]). So the Taylor series for the moment

Lyapunov exponent of the two-dimensional linear SDE can be found in quadratures. However, these formulas are rather complicated to compute. Therefore, a series expansion of the invariant measure has been used in [53] to get an asymptotic expansion of $\ddot{g}(0)$ and obtain asymptotics of $g(p)$ for the case of small noise intensity and small p. The case of real noise has also been considered in this paper. But an approximation valid for small p does not allow us to find, for example, the stability index.

In general, it is difficult to find a uniform asymptotic expansion of $g_\varepsilon(p)$ because the stationary distribution of the corresponding process on the unit circle may concentrate in the vicinity of certain points as $\varepsilon \to 0$ (see Sect. 6 in [224]). But in the important case of a weakly perturbed conservative system, the limiting distribution on the circle is uniform and the asymptotic series expansion of $g_\varepsilon(p)$ can be obtained.

The main purpose of this section is to prove the existence of asymptotic expansion for $g_\varepsilon(p)$, γ_ε^*, and to propose a recursive procedure for their determination in the interesting for applications case of a weakly perturbed conservative system.

Assume that diffusion in the system (B.111) is small and matrix B is a small perturbation of the matrix with two pure imaginary eigenvalues. So the unperturbed system represent a simple harmonic oscillator. In this case the original SDE may be represented in the form

$$dX^\varepsilon(t) = (B_0 + \varepsilon B_1)X^\varepsilon(t)\,dt + \sqrt{\varepsilon}\sum_{r=1}^{k}\sigma_r X^\varepsilon(t)\,d\xi_r(t). \qquad (\text{B.115})$$

Without loss of generality we may assume that B_0 has the form

$$B_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Note that for the system (B.115) the Lie algebra condition (B.113) is satisfied for arbitrary matrices $\sigma_1, \ldots, \sigma_k$. The operator $L(p)$ and its eigenfunction $e(\theta; p)$ ($\theta \bmod \pi$ is a local coordinate on the unit circle $\mathbf{S}^1$) now depend on ε: $L(p) = L_\varepsilon(p)$, $e(\theta; p) = e_\varepsilon(\theta; p)$, and

$$L_\varepsilon(p) = L_1 + \varepsilon L_2(p). \qquad (\text{B.116})$$

Here $L_1 = -d/d\theta$ and $L_2(p)$ is the following second-order differential operator:

$$L_2(p) = a(\theta)\frac{d^2}{d\theta^2} + b(\theta; p)\frac{d}{d\theta} + c(\theta; p), \qquad (\text{B.117})$$

where

$$a(\theta) = \frac{1}{2} \sum_{r=1}^{k} (\sigma_r \lambda, \hat{\lambda})^2,$$

$$b(\theta; p) = -(B_1 \lambda, \hat{\lambda}) - (p-1) \sum_{r=1}^{k} (\sigma_r \lambda, \lambda)(\sigma_r \lambda, \hat{\lambda}),$$

(B.118)

$$c(\theta; p) = p \left[(B_1 \lambda, \lambda) + \frac{1}{2} \sum_{r=1}^{k} (\sigma_r \lambda, \hat{\lambda})^2 + \frac{p-1}{2} \sum_{r=1}^{k} (\sigma_r \lambda, \lambda)^2 \right],$$

$$\lambda = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \qquad \hat{\lambda} = -\frac{d\lambda}{d\theta} = \begin{pmatrix} \sin\theta \\ -\cos\theta \end{pmatrix}$$

(see Sect. 6.8). The eigenvalue problem (B.114) now takes the form

$$L_\varepsilon(p) e_\varepsilon(\theta; p) = g_\varepsilon(p) e_\varepsilon(\theta; p),$$

$$e_\varepsilon(0; p) = e_\varepsilon(\pi; p) = 1, \qquad e'_\varepsilon(0; p) = e'_\varepsilon(\pi; p).$$

(B.119)

B.4.2 Method of Asymptotic Expansion

We obtain here an asymptotic expansion of $g_\varepsilon(p)$,

$$g_\varepsilon = g_0 + \varepsilon g_1 + \cdots + \varepsilon^n g_n + O(\varepsilon^{n+1}),$$

(B.120)

whenever such an expansion holds, and we evaluate the coefficients $g_m = g_m(p)$. For this purpose we consider an asymptotic expansion of the eigenfunction $e_\varepsilon(\theta; p)$ as well:

$$e_\varepsilon(\theta; p) = e_0(\theta; p) + \varepsilon e_1(\theta; p) + \cdots + \varepsilon^n e_n(\theta; p) + \cdots.$$

Here $e_n(\theta; p)$ is the n-th term of the asymptotic expansion for $e_\varepsilon(\theta; p)$. The termal expressions

$$e_\varepsilon = e_0 + \varepsilon e_1 + \cdots + \varepsilon^n e_n + \cdots,$$

$$g_\varepsilon = g_0 + \varepsilon g_1 + \cdots + \varepsilon^n g_n + \cdots$$

after inserting in (B.119) and equating the coefficients of like power of ε lead to the equations

$$L_1 e_0 = g_0 e_0,$$

$$L_1 e_1 + L_2 e_0 = g_0 e_1 + g_1 e_0,$$

$$L_1 e_2 + L_2 e_1 = g_0 e_2 + g_1 e_1 + g_2 e_0, \tag{B.121}$$

$$\vdots$$

$$L_1 e_n + L_2 e_{n-1} = g_0 e_n + g_1 e_{n-1} + \cdots + g_n e_0.$$

Each function e_m here is a periodic function in θ. Consider the first equation of (B.121) (hereafter a prime denotes differentiation with respect to θ):

$$-e_0' = g_0 e_0, \qquad e_0(0) = 1.$$

Function e_0 must be positive and periodic in θ. It is possible only if $g_0 = 0$ and $e_0 = 1$. The second equation in (B.121) has a periodic solution if and only if

$$\int_0^\pi (L_2 \cdot 1 - g_1) \, d\theta = 0.$$

So we have

$$g_1 = g_1(p) = \overline{L_2 \cdot 1} = \overline{c_p(\theta)}$$

and

$$e_1 = e_1(\theta; p) = \int_0^\theta (L_2 \cdot 1 - g_1) \, ds.$$

Hereafter, we use the notation

$$\overline{(\cdot)} = \frac{1}{\pi} \int_0^\pi (\cdot) \, d\theta.$$

Analogously from the periodicity of $e_2 = e_2(\theta; p)$ we obtain

$$g_2 = g_2(p) = \overline{L_2(p) e_1(\theta; p)} - g_1 \overline{e_1(\theta; p)},$$

$$e_2(\theta; p) = \int_0^\theta (L_2(p) e_1(s; p) - g_1(p) e_1(s; p) - g_2(p)) \, ds,$$

and finally

$$g_n = \overline{L_2 e_{n-1}} - g_1 \overline{e_{n-1}} - \cdots - g_{n-1} \overline{e_1},$$

$$e_n = \int_0^\theta (L_2 e_{n-1} - g_1 e_{n-1} - \cdots - g_n e_0) \, ds. \tag{B.122}$$

Note that the asymptotic expansion formalism described here is similar to the one used in [224] for determination of the Lyapunov exponent.

Theorem B.11 *Let matrix B_0 in (B.115) have two pure imaginary eigenvalues and $g_0, g_1, \ldots, g_n$ and $e_0, e_1, \ldots, e_n$ be the functions obtained from the recursive procedure (B.122). Then for any $n > 0$*

$$g_\varepsilon = g_0 + \varepsilon g_1 + \cdots + \varepsilon^n g_n + O(\varepsilon^{n+1}) \quad (\varepsilon \to 0).$$

Here $O(\varepsilon^{n+1})/\varepsilon^{n+1}$ is bounded uniformly in $p \in \Delta$, and Δ is any compact in $\mathbb{R}$.

For proving the theorem we need the following lemma.

Lemma B.1 *Let $e_\varepsilon(\theta; p)$ be a function satisfying (B.119) and $e_\varepsilon(0; p) = 1$. Then for*

$$\varepsilon < \frac{M_1}{1 + p^2}, \tag{B.123}$$

the following estimates hold:

$$\|e'_\varepsilon(\,\cdot\,; p)\| < M_2 \sqrt{\varepsilon} \sqrt{p^2 + |p|}, \tag{B.124}$$

$$|e_\varepsilon(\theta; p) - 1| < \pi M_2 \sqrt{\varepsilon} \sqrt{p^2 + |p|}. \tag{B.125}$$

Here M_1, M_2 are positive constants independent of ε and p and

$$\|F\| = \sqrt{\frac{1}{\pi} \int_0^\pi F^2(\theta) \, d\theta}.$$

Proof Taking the inner product of both sides of (B.119) and e'_ε, we obtain

$$-\int_0^\pi e'^2_\varepsilon d\theta + \varepsilon \left(\int_0^\pi a e'_\varepsilon e''_\varepsilon \, d\theta + \int_0^\pi b e'^2_\varepsilon \, d\theta + \int_0^\pi c e'_\varepsilon e_\varepsilon \, d\theta \right) = g_\varepsilon \int_0^\pi e'_\varepsilon e_\varepsilon \, d\theta.$$

This equation can be rewritten (after integration by parts) in the form

$$\int_0^\pi \left[1 - \varepsilon \left(b - \frac{a'}{2} \right) \right] e'^2_\varepsilon \, d\theta = -\varepsilon \int_0^\pi \frac{c'}{2} e^2_\varepsilon \, d\theta. \tag{B.126}$$

Define $K_1(p)$, $K_2(p)$ by

$$K_1(p) = \max_{0 \le \theta \le \pi} \left| b(\theta, p) - \frac{a'(\theta)}{2} \right|, \qquad K_2(p) = \max_{0 \le \theta \le \pi} |c'(\theta, p)|. \tag{B.127}$$

It follows from (B.118) that

$$K_1(p) \le N_1(|p| + 1), \qquad K_2(p) \le N_2|p|(|p| + 1), \tag{B.128}$$

where N_1, N_2 are some constants independent of ε and p. It is clear from (B.126) that for $0 < \varepsilon < 1/(2K_1)$

$$\|e'_\varepsilon\| \le \sqrt{\varepsilon K_2} \, \|e_\varepsilon\|. \tag{B.129}$$

Using (B.129), we have

$$|e_\varepsilon(\theta; p) - 1| = \left| \int_0^\theta e_\varepsilon'(s; p)\,ds \right| \leq \pi \|e_\varepsilon'\| \leq \pi \sqrt{\varepsilon K_2} \|e_\varepsilon\|. \tag{B.130}$$

It follows from (B.130) that

$$\|e_\varepsilon\| \leq 1 + 2\pi \sqrt{\varepsilon K_2} \leq 2 \tag{B.131}$$

if ε is small enough and

$$\varepsilon \leq \frac{1}{4\pi^2 K_2}. \tag{B.132}$$

From (B.129), (B.131), and (B.132) it follows that

$$\|e_\varepsilon'\| \leq 2\sqrt{\varepsilon K_2} \tag{B.133}$$

and

$$|e_\varepsilon(\theta; p) - 1| \leq 2\pi \sqrt{\varepsilon K_2} \tag{B.134}$$

if ε satisfies (B.132) and $0 < \varepsilon < 1/(2K_1)$. Now the assertion of lemma follows from (B.128), (B.133), and (B.134). □

Proof of Theorem B.11 First, note that for the adjoint problem

$$L_\varepsilon^*(p)\psi_\varepsilon(\theta; p) = g_\varepsilon(p)\psi_\varepsilon(\theta; p), \quad \psi_\varepsilon(0; p) = \psi_\varepsilon(\pi; p) = 1,$$
$$(L_\varepsilon^*(p) = -L_1 + \varepsilon L_2^*(p)), \tag{B.135}$$

the same algorithm of asymptotic expansion can be applied. It results in the formal series

$$\psi_\varepsilon = \psi_0 + \varepsilon \psi_1 + \cdots + \varepsilon^n \psi_n + \cdots.$$

Functions $\psi_m = \psi_m(\theta; p)$ can be found analogously to (B.121). In particular, $\psi_0(\theta; p) = 1$. Introduce the notation

$$e_\varepsilon^{(n)} = e_0 + \varepsilon e_1 + \cdots + \varepsilon^n e_n,$$
$$g_\varepsilon^{(n)} = g_0 + \varepsilon g_1 + \cdots + \varepsilon^n g_n,$$
$$\psi_\varepsilon^{(n)} = \psi_0 + \varepsilon \psi_1 + \cdots + \varepsilon^n \psi_n.$$

Recall that $g_0 = g_0(p) = 0$, $e_0 = e_0(\theta; p) = 1$. It follows from the procedure of evaluation e_m, g_m, ψ_m that

$$L_\varepsilon(p)e_\varepsilon^{(n)}(\theta; p) = g_\varepsilon^{(n)}(p)e_\varepsilon^{(n)}(\theta; p) + O(\varepsilon^{n+1}), \tag{B.136}$$

$$L_\varepsilon^*(p)\psi_\varepsilon^{(n)}(\theta; p) = g_\varepsilon^{(n)}(p)\psi_\varepsilon^{(n)}(\theta; p) + O(\varepsilon^{n+1}). \tag{B.137}$$

Denote $F_\varepsilon^{(n)}(\theta; p) = e_\varepsilon(\theta; p) - e_\varepsilon^{(n)}(\theta; p)$. Then from (B.136) and (B.119) it follows that

$$L_\varepsilon(p) F_\varepsilon^{(n)}(\theta; p) - g_\varepsilon^{(n)}(p) F_\varepsilon^{(n)}(\theta; p) = (g_\varepsilon(p) - g_\varepsilon^{(n)}(p)) e_\varepsilon(\theta; p) + O(\varepsilon^{n+1}).$$
(B.138)

Taking the inner product of (B.138) and $\psi_\varepsilon^{(n)}(\theta; p)$ gives

$$(\psi_\varepsilon^{(n)}(\cdot\,; p), L_\varepsilon(p) F_\varepsilon^{(n)}(\cdot\,; p)) - g_\varepsilon^{(n)}(p)(\psi_\varepsilon^{(n)}(\cdot\,; p), F_\varepsilon^{(n)}(\cdot\,; p))$$
$$= (g_\varepsilon(p) - g_\varepsilon^{(n)}(p))(\psi_\varepsilon^{(n)}(\cdot\,; p), e_\varepsilon(\cdot\,; p)) + O(\varepsilon^{n+1}).$$
(B.139)

Using (B.137), the first term on the left side of (B.139) can be written in the form

$$(\psi_\varepsilon^{(n)}(\cdot\,; p), L_\varepsilon(p) F_\varepsilon^{(n)}(\cdot\,; p)) = (L_\varepsilon^*(p)\psi_\varepsilon^{(n)}(\cdot\,; p), F_\varepsilon^{(n)}(\cdot\,; p))$$
$$= g_\varepsilon^{(n)}(p)(\psi_\varepsilon^{(n)}(\cdot\,; p), F_\varepsilon^{(n)}(\cdot\,; p)) + O(\varepsilon^{n+1}).$$

Now (B.139) yields the estimate

$$(g_\varepsilon(p) - g_\varepsilon^{(n)}(p))(\psi_\varepsilon^{(n)}(\cdot\,; p), e_\varepsilon(\cdot\,; p)) = O(\varepsilon^{n+1}) \quad (\varepsilon \to 0).$$

We proved in the lemma that $e_\varepsilon(\theta; p) \to 1$ if $\varepsilon \to 0$ uniformly in θ and $p \in \Delta$, for any compact set $\Delta \in \mathbb{R}$. Therefore

$$\lim_{\varepsilon \to 0} (\psi_\varepsilon^{(n)}(\cdot\,; p), e_\varepsilon(\cdot\,; p)) = 1$$

and

$$g_\varepsilon(p) - g_\varepsilon^{(n)}(p) = O(\varepsilon^{n+1})$$

uniformly in $p \in \Delta$. This ends the proof of Theorem B.11.   □

B.4.3 Stability Index

Let us show now that the asymptotic expansion for the stability index γ_ε^* follows from the asymptotic expansion for the moment Lyapunov exponent $g_\varepsilon(p)$. Stability index γ_ε^* is a nonzero root of the equation

$$g_\varepsilon(\gamma_\varepsilon^*) = 0.$$
(B.140)

Let us use for γ_ε^* the similar formal asymptotic expansion,

$$\gamma_\varepsilon^* = \gamma_0 + \varepsilon\gamma_1 + \cdots + \varepsilon^n\gamma_n + \cdots.$$
(B.141)

After inserting (B.141) into the formal equation

$$g_1(\gamma_\varepsilon^*) + \varepsilon g_2(\gamma_\varepsilon^*) + \cdots + \varepsilon^{n-1} g_n(\gamma_\varepsilon^*) + \cdots = 0$$
(B.142)

and equating the coefficients of like power of ε, the following sequence of problems can be obtained:

$$g_1(\gamma_0) = 0,$$

$$\dot{g}_1(\gamma_0)\gamma_1 + g_2(\gamma_0) = 0,$$

$$\dot{g}_1(\gamma_0)\gamma_2 + \frac{1}{2}\ddot{g}_1(\gamma_0)\gamma_1^2 + \dot{g}_2(\gamma_0)\gamma_1 + g_3(\gamma_0) = 0,$$

$$\dot{g}_1(\gamma_0)\gamma_3 + \frac{1}{6}\dddot{g}_1(\gamma_0)\gamma_1^3 + \ddot{g}_1(\gamma_0)\gamma_1\gamma_2 \qquad\qquad (\text{B.143})$$

$$+ \frac{1}{2}\ddot{g}_2(\gamma_0)\gamma_1^2 + \dot{g}_2(\gamma_0)\gamma_2 + \dot{g}_3(\gamma_0)\gamma_1 + g_4(\gamma_0) = 0,$$

$$\vdots$$

As before, a dot denotes differentiation with respect to p.

Note that due to the procedure of evaluation of g_m, any function g_m is a polynomial in p. Consider the case when

$$d = \ddot{g}_1(0) = \sum_{r=1}^{k} \overline{(\sigma_r\lambda, \lambda)^2} > 0. \qquad\qquad (\text{B.144})$$

Then g_1 is a quadratic function of p, so it can be represented in the form $g_1 = \frac{1}{2}dp(p - \gamma_0)$. Therefore $\dot{g}_1(\gamma_0) = \frac{1}{2}dg\gamma_0 \neq 0$ and system (B.143) can be successively solved for $\gamma_0, \gamma_1, \ldots, \gamma_n, \ldots$.

Assume that n terms of the asymptotics of the moment Lyapunov exponent $g_\varepsilon(p)$ are used:

$$g_\varepsilon^{(n)}(p) = \varepsilon g_1(p) + \varepsilon^2 g_2(p) + \cdots + \varepsilon^n g_n(p). \qquad\qquad (\text{B.145})$$

So we can obtain $\gamma_0, \gamma_1, \ldots, \gamma_{n-1}$ by solving the first n equations of (B.143). Denote by

$$\gamma_\varepsilon^{(n-1)} = \gamma_0 + \varepsilon\gamma_1 + \cdots + \varepsilon^{n-1}\gamma_{n-1}.$$

Let us prove that $\gamma_\varepsilon^* = \gamma_\varepsilon^{(n-1)} + O(\varepsilon^n)$. It follows from (B.143), (B.145) that

$$g_\varepsilon^{(n)}(\gamma_\varepsilon^{(n-1)}) = O(\varepsilon^{n+1}) \qquad\qquad (\text{B.146})$$

uniformly in p for any compact set in $\mathbb{R}$. Function $g_1(p)$ has nonzero root at γ_0. So it follows from Theorem B.11 that $|\gamma_\varepsilon^*| < K$, where K does not depend on ε, and

$$g_\varepsilon(\gamma_\varepsilon^*) - g_\varepsilon^{(n)}(\gamma_\varepsilon^*) = O(\varepsilon^{n+1})$$

or

$$g_\varepsilon^{(n)}(\gamma_\varepsilon^*) = O(\varepsilon^{n+1}). \qquad\qquad (\text{B.147})$$

From (B.146) and (B.147) we now have

$$g_\varepsilon^{(n)}(\gamma_\varepsilon^*) - g_\varepsilon^{(n)}(\gamma_\varepsilon^{(n-1)}) = O(\varepsilon^{n+1}), \tag{B.148}$$

and according to the mean value theorem, there is a $\bar\gamma \in (\gamma_\varepsilon^*, \gamma_\varepsilon^{(n-1)})$ such that

$$g_\varepsilon^{(n)}(\gamma_\varepsilon^*) - g_\varepsilon^{(n)}(\gamma_\varepsilon^{(n-1)}) = \dot g_\varepsilon^{(n)}(\bar\gamma)(\gamma_\varepsilon^* - \gamma_\varepsilon^{(n-1)}). \tag{B.149}$$

For $\dot g_\varepsilon^{(n)}$ we have expression

$$\dot g_\varepsilon^{(n)} = \varepsilon(\dot g_1 + \varepsilon \dot g_2 + \cdots + \varepsilon^{n-1} \dot g_n). \tag{B.150}$$

So

$$\dot g_\varepsilon^{(n)}(\bar\gamma) = O(\varepsilon). \tag{B.151}$$

From (B.148), (B.149), and (B.151) it now follows that

$$\gamma_\varepsilon^* - \gamma_\varepsilon^{(n-1)} \sim O(\varepsilon^n).$$

So we obtain the following theorem.

Theorem B.12 *Let the matrix B_0 in (B.115) have two pure imaginary eigenvalues and condition (B.144) be fulfilled. Then the stability index of the system (B.115) has an asymptotic expansion*

$$\gamma_\varepsilon^* = \gamma_0 + \varepsilon\gamma_1 + \cdots + \varepsilon^n \gamma_n + O(\varepsilon^{n+1}),$$

where $\gamma_0, \gamma_1, \ldots, \gamma_n$ can be found recursively from (B.122), (B.143).

Remark B.8 Consider now the case when the original system (B.115) is not pure deterministic and $d = 0$ (so the condition (B.144) does not hold). This is possible if and only if each matrix σ_r is of the form: $\sigma_r = \nu_r B_0$. In this specific case random excitations act only along the circle on the phase plane. Stability properties of the system (B.115) depend crucially on the structure of the matrix B_1. Let

$$B_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

The coefficients of the operator $L_2(p)$ are evaluated as follows (see (B.118)):

$$a(\theta) = \frac{\nu^2}{2}, \quad \nu^2 = \sum_{r=1}^k \nu_r^2,$$

$$b(\theta; p) = -\frac{1}{2}[a_{12} - a_{21} + (a_{11} - a_{22})\sin 2\theta - (a_{12} + a_{21})\cos 2\theta], \tag{B.152}$$

$$c(\theta; p) = p\tilde c(\theta),$$

$$\tilde c(\theta) = \frac{1}{2}[\nu^2 + a_{11} + a_{22} + (a_{12} + a_{21})\sin 2\theta + (a_{11} - a_{22})\cos 2\theta],$$

and equation for $r := |X^\varepsilon(t)|$ is

$$dr = \varepsilon r \tilde{c}(\theta)\, dt. \tag{B.153}$$

We have three possibilities here.

(i) If $a_{11} - a_{22} = 0$ and $a_{12} + a_{21} = 0$, then (B.153) for r becomes deterministic and $g(p) = \frac{1}{2}\varepsilon(v^2 + a_{11} + a_{22})p$ as in deterministic theory ($e_\varepsilon(\theta; p) \equiv 1$).

(ii) If $(a_{11} - a_{22})^2 + (a_{12} + a_{21})^2 \neq 0$, and $v^2 + a_{11} + a_{22} \neq 0$, then

$$g(p) = \frac{1}{2}\varepsilon p(v^2 + a_{11} + a_{22}) + \frac{1}{16}\varepsilon^3 v^2 p(p+2)[(a_{11} - a_{22})^2 + (a_{12} + a_{21})^2]$$

$$+ \frac{1}{16}\varepsilon^4 v^2 p(p+2)(a_{21} - a_{12})[(a_{11} - a_{22})^2 + (a_{12} + a_{21})^2] + \cdots. \tag{B.154}$$

Asymptotic expansion for the stability index takes the form

$$\gamma_\varepsilon^* = \frac{\gamma_{-2}}{\varepsilon^2} + \frac{\gamma_{-1}}{\varepsilon} + \gamma_0 + \varepsilon\gamma_1 + \cdots, \tag{B.155}$$

where the coefficients $\gamma_{-2}, \gamma_{-1}, \gamma_0, \ldots$ can be found recursively by substituting formal series (B.155) into the equation $g(p) = 0$. In particular,

$$\gamma_{-2} = -\frac{8(v^2 + a_{11} + a_{22})}{v^2[(a_{11} - a_{22})^2 + (a_{12} + a_{21})^2]}, \quad \gamma_{-1} = (a_{12} - a_{21})\gamma_{-2}.$$

The proof is analogous.

(iii) If $(a_{11} - a_{22})^2 + (a_{12} + a_{21})^2 \neq 0$, and $v^2 + a_{11} + a_{22} = 0$, then for the moment Lyapunov exponent $g(p)$ we have expansion (B.154) where the first term vanishes. It is not difficult to prove that $\gamma_\varepsilon^* = -2$ in this case.

B.4.4 Applications

Consider the following stochastic oscillator:

$$\ddot{Y} + Y = -\varepsilon\alpha\,\dot{Y} + \sqrt{\varepsilon\beta}Y * \dot{\xi}_1(t) + \sqrt{\varepsilon\delta}\,\dot{Y} * \dot{\xi}_2(t) \quad (\beta^2 + \delta^2 > 0). \tag{B.156}$$

Now a dot denotes a time derivative, ξ_1 and ξ_2 are independent standard Wiener processes, and we consider this equation in the sense of Stratonovich. Equation (B.156) is a mathematical model of a one-degree-of-freedom mechanical system parametrically excited in stiffness and damping. The damping coefficient and intensities of excitations are assumed to be small, and this is indicated by the parameter ε. Note that the condition (B.144) is fulfilled since $\beta^2 + \delta^2 > 0$. Equivalently, (B.156) can be rewritten in the form

$$\ddot{Y} + Y = -\varepsilon(\alpha - \delta/2)\,\dot{Y} + \sqrt{\varepsilon\beta}\,Y\dot{\xi}_1(t) + \sqrt{\varepsilon\delta}\,\dot{Y}\dot{\xi}_2(t). \tag{B.157}$$

The operator $L_\varepsilon(p)$ (see (B.114)) has the form (B.116)–(B.118) where coefficients $a(\theta)$, $b(\theta; p)$, $c(\theta; p)$ can be evaluated as follows:

$$a(\theta) = \frac{1}{2}\cos^2\theta(\beta\cos^2\theta + \delta\sin^2\theta),$$

$$b(\theta; p) = -[(1-p)(\beta\cos^2\theta + \delta\sin^2\theta) + (\alpha - \delta/2)]\sin\theta\cos\theta,$$

$$c(\theta; p) = p[-(\alpha - \delta/2)]\sin^2\theta$$

$$+ \frac{1}{2}(\beta\cos^2\theta + \delta\sin^2\theta)(\cos^2\theta - \sin^2\theta) + p\sin^2\theta.$$

Following the procedure of the asymptotic analysis described in Sect. B.4.2, we obtain the following expressions for the moment Lyapunov exponent $g(p)$ and stability index γ_ε^* (we write here only three nonzero terms of the expansion):

$$\begin{aligned}
g(p) = &\ \varepsilon[2^{-4}(\beta + 3\delta)p^2 + 2^{-3}(\beta + \delta)p - 2^{-1}\alpha p]\\
&+ \varepsilon^3\{2^{-13}[33\delta^2(\beta + \delta) + \beta^2(11\beta - 13\delta)]p^4\\
&+ 2^{-11}[\delta(\beta + \delta)(17\delta - 32\alpha) + \beta^2(11\beta - 13\delta)]p^3\\
&+ 2^{-11}[(\beta + \delta)(32\alpha^2 - 64\alpha\delta - \delta^2) + \beta^2(-19\beta + 5\delta)]p^2\\
&+ 2^{-9}[(\beta + \delta)(16\alpha^2 - \delta^2) + \beta^2(-15\beta + 9\delta)]p\} + O(\varepsilon^5),\\
\gamma_\varepsilon^* = &\ 2(4\alpha - \beta - \delta)/(\beta + 3\delta) + \varepsilon^2(\beta - \delta)(\beta + 2\delta - 2\alpha)(\delta + 2\alpha)\\
&\times \frac{15\beta^3 + 52\beta^2\alpha + 77\beta^2\delta + 17\beta\delta^2 - 8\beta\delta\alpha + 3\delta^3 - 12\delta^2\alpha}{16(\beta + 3\delta)^4}\\
&+ O(\varepsilon^4).
\end{aligned} \tag{B.158}$$

The following expression for the Lyapunov exponent now follows from (B.158):

$$\begin{aligned}
a^* = &\ \varepsilon[2^{-3}(\beta + 3\delta) - 2^{-1}\alpha]\\
&+ 2^{-9}\varepsilon^3[(\beta + \delta)(16\alpha^2 - \delta^2) + \beta^2(-15\beta + 9\delta)] + O(\varepsilon^5).
\end{aligned}$$

The system is stable if $a^* < 0$. In terms of the system parameters α, β, δ it means that the system is stable if (here we use expansion of the Lyapunov exponent up to fifth order):

$$\begin{aligned}
\alpha > &\ 2^{-2}(\beta + \delta) + 2^{-7}\varepsilon^2\beta(-7\beta^2 + 6\beta\delta + \delta^2)\\
&- 2^{-12}\varepsilon^4\beta(-402\beta^4 + 465\beta^3\delta - 71\beta^2\delta^2 + 3\beta\delta^3 + 5\delta^4) + O(\varepsilon^6).
\end{aligned}$$

All expressions become simpler for some special cases.

Case I Parametrically excited system in the stiffness ($\delta = 0$).

Stability index:

$$\gamma_\varepsilon^* = 8\frac{\alpha}{\beta} - 2 + \varepsilon^2\left(\frac{15}{8}\alpha\beta - 13\frac{\alpha^3}{\beta} + \frac{11}{4}\alpha^2\right)$$

$$+ \varepsilon^4\left(-\frac{735}{256}\alpha\beta^3 - \frac{404}{64}\alpha^2\beta^2 + \frac{699}{32}\alpha^3\beta - \frac{97}{8}\frac{\alpha^5}{\beta} + \frac{337}{32}\alpha^4\right) + O(\varepsilon^6).$$

Lyapunov exponent:

$$a^* = \varepsilon\left(\frac{1}{8}\beta - \frac{1}{2}\alpha\right) + \varepsilon^3\left(-\frac{15}{512}\beta^3 + \frac{1}{32}\alpha^2\beta\right)$$

$$+ \varepsilon^5\left(\frac{1695}{32768}\beta^5 - \frac{15}{512}\alpha^2\beta^3 + \frac{1}{128}\alpha^4\beta\right) + O(\varepsilon^7).$$

Stability condition:

$$\alpha > \frac{1}{4}\beta - \frac{7}{128}\varepsilon^2\beta^3 + \frac{201}{2048}\varepsilon^4\beta^5 + O(\varepsilon^6).$$

Note that the first term in this condition coincides with the one obtained in [124].

Case II Parametrically excited system in the damping ($\beta = 0$).

Stability index:

$$\gamma_\varepsilon^* = \frac{8\alpha}{3\delta} - \frac{2}{3} + \varepsilon^2\left(-\frac{1}{216}\delta^2 + \frac{1}{72}\alpha\delta - \frac{\alpha^3}{27\delta} + \frac{1}{36}\alpha^2\right)$$

$$+ \varepsilon^4\left(\frac{55}{62208}\delta^4 - \frac{155}{62208}\alpha\delta^3\right.$$

$$\left. - \frac{95}{15552}\alpha^2\delta^2 + \frac{55}{7776}\alpha^3\delta - \frac{5\alpha^5}{1944\delta} + \frac{25\alpha^4}{7776}\right) + O(\varepsilon^6).$$

Lyapunov exponent:

$$a^* = \varepsilon\left(\frac{1}{8}\delta - \frac{1}{2}\alpha\right) + \varepsilon^3\left(-\frac{1}{512}\delta^3 + \frac{1}{32}\alpha^2\delta\right)$$

$$+ \varepsilon^5\left(\frac{13}{32768}\delta^5 - \frac{7}{1024}\alpha^2\delta^3 + \frac{1}{128}\alpha^4\delta\right) + O(\varepsilon^7).$$

Stability condition:

$$\alpha > \frac{1}{4}\delta + O(\varepsilon^6).$$

Remark B.9 The interesting feature of the system is that $g_\varepsilon(p)$ is the odd function of ε, so it contains only odd terms in the expansion. Let us show that it happens

when $a(\theta)$ and $c(\theta; p)$ are even, and $b(\theta; p)$ is an odd function of θ. In fact if we replace θ by $-\theta$ and ε by $-\varepsilon$, then the operator $L_\varepsilon(p)$ (see (B.116)–(B.118)) will convert into the operator $-L_\varepsilon(p)$, which has eigenvalue $-g_\varepsilon(p)$. Operator $-L_\varepsilon(p)$ has eigenvalue $g_{-\varepsilon}(p)$. Since the transformation of coordinate $\theta \to -\theta$ does not change the eigenvalue, then $g_{-\varepsilon} = -g_\varepsilon$.

Remark B.10 The results of this section are carried over for conservative systems with small random and periodic excitations in [102] (see [20] as well). Also in [102], an outline of general theory for stochastic systems with periodic in t coefficients is given. In [135], gyroscopic systems with two degrees of freedom under small random perturbations are investigated by use of the stochastic averaging principle. Explicit formulas for the averaged Lyapunov and moment Lyapunov exponents are derived. It turned out that the averaged exponents are the first-order terms of the expansions for the exponents of the original gyroscopic system.

References

1. Akhmetkaliev, T.: On the connection between the stability of stochastic difference and differential systems. Differentsial'nye Uravneniya **1**, 1016–1026 (1965). English transl.: Differential Equations **1**, 790–798 (1965)
2. Akhmetkaliev, T.: Stability of difference stochastic systems under constantly acting perturbations. Differentsial'nye Uravneniya **2**, 1161–1169 (1966). English transl.: Differential Equations **2**, 600–604 (1966)
3. Arnold, L.: Stabilization by noise revisited. Z. Angew. Math. Mech. **70**(7), 235–246 (1990)
4. Arnold, L.: A formula connecting sample and moment stability of linear stochastic systems. SIAM J. Appl. Math. **44**, 793–802 (1984)
5. Arnold, L.: Random Dynamical Systems. Springer, Berlin (1998)
6. Arnold, L., Khasminskii, R.Z.: Stability index for nonlinear stochastic differential equations. Proc. Symp. Pure Math. **57**, 543–551 (1995)
7. Arnold, L., Kliemann, W., Oeljeklaus, E.: Lyapunov exponents of linear stochastic systems. In: Arnold, L., Wihstutz, V. (eds.) Lyapunov Exponents. Lecture Notes in Math., vol. 1186, pp. 85–125. Springer, Berlin (1986)
8. Arnold, L., Oeljeklaus, E., Pardoux, E.: Almost sure and moment stability for linear Ito equations. In: Arnold, L., Wihstutz, V. (eds.) Lyapunov Exponents. Lecture Notes in Math., vol. 1186, pp. 129–159. Springer, Berlin (1986)
9. Arnold, L., Papanicolaou, E., Wihstutz, V.: Asymptotic analysis of the Lyapunov exponent and rotation of the random oscillator and applications. SIAM J. Appl. Math. **46**, 427–450 (1986)
10. Arnold, L., Wihstutz, V. (eds.) Lyapunov Exponents. Lecture Notes in Math., vol. 1186. Springer, Berlin (1986)
11. Arnold, L., Crauel, H., Eckmann, J.-P. (eds.) Lyapunov Exponents. Lecture Notes in Math., vol. 1486. Springer, Berlin (1991)
12. Arnold, V.I.: Mathematical Methods of Classical Mechanics. Springer, Berlin (1978)
13. Ariaratnam, S., Graefe, P.: Linear systems with stochastic coefficients. I. Internat. J. Control **1**(1), 239–248 (1965)
14. Ariaratnam, S., Graefe, P.: Linear systems with stochastic coefficients. II. Internat. J. Control **2**(1), 161–169 (1965)
15. Ariaratnam, S., Graefe, P.: Linear systems with stochastic coefficients. III. Internat. J. Control **2**, 205–211 (1965)
16. Aström, K.J.: On a first-order stochastic differential equation. Internat. J. Control **1**(1), 301–326 (1965)
17. Auslender, E.I., Milstein, G.N.: Asymptotic expansions of the Lyapunov index for linear stochastic systems with small noise. J. Appl. Math. Mech. **46**, 277–283 (1983)
18. Barbashin, E.A., Krasovskii, N.N.: On stability of motion in the large. Dokl. Akad. Nauk USSR **86**, 453–456 (1952) (Russian)

19. Baxendale, P.: Moment stability and large deviations for linear stochastic differential equations. In: Ikeda, N. (ed.) Proc. Taniguchi Symp. on Probab. Meth. in Math. Physics, Katata and Kyoto, 1985, pp. 31–54. Kinokuniya, Tokyo (1987)

20. Baxendale, P.: Lyapunov exponents and resonance for small periodic and random perturbations of a conservative linear system. Stochastics and Dynamics **2**(1), 49–66 (2002)

21. Baxendale, P.H., Khasminskii, R.Z.: Stability index for products of random transformations. Adv. Appl. Prob. **30**, 968–988 (1998)

22. Bedelbaev, A.K.: On the structure of Lyapunov functions which are quadratic forms and their application to the stability of control systems. Avtomatika (Kiev) **1**, 37–43 (1958) (Ukrainian)

23. Bellman, R.E.: Stability of Differential Equations. McGraw-Hill, New York (1953). Russian transl.: IL, Moscow (1954)

24. Bellman, R.E.: Dynamic Programming. Princeton University Press, Princeton (1957). Russian transl.: IL, Moscow (1960)

25. Bellman, R.E.: Stochastic transformation and functional equations. In: Proc. Symp. Appl. Math., vol. 16, pp. 171–177. Am. Math. Soc., Providence (1964)

26. Benderskii, M.M.: Determination of the region of stability of a second order linear differential equation with random coefficients of a special form. Differentsial'nye Uravneniya **5**, 1885–1888 (1969) (Russian)

27. Benderskii, M.M.: Determination of the moments of the solutions of a system of linear differential equations with random Markov coefficients. Trudy FTINT, Matematicheskaya Fizika i Funktsional'nyi Analiz (Har'kov) **1**, 64–71 (1969) (Russian)

28. Benderskii, M.M.: On the asymptotic behavior of the moments of the solutions of a linear system with random coefficients. Trudy FTINT, Matematicheskaya Fizika i Funktsional'nyi Analiz (Har'kov) **3**, 15–21 (1972) (Russian)

29. Benderskii, M.M., Pastur, L.A.: Asymptotical behavior of the solutions of a second order equation with random coefficients. Teoriya Funktsii i Funktsional'nyi Analiz (Har'kov) **22**, 3–14 (1973) (Russian)

30. Benderskii, M.M., Pastur, L.A.: The spectrum of the one-dimensional Schrodinger equation with random potential. Mat. Sb. (N.S.) **82**(124), 273–284 (1970) (Russian)

31. Bertram, J.E., Sarachik, P.E.: Stability of circuits with randomly time-varying parameters. In: Proc. Internat. Sympos. on Circuit and Information Theory, Los Angeles, Calif. IRE Trans., vol. CT-6, pp. 809–823 (1959)

32. Bharucha-Redi, A.T.: On the theory of random equations. In: Proc. Symp. Appl. Math., vol. 16, pp. 40–69. Am. Math. Soc., Providence (1964)

33. Blagoveshchenskii, Yu.N., Freidlin, M.I.: Certain properties of diffusion processes depending on a parameter. Dokl. Akad. Nauk USSR **138**, 508–511 (1961). English transl.: Soviet Math. Dokl. **2**, 633–636 (1961)

34. Bogdanov, J.L., Kozin, F.: Moments of the output of linear random systems. J. Acoust. Soc. Amer. **34**, 1063–1066 (1962)

35. Borovkov, A.A.: Ergodicity and Stability of Stochastic Processes. Wiley, New York (1998)

36. Brockett, R.W., Willems, J.C.: Average Value Criteria for Stochastic Stability. Lecture Notes in Mathematics, vol. 294 (1972)

37. Bucy, R.S.: Stability and positive supermartingales. J. Differential Equations **1**, 151–155 (1965)

38. Bunke, H.: Über das asymptotische Verhalten von Losungen linearer stochastischer Differentialgleichungssysteme. Z. Angew. Math. Mech. **45**, 1–9 (1965)

39. Bunke, H.: Über die fast sichere Stabilität linearer stochastischer Systeme. Z. Angew. Math. Mech. **43**, 533–535 (1963)

40. Bunke, H.: Stabilität bei stochastischen Differentialgleichungssysteme. Z. Angew. Math. Mech. **43**, 63–70 (1963)

41. Bunke, H.: Gewöhnliche Differentialgleiohungen mit zufälligen Parametern. Akademie Verlag, Berlin (1972)

42. Caughey, T.K., Dienes, J.K.: The behavior of linear systems with random parametric excitation. J. Math. Phys. **41**, 300–318 (1962)

43. Caughey, T.K., Gray, A.H. Jr.: On the almost sure stability of linear dynamic systems with stochastic coefficients. Trans. ASME Ser. E J. Appl. Mech. **32**, 365–372 (1965)

44. Caughey, T.K., Gray, A.H. Jr.: A controversy in problems involving random parametric excitation. J. Math. Phys. **44**, 288–296 (1965)

45. Chelpanov, I.B.: Vibration of a second-order system with a randomly varying parameter. Prikl. Mat. Meh. **26**, 762–766 (1962). English transl.: J. Appl. Math. Mech. **26**, 1145–1152 (1962)

46. Chetaev, N.G.: Stability of Motion, 3rd edn. Nauka, Moscow (1965). English transl. of 2nd edn. Pergamon Press, Oxford (1961)

47. Chan, S.Y., Chuang, K.: A study of linear time-varying systems subject to stochastic disturbances. Automatica **4**(1), 31–48 (1966)

48. Chung, K.L.: Contributions to the theory of Markov chains. II. Trans. Amer. Math. Soc. **76**, 397–419 (1954)

49. Clément, Ph., Heijmans, H.J.A.M., et al.: One-Parameter Semigroups. North-Holland, Amsterdam (1987)

50. Darkovskii, B.S., Leibovich, V.S.: Statistical stability and output moments of a certain class of systems with at random varying structure. Avtom. i Telemekh. **10**, 36–43 (1971) (Russian)

51. Demidovich, B.P.: On the dissipativity of a certain non-linear system of differential equations. I. Vestnik Moskov. Univ. Ser. I Mat. Meh. **6**, 19–27 (1961) (Russian)

52. Demidovich, B.P.: On the dissipativity of a certain non-linear system of differential equations. II. Vestnik Moskov. Univ. Ser. I Mat. Meh. **1**, 3–8 (1962) (Russian)

53. Doyle, M., Namachivaya, N.Sri., Arnold, L.: Small noise expansion of moment Lyapunov exponents for two-dimensional systems. Dynam. Stability Systems **12**, 187–211 (1997)

54. Doob, J.L.: Asymptotic properties of Markov transition probabilities. Trans. Amer. Math. Soc. **63**, 393–421 (1948)

55. Doob, J.L.: Martingales and one-dimensional diffusion. Trans. Amer. Math. Soc. **78**, 168–208 (1955)

56. Doob, J.L.: Stochastic Processes. Wiley/Chapman & Hall, New York (1953). Russian transl.: IL, Moscow (1956)

57. Dorogovtsev, A.Ya.: Some remarks on differential equations perturbed by periodic random processes. Ukrain. Mat. Zh. **14**(2), 119–128 (1962). English transl.: In: Selected Transl. Math. Statist. and Probability, vol. 5, pp. 259–269. Am. Math. Soc, Providence (1965)

58. Driml, M., Hans, O.: Continuous stochastic approximations. In: Trans. Second Prague Conf. on Information Theory, pp. 113–122. Publ. House Czech. Akad. Sci., Prague (1960). Academic Press, New York (1961)

59. Driml, M., Nedoma, J.: Stochastic approximations for continuous random processes. In: Trans. Second Prague Conf. on Information Theory, pp. 145–158. Publ. House Czech. Akad. Sci., Prague (1960). Academic Press, New York (1961)

60. Dubrovin, B.A., Novikov, S.P., Fomenko, A.T.: The Contemporary Geometry. Nauka, Moscow (1979)

61. Dym, H.: Stationary measures for the flow of a linear differential equation driven by white noise. Trans. Amer. Math. Soc. **123**, 130–164 (1966)

62. Dynkin, E.B.: Infinitesimal operators of Markov processes. Teor. Veroyatnost. i Primenen. **1**, 38–60 (1956). English transl.: Theor. Probability Appl. **1**, 34–54 (1956)

63. Dynkin, E.B.: Foundations of the Theory of Markov Processes. Fizmatgiz, Moscow (1959). English transl.: Prentice Hall, Englewood Cliffs, Pergamon Press, Oxford (1961)

64. Dynkin, E.B.: Markov Processes. Fizmatgiz, Moscow (1963). English transl.: Die Grundlehren der Math. Wissenschaften, Bände 121, 122. Academic Press, New York. Springer, Berlin (1965)

65. Dynkin, E.B., Yushkevich, A.A.: Strong Markov processes. Teor. Veroyatnost. i Primenen. **1**, 149–155 (1956). English transl.: Theor. Probability Appl. **1**, 134–139 (1956)

66. Eidelman, S.D.: Parabolic Systems. Nauka, Moscow (1964). English transl.: Noordhoff, Groningen. North-Holland, Amsterdam (1969)

67. Erdös, P., Kac, M.: On the number of positive sums of independent random variables. Bull. Amer. Math. Soc. **53**, 1011–1020 (1947)

68. Fabian, V.: A survey of deterministic and stochastic approximation methods for the minimization of functions. Kibernetika (Prague) **1**, 499–523 (1965) (Czech.)
69. Fabian, V.: Stochastic approximation method. Czechoslovak Math. J. **10**(85), 123–159 (1960)
70. Fabian, V.: A new one-dimensional stochastic approximation method for finding a local minimum of a function. In: Trans. Third Prague Conf. on Information Theory, Statist. Decision Functions, Random Processes, Liblice, 1962, pp. 85–105. Publ. House Czech. Akad. Sci., Prague (1964)
71. Feller, W.K.: Fluctuation theory of recurrent events. Trans. Amer. Math. Soc. **67**, 98–119 (1949)
72. Feller, W.K.: The parabolic equations and the associated semi-groups of transformations. Ann. of Math. (2) **55**, 468–519 (1952). Russian transl.: Matematika **1**(4), 105–153 (1957)
73. Feller, W.K.: Diffusion processes in one dimension. Trans. Amer. Math. Soc. **77**, 1–31 (1954)
74. Feller, W.K.: An Introduction to Probability Theory and Its Applications, vol. 2. Wiley, New York (1966). Russian transl.: Mir, Moscow (1967)
75. Fleming, W.H.: Duality and a priori estimates in Markovian optimization problems. J. Math. Anal. Appl. **16**, 254–279 (1966)
76. Friedman, A.: Interior estimates for parabolic systems of partial differential equations. J. Math. Mech. **7**, 393–417 (1958)
77. Friedman, A.: Stability and angular behavior of solutions of stochastic differential equations. In: Lecture Notes in Math., vol. 294, pp. 14–20 (1972)
78. Friedman, A., Pinsky, M.A.: Asymptotic behavior of solutions of linear stochastic differential systems. Trans. Amer. Math. Soc. **181**, 1–22 (1973)
79. Friedman, A., Pinsky, M.A.: Asymptotic behavior and spiraling properties of stochastic equations. Trans. Amer. Math. Soc. **186**, 331–358 (1973)
80. Friedman, A., Pinsky, M.A.: Dirichlet problem for degenerate elliptic equations. Trans. Amer. Math. Soc. **186**, 359–383 (1973)
81. Frisch, U.: Sur la résolution des équations différentielles stochastiques à coefficients markoviens. C. R. Acad. Sci. Paris Sér. A–B **262**, A762–A765 (1966)
82. Furstenberg, H.: Noncommuting random products. Trans. Amer. Math. Soc. **108**, 377–428 (1963)
83. Gabasov, R.: On the stability of stochastic systems with a small parameter multiplying the derivatives. Uspekhi Mat. Nauk **20**1(121), 189–196 (1965) (Russian)
84. Gantmakher, F.R.: The Theory of Matrices, 2nd edn. Nauka, Moscow (1966). English transl. of 1st edn., vols. 1, 2. Chelsea, New York (1959)
85. Gelfand, I.M.: Lectures on Linear Algebra, 2nd edn. GITTL, Moscow (1951). English transl.: Pure and Appl. Math., vol. 9. Interscience, New York (1961)
86. Gelfand, I.M.: Generalized random processes. Dokl. Akad. Nauk SSSR **100**, 853–856 (1955) (Russian)
87. Germaidze, V.E., Kac, I.Ya.: A problem of stability in the first approximation for stochastic systems. Ural. Politekhn. Inst. Sb. **139**, 27–35 (1964) (Russian)
88. Gikhman, I.I.: On the theory of differential equations for stochastic processes, I. Ukrain. Mat. Zh. **2**(3), 37–63 (1950). English transl.: Amer. Math. Soc. Transl. (2) **1**, 111–137 (1955)
89. Gikhman, I.I.: Differential equations with random functions. In: Winter School in Theory of Probability and Math. Statist., Uzhgorod, 1964, pp. 41–86. Izdat. Akad. Nauk Ukrain. SSR, Kiev (1964) (Russian)
90. Gikhman, I.I.: Stability of solutions of stochastic differential equations. In: Limit Theorems Statist. Inference, pp. 14–45. Izdat. "Fan", Tashkent (1966). English transl.: Selected Transl. Statist. and Probability, vol. 12, pp. 125–154. Am. Math. Soc., Providence (1973)
91. Gikhman, I.I., Dorogovtsev, A.Ya.: On stability of solutions of stochastic differential equations. Ukrain. Mat. Zh. **17**(6), 3–21 (1965). English transl.: Amer. Math. Soc. Transl. (2) **72**, 229–250 (1968)
92. Gikhman, I.I., Skorokhod, A.V.: Introduction to the Theory of Random Processes. Nauka, Moscow (1965). English transl.: Saunders, Philadelphia (1969)

93. Gikhman, I.I., Skorokhod, A.V.: Stochastic Differential Equations. Springer, Berlin (1972)
94. Gladyshev, E.G.: On stochastic approximation. Teor. Veroyatnost. i Primenen. **10**, 297–300 (1965). English transl.: Theor. Probability Appl. **10**, 275–278 (1965)
95. Gnedenko, B.V.: A Course in the Theory of Probability, 3rd edn. Fizmatgiz, Moscow (1961). English transl.: Chelsea, New York (1962)
96. Grenander, U.: Probabilities on Algebraic Structures. Wiley, New York (1963)
97. Halmos, P.: Measure Theory. Van Nostrand, Princeton (1950). Russian transl.: IL, Moscow (1953)
98. Hopf, E.: Ergodentheorie. Ergebnisse der Mathematik und Ihrer Grenzgebiete, Band 5. Springer, Berlin (1937). Reprint. Chelsea, New York (1948). Russian transl.: Uspekhi Mat. Nauk **4**(1(29)), 113–182 (1949)
99. Ibragimov, I.A., Linnik, Yu.V.: Independent and Stationarily Connected Variables. Nauka, Moscow, (1965). English transl.: Groningen, Noordhoff (1971)
100. Il'in, A.M., Khasminskii, R.Z.: Asymptotic behavior of solutions of parabolic equations and an ergodic property of nonhomogeneous diffusion processes. Mat. Sb. **60**(102), 366–392 (1963). English transl.: Amer. Math. Soc. Transl. (2) **49**, 241–268 (1965)
101. Il'in, A.M., Kalashnikov, A.S., Oleinik, O.A.: Second-order linear equations of parabolic type. Uspekhi Mat. Nauk **17**(3(105)), 3–146 (1962). English transl.: Russian Math. Surveys **17**(3), 1–143 (1962)
102. Imkeller, P., Milstein, G.N.: Moment Lyapunov exponent for conservative systems with small periodic and random perturbations. Stochastics and Dynamics **2**(1), 25–48 (2002)
103. Itô, K.: On stochastic differential equations. Mem. Amer. Math. Soc. **4**, 1 (1951). Russian transl.: Matematika **1**(1), 78–116 (1957)
104. Itô, K.: On a formula concerning stochastic differentials. Nagoya Math. J. **3**, 55–65 (1951). Russian transl.: Matematika **3**(5), 131–141 (1959)
105. Itô, K.: Stationary random distributions. Mem. Coll. Sci. Univ. Kyoto Ser. A, Math. **28**, 209–223 (1954)
106. Itô, K., Nisio, M.: On stationary solutions of a stochastic differential equation. J. Math. Kyoto Univ. **4**, 1–75 (1964)
107. Kac, I.Ya.: On stability in the first approximation of systems with random parameters. Ural. Gos. Univ. Mat. Zap. **3**(2), 30–37 (1962) (Russian)
108. Kac, I.Ya.: On the stability of stochastic systems in the large. Prikl. Mat. Meh. **28**, 366–372 (1964). English transl.: J. Appl. Math. Mech. 28, 449–456 (1964)
109. Kac, I.Ya.: Converse of the theorem on the asymptotic stability of stochastic systems. Ural. Gos. Univ. Mat. Zap. **5**(2), 43–51 (1965) (Russian)
110. Kac, I.Ya.: On the stability in the first approximation of systems with random lag. Prikl. Mat. Meh. **31**, 447–452 (1967). English transl.: J. Appl. Math. Mech. **31**, 478–482 (1967)
111. Kac, I.Ya., Krasovskii, N.N.: On stability of systems with random parameters. Prikl. Mat. Meh. **24**, 809–823 (1960). English transl.: J. Appl. Math. Mech. **24**, 1225–1246 (1960)
112. Kalman, R.E.: Control of randomly varying linear dynamical systems. In: Proc. Symp. Appl. Math., vol. 13, pp. 287–298. Am. Math. Soc., Providence (1962)
113. Kamke, E.: Differentialgleichungen. Lösungsmethoden und Lösungen. Teil I. Gëwohnliche Differentialgleichungen, 3rd edn. Math. und Ihrer Anwendungen in Physik und Technik, Band 18. Geest & Portig, Leipzig (1944). Russian transl.: IL, Moscow (1951)
114. Kampé de Fériet, J.: La mecanique statistique des milieux continus. In: Congrès Internat. de Philosophie des Sciences, vol. III Paris, 1949. Actualités Sci. Indust., vol. 1137, pp. 129–142. Hermann, Paris (1951). Russian transl.: Mir, Moscow (1964)
115. Kato, T.: Perturbation Theory of Linear Operators. Springer, Berlin (1966)
116. Kesten, J., Furstenberg, H.: Products of random matrices. Ann. Math. Statist. **31**, 457–469 (1960)
117. Khasminskii, R.Z.: Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations. Teor. Veroyatnost. i Primenen. **5**, 196–214 (1960). English transl.: Theor. Probability Appl. **5**, 179–196 (1960)
118. Khasminskii, R.Z.: On the stability of the trajectory of Markov processes. Prikl. Mat. Meh. **26**, 1025–1032 (1962). English transl.: J. Appl. Math. Mech. **26**, 1554–1565 (1962)

119. Khasminskii, R.Z.: The averaging principle for parabolic and elliptic differential equations and Markov processes with small diffusion. Teor. Veroyatnost. i Primenen. **8**, 3–25 (1963). English transl.: Theor. Probability Appl. **8**, 1–21 (1963)
120. Khasminskii, R.Z.: The behavior of a self-oscillating system acted upon by slight noise. Prikl. Mat. Meh. **27**, 683–688 (1963). English transl.: J. Appl. Math. Mech. **27**, 1035–1044 (1963)
121. Khasminskii, R.Z.: On the dissipativeness of random processes defined by differential equations. Problemy Peredachi Informatsii **1**(1), 88–104 (1965). English transl.: Problems of Information Transmission **1**(1), 63–77 (1965)
122. Khasminskii, R.Z.: On equations with random perturbations. Teor. Veroyatnost. i Primenen. **10**, 394–397 (1965). English transl.: Theor. Probability Appl. **10**, 361–364 (1965)
123. Khasminskii, R.Z.: On the stability of systems under persistent random perturbations. In: Theory of Information Transmission. Pattern Recognition, pp. 74–87. Nauka, Moscow (1965) (Russian)
124. Khasminskii, R.Z.: A limit theorem for the solutions of differential equations with random right-hand sides. Teor. Veroyatnost. i Primenen. **11**, 444–462 (1966). English transl.: Theor. Probability Appl. **11**, 390–406 (1966)
125. Khasminskii, R.Z.: On the stability of nonlinear stochastic systems. Prikl. Mat. Meh. **30**, 915–921 (1966). English transl.: J. Appl. Math. Mech. **30**, 1082–1089 (1966)
126. Khasminskii, R.Z.: Remark on the article of M.G. Shur "On linear differential equations with randomly perturbed parameters". Izv. Akad. Nauk SSSR Ser. Mat. **30**, 1311–1314 (1966) (Russian)
127. Khasminskii, R.Z.: Necessary and sufficient conditions for the asymptotic stability of linear stochastic systems. Teor. Veroyatnost. i Primenen. **12**, 167–172 (1967). English transl.: Theor. Probability Appl. **12**, 144–147 (1967)
128. Khasminskii, R.Z.: Stability in the first approximation for stochastic systems. Prikl. Mat. Meh. **31**, 1021–1027 (1967). J. Appl. Math. Mech. **31**, 1025–1030 (1967)
129. Khasminskii, R.Z.: The limit distribution of additive functionals of diffusion processes. Mat. Zametki **4**, 599–610 (1968) (Russian)
130. Khasminskii, R.Z.: On positive solutions of the equation $Au + v(x)u = 0$. Theory Probability and Applications **4**, 3 (1959)
131. Khasminskii, R.Z.: Arcsine law and one generalization. Acta Applicandae Mathematical **58**, 1–7 (1999)
132. Khasminskii, R.Z.: Limit distributions of some integral functionals for null-recurrent diffusions. Stochastic Processes and Applications **92**, 1–9 (2001)
133. Khasminskii, R.Z.: Stochastic Stability of Differential Equations. Nauka, Moscow (1969). English transl.: Sijthoff and Noordhoff, Alphen, 1980
134. Khasminskii, R.: On robustness of some concepts in stability of stochastic differential equations. Fields Institute Communications **9**, 131–137 (1996)
135. Khasminskii, R., Milstein, G.N.: Stability of gyroscopic systems under small random excitations. Stochastics and Dynamics **4**(1), 107–133 (2004)
136. Khasminskii, R., Moshchuk, N.: Moment Lyapunov exponent and stability index for linear conservative system with small random perturbation. SIAM J. Appl. Math. **58**, 245–256 (1998)
137. Khasminskii, R.Z., Nevelson, M.B.: The stability of the solutions of one-dimensional stochastic equations. Dokl. Akad. Nauk SSSR **200**, 785–788 (1971). English transl.: Soviet Math. Dokl. **200**(4) (1971)
138. Khasminskii, R.Z., Nevelson, M.B.: Stochastic Approximation and Recursive Estimation. Translations of Mathematical Monographs, vol. 47 (1976)
139. Khasminskii, R.Z., Yin, G.: Asymptotic behavior of parabolic equations arising from null-recurrent diffusion. J. Differential Equations **161**, 154–173 (2000)
140. Khasminskii, R.Z., Zhu, C., Yin, G.: Stability of regime-switching diffusions. Stochastic Processes and Applications **117**, 1037–1051 (2007)
141. Khinchin, A.Ya.: Asymptotische Gesetze der Wahrscheinlichkeitsrechnung. Springer, Berlin (1933). Russian transl.: ONTI, Moscow (1936)

142. Kolmanovskii, V.B.: Stationary solutions of equations with delay. Problemy Peredachi Informatsii **3**(1), 64–72 (1967). English transl.: Problems of Information Transmission **1**, 50–57 (1967)

143. Kolmanovskii, V.B.: The stability of non-linear systems with lag. Mat. Zametki **7**, 743–751 (1970) (Russian)

144. Kolmogorov, A.N.: Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. Math. Ann. **104**, 415–458 (1931). Russian transl.: Uspekhi Mat. Nauk **5**, 5–41 (1938)

145. Kolmogorov, A.N.: Markov chains with countably many possible states. Bull. Moskov. Gos. Univ. Sec. A **1**(3), 1–16 (1937) (Russian)

146. Kolmogorov, A.N., Prokhorov, Yu.V.: On sum of random number of random variables. Uspekhi Mat. Nauk **4**(4), 168–172 (1949)

147. Kolomiets, V.G.: Parametric action of a random force on a non-linear oscillatory system. Ukrain. Mat. Zh. **14**, 211–214 (1962) (Russian)

148. Kolomiets, V.G.: Parametric random effect on linear and non-linear oscillatory systems. Ukrain. Mat. Zh. **15**, 199–205 (1963) (Russian)

149. Kolomiets, V.G.: Parametric random oscillations in linear and nonlinear systems. Sb. Dokl. Tashkent. Politekhn. Inst. **6**, 49–59 (1964) (Russian)

150. Kozin, F.: On almost sure stability of linear systems with random coefficients. J. Math. Phys. **42**, 59–67 (1963)

151. Kozin, F.: On relations between moment properties and almost sure Lyapunov stability for linear stochastic systems. J. Math. Anal. Appl. **10**, 342–353 (1965)

152. Kozin, F.: On almost sure asymptotic sample properties of diffusion processes defined by stochastic differential equations. J. Math. Kyoto Univ. **4**, 515–528 (1964/65)

153. Kozin, F.: Stability of the linear stochastic system. In: Lecture Notes in Mathematics, vol. 294, pp. 186–229 (1972)

154. Krasnoselskii, M.A.: The Operator of Translation Along the Trajectories of Differential Equations. Nauka, Moscow (1966). English transl.: Transl. Math. Monographs, vol. 19. Am. Math. Soc., Providence (1968)

155. Krasovskii, A.A.: Sufficient conditions for statistical stability of motion. Izv. Akad. Nauk USSR Tekhn. Kibernet. **2**, 107–113 (1966) (Russian)

156. Krasovskii, N.N.: Certain Problems in the Theory of Stability of Motion. Fizmatgiz, Moscow (1959). English transl.: Stanford University Press, Stanford (1963)

157. Krasovskii, N.N.: Problems in stabilization of controlled motion. Appendix IV to I.G. Malkin [191], 2nd edn., pp. 475–514 (Russian)

158. Krasovskii, N.N.: The stabilization of systems in which the noise depends on the magnitude of the action of the controller. Izv. Akad. Nauk USSR Tekhn. Kibernet. **2**, 102–109 (1965) (Russian)

159. Krasovskii, N.N., Lidskii, E.A.: Analytic design of controllers in systems with random attributes, I. Avtomat. i Telemekh. **22**, 1145–1150 (1961). Automat. Remote Control **22**, 1021–1025 (1961)

160. Krasovskii, N.N., Lidskii, E.A.: Analytic design of controllers in systems with random attributes, II. Avtomat. i Telemekh. **22**, 1273–1278 (1961). Automat. Remote Control **22**, 1145–1146 (1961)

161. Krasovskii, N.N., Lidskii, E.A.: Analytic design of controllers in systems with random attributes, III. Avtomat. i Telemekh. **22**, 1425–1431 (1961). Automat. Remote Control **22**, 1289–1294 (1961)

162. Krasulina, T.P.: The Robbins–Monro process in the case of several roots. Teor. Veroyatnost. i Primenen. **12**, 386–390 (1967). English transl.: Theor. Probability Appl. **12**, 333–337 (1967)

163. Krylov, N.M., Bogolyubov, N.N.: Introduction to Non-linear Mechanics. Princeton University Press, Princeton (1947)

164. Krylov, N.V.: Introduction to the Theory of Diffusion Processes. Am. Math. Soc., Providence (1991)

165. Kushner, H.J.: On the stability of stochastic dynamical systems. Proc. Natl. Acad. Sci. USA **53**, 8–12 (1964)

166. Kushner, H.J.: On the construction of stochastic Lyapunov functions. IEEE Trans. Automatic Control **AC-10**, 477–478 (1965)
167. Kushner, H.J.: Finite time stochastic stability and the analysis of tracking systems. IEEE Trans. Automatic Control **AC-11**, 219–227 (1966)
168. Kushner, H.J.: Stochastic stability and control. In: Math Sci. and Engineering, vol. 33. Academic Press, New York (1967)
169. Kushner, H.J.: Converse theorems for stochastic Lyapunov functions. SIAM J. Control **5**, 228–233 (1967)
170. Kushner, H.J.: The Concept of Invariant Set for Stochastic Dynamical Systems and Applications to Stochastic Stability. Stochastic Optimization and Control. Wiley, New York (1968)
171. Kushner, H.J.: Stochastic stability. In: Lecture Notes in Mathematics, vol. 294, pp. 97–124 (1972)
172. Kushner, H.J.: Stochastic Stability and Control. Academic Press, New York (1967)
173. Kushner, H.J.: On the stability of processes defined by stochastic difference-differential equations. J. Differential Equations **4**, 424–443 (1968)
174. Lamperti, J.: An occupation time theorem for a class of stochastic processes. Trans. Amer. Math. Soc. **88**, 380–387 (1958)
175. Lamperti, J.: Stochastic Processes. Springer, Berlin (1977)
176. Langevin, P.: Sur la théorie du mouvement brownien. C. R. Acad. Sci. Paris **146**, 530–533 (1908)
177. Laning, J.H., Battin, R.H.: Random Processes in Automatic Control. McGraw-Hill, New York (1956)
178. LaSalle, J.P., Lefschetz, S.: Stability by Lyapunov's Direct Methods with Applications. Academic Press, New York (1961). Russian transl.: Mir, Moscow (1964)
179. Lavrentiev, M.A., Shabat, B.V.: Methods of the Theory of Functions of a Complex Variable, 2nd rev. edn. Fizmatgiz, Moscow (1958). German transl.: Mathematik für Naturwiss. und Technik, Band 13, VEB. Deutscher Verlag, Berlin (1967)
180. Leibowitz, M.A.: Statistical behavior of linear systems with randomly varying parameters. J. Math. Phys. **4**, 852–858 (1963)
181. Levit, M.V.: A partial criterion for absolute stochastic stability of non-linear systems of differential equations. Uspekhi Mat. Nauk **4** (1961); **27** (1972) (Russian)
182. Levy, P.: Sur certains processus stochastiques homogines. Compositio Mathematica **7**, 283–339 (1939)
183. Lidskii, E.A.: On the stability of solutions of a stochastic system. In: Proc. Interuniversity Conf. on Applications of Stability Theory, Kazan', pp. 99–102 (1964) (Russian)
184. Linnik, Yu.V.: Decomposition of Probability Distributions. Leningrad State University, Leningrad (1960). English transl.: Oliver and Boyd, Edinburgh/London (1964)
185. Loève, M.: Probability Theory. Foundations. Random sequences. Van Nostrand, Princeton (1960). Russian transl.: IL, Moscow (1962)
186. Lukacs, E.: Characteristic Functions. Griffin, London (1970)
187. Lur'e, A.I.: Minimal quadratic performance index of a controlled system. Izv. Akad. Nauk SSSR Tekhn. Kibernet. **4**, 140–146 (1963) (Russian)
188. Lyapunov, A.M.: Problème général de la stabilité du mouvement, Kharkov, 1892. Reprint. GITTL, Moscow (1950)
189. Malakhov, A.N.: Statistical stability of motion. Izv. Vyssh. Uchebn. Zaved. Radiofizika **6**(1), 42–53 (1963) (Russian)
190. Malkin, I.G.: Das Existenzproblem von Liapounoffschen Funktionen. Izv. Kazan. Fiz.-Mat. Obshch. **5**, 63–84 (1931)
191. Malkin, I.G.: Theory of Stability of Motion, 2nd rev. edn. Nauka, Moscow (1966). English transl.: Atomic Energy Commission Transl. **3352**
192. Mao, X.: Exponential Stability of Stochastic Differential Equations. Dekker, New York (1994)
193. Maruyama, G., Tanaka, H.: Some properties of one-dimensional diffusion processes. Mem. Fac. Sci. Kyushu Univ. Ser. A, Math. **11**, 117–141 (1957)

194. Maruyama, G., Tanaka, H.: Ergodic property of N-dimensional recurrent Markov processes. Mem. Fac. Sci. Kyushu Univ. Ser. A, Math. **13**, 157–172 (1959)

195. McKenna, J., Morrison, J.: Moments of solutions of a class of stochastic differential equations. J. Math. Phys. **12**(10), 2126–2135 (1971)

196. Meyn, S., Tweedie, R.: Markov Chains and Stochastic Stability, 2nd edn. Cambridge Math. Library (2009)

197. Mikhailov, V.P.: A solution of a mixed problem for a parabolic system using the method of potentials. Dokl. Akad. Nauk USSR **132**, 291–294 (1960). English transl.: Soviet Math. Dokl. **1**, 556–559 (1960)

198. Milstein, G.N.: Evaluation of moment Lyapunov exponents for second order stochastic systems. Random & Comput. Dynam. **4**(4), 301–315 (1996)

199. Milstein, G.N.: Stability index for invariant manifolds of stochastic systems. Random & Comput. Dynam. **5**(4), 177–211 (1997)

200. Milstein, G.N.: Orbital stability index for stochastic systems. Stochastic Analysis and Applications **18**(5), 777–809 (2000)

201. Milstein, G.N.: The asymptotic behavior of semi-invariants for linear stochastic systems. Stochastics and Dynamics **2**(2), 281–294 (2002)

202. Miranda, C.: Equazioni alle derivate parziali di tipo ellittico. Ergebnisse der Mathematik und ihrer Grenzgebiete, Heft 2. Springer, Berlin (1955). Russian transl.: IL, Moscow (1957); English transl.: Springer, Berlin (1970)

203. Moiseev, N.N.: On a probability treatment of the concept "stability of motion". Rostov. Gos. Univ. Uch. Zap. Fiz-Mat. Fak. **18**(3), 79–82 (1953) (Russian)

204. Molchanov, S.A.: The structure of eigenfunctions of one-dimensional unordered structures. Math. USSR, Izvestija, **12**, 69–101 (1978)

205. Morozan, T.: La stabilité des solutions des systèmes d'équations différentielles aux paramètres aléatoires. Rev. Roumaine Math. Pures Appl. **11**, 211–238 (1966)

206. Morozan, T.: Stability of some linear stochastic systems. J. Differential Equations **3**, 153–169 (1967)

207. Morozan, T.: Stability of controlled systems with random parameters. Rev. Roumaine Math. Pures Appl. **12**, 545–552 (1967) (Russian)

208. Morozan, T.: Sur l'approximation stochastique. C. R. Acad. Sci. Paris Sér. A–B **264**, A633–A635 (1967)

209. Morozan, T.: Stability in the case of small stochastic perturbations. Rev. Roumaine Math. Pures Appl. **13**, 71–74 (1968) (Russian)

210. Morozan, T.: Stability of stochastic discrete systems. J. Math. Anal. Appl. **23**, 1–9 (1968)

211. Morozan, T.: Stabilitatea sistemelor cu parametri aleatori. Edit. Acad. R.S.R., Bucaresti (1969)

212. Morozan, T.: Boundedness properties for stochastic systems. In: Lecture Notes in Mathematics, vol. 294, pp. 21–34 (1972)

213. Nagaev, S.V.: Ergodic theorems for discrete-time Markov processes. Sibirsk. Mat. Zh. **6**, 413–432 (1965) (Russian)

214. Nagel, R. (ed.): One-parameter Semigroups of Positive Operators. Lecture Notes in Math., vol. 1184. Springer, Berlin (1986)

215. Nevelson, M.B.: Stability in the large of a trajectory of the Markov processes of diffusion-type. Differentsial'nye Uravneniya **2**, 1052–1060 (1966). English transl.: Differential Equations **2**, 544–548 (1966)

216. Nevelson, M.B.: Some remarks on the stability of a linear stochastic system. Prikl. Mat. Meh. **30**, 1124–1127 (1966). English transl.: J. Appl. Math. Mech. **30**, 1332–1335 (1966)

217. Nevelson, M.B.: Behavior of a linear system under small random excitation of its parameters. Prikl. Mat. Meh. **31**, 527–530 (1967). J. Appl. Math. Mech. **31**, 552–555 (1967)

218. Nevelson, M.B.: On the question of the recurrence and ergodicity of a Gaussian Markov process. Teor. Veroyatnost. i Primenen. **14**, 35–43 (1969). English transl.: Theor. Probability Appl. **14**, 35–43 (1969)

219. Nevelson, M.B.: Criterion of existence of an optimal control for a class of linear stochastic systems. Prikl. Mat. Meh. **33**, 573–577 (1969). English transl.: J. Appl. Math. Mech. **33**, 561–565 (1969)
220. Nevelson, M.B., Khasminskii, R.Z.: Stability of a linear system with random disturbances of its parameters. Prikl. Mat. Meh. **30**, 487–493 (1966)
221. Nevelson, M.B., Khasminskii, R.Z.: Stability of stochastic systems. Problemy Peredachi Informatsii **2**(3), 76–91 (1966). English transl.: Problems of Information Transmission **2**(3), 61–74 (1966)
222. Nevelson, M.B., Khasminskii, R.Z.: Stability and stabilization of stochastic differential equations. In: Proc. Sixth Math. Summer School on Probability Theory and Math. Statist., Kaciveli, 1968, pp. 59–122. Izdat. Akad. Nauk Ukrain, SSR., Kiev (1969) (Russian)
223. Oksendal, B.: Stochastic Differential Equations. Springer, Berlin (2007)
224. Pardoux, E., Wihstutz, V.: Lyapunov exponent and rotation number of two-dimensional linear stochastic systems with small diffusion. SIAM J. Appl. Math. **48**, 442–457 (1988)
225. Petrovskii, I.G.: Lectures on Partial Differential Equations. GITTL, Moscow (1950). English transl.: Interscience, New York (1954)
226. Pinsky, M.A.: Instability of the harmonic oscillator with small noise. SIAM J. Appl. Math. **46**, 451–463 (1986)
227. Pinsky, M.A.: Stochastic stability and Dirichlet problem. Comm. Pure Appl. Math. **27**, 311–350 (1974)
228. Pliss, V.A.: Non-local Problems in the Theory of Oscillations. Nauka, Moscow (1964). English transl.: Academic Press, New York (1966)
229. Pontryagin, L.S., Andronov, A.A., Vitt, A.A.: On the statistical investigation of dynamic systems. Zh. Eksper. Teoret. Fiz. **3**(3), 165–180 (1933) (Russian)
230. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., Mishchenko, E.F.: The Mathematical Theory of Optimal Processes. Fizmatgiz, Moscow (1961). English transl.: Wiley, New York (1962). Macmillan, New York (1964)
231. Popov, V.M.: Hiperstabilitatea sistemelor automate. Edit. Acad. R.S.R., Bucaresti (1966)
232. Prokhorov, Yu.V.: Convergence of random processes and limit theorems in probability theory. Teor. Veroyatnost. i Primenen. **1**, 177–238 (1956). English transl.: Theor. Probability Appl. **1**, 106–134 (1956)
233. Pugachev, V.S.: The Theory of Random Functions and Its Application to Control Problems. Fizmatgiz, Moscow (1960). English transl.: Internat. Series of Monographs on Automation and Automatic Control, vol. 5. Pergamon Press, Oxford. Addison-Wesley, Reading (1965)
234. Rabotnikov, Yu.L.: On the impossibility of stabilizing a system in the mean-square by random perturbations of its parameters. Prikl. Mat. Meh. **28**, 935–940 (1964). English transl.: J. Appl. Math. Mech. **28**, 1131–1136 (1964)
235. Rabotnikov, Yu.L.: Boundedness of solutions of differential equations with random coefficients whose averages are constancy. Zap. Mekh.-Mat. Fak. Har'kov. Gos. Univ. i Har'kov. Mat. Obshch. (4) **30**, 75–84 (1964) (Russian)
236. Ramachandran, B.: Advanced Theory of Characteristic Functions. Statistical Publishing Society, Calcutta (1967)
237. Robbins, H., Monro, S.: A stochastic approximation method. Ann. Math. Statist. **22**, 400–407 (1951)
238. Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes, and Martingales, vol. 2: Ito Calculus. Wiley, New York (1987)
239. Romanovskii, Yu.M.: Parametric random perturbations in some aero elasticity problems. Izv. Akad. Nauk USSR Ser. Mekh. Mat. **4**, 133–135 (1960) (Russian)
240. Rosenbloom, A.: Analysis of linear systems with randomly tune-varying parameters. In: Proc. Sympos. on Information Networks, New York, 1954, pp. 145–153. Polytechnic Institute of Brooklyn, Brooklyn (1955)
241. Rozanov, Yu.A.: Stationary Random Processes. Fizmatgiz, Moscow (1963). English transl.: Holden-Day, San Francisco (1967)
242. Rytov, S.M.: Introduction to Statistical Radiophysics. Nauka, Moscow (1966) (Russian)

243. Sagirov, P.: Stochastic Methods in the Dynamics of Satellites, Intern. Center for Mech. Course Lectures, vol. 57. SIAM, Udine (1970)

244. Shaikhet, L.E.: Stability investigations of stochastic systems with lag by the method of Lyapunov functionals. Problemy Peredachi Informatsii **4**(11), 70–75 (1975) (Russian)

245. Samuels, J.C.: On the stability of random systems and the stabilization of deterministic systems. J. Acoust. Soc. Amer. **32**, 594–601 (1960)

246. Samuels, J.C., Eringen, A.C.: On stochastic linear systems. J. Math. Phys. **38**, 83–103 (1959/60)

247. Sazonov, V.V., Tutubalin, V.N.: Probability distributions on topological groups. Teor. Veroyatnost. i Primenen. **11**, 3–55 (1966). English transl.: Theor. Probability Appl. **11**, 1–45 (1966)

248. Šefl, O.: On stability of a randomized linear system. Sci. Sinica **7**, 1027–1034 (1958)

249. Serrin, J.: On the Harnaok inequality for linear elliptic equations. J. Analyse Math. **4**, 292–308 (1955/56). Russian transl.: Matematika **2**(6), 49–62 (1958)

250. Shiryaev, A.N.: On stochastic equations in the theory of conditional Markov processes. Teor. Veroyatnost. i Primenen. **11**, 200–206 (1966). English transl.: Theor. Probability Appl. **11**, 179–184 (1966)

251. Skorokhod, A.V.: Studies in the Theory of Random Processes. Izdat. Kiev. Univ., Kiev (1961). English transl.: Addison-Wesley, Reading (1965)

252. Skorokhod, A.V.: Stochastic Processes with Independent Increments. Nauka, Moscow (1964) (Russian)

253. Stepanov, V.V.: Zur Definition der Stabilitdtswahrscheinlichkeit. C. R. (Dokl.) Acad. Sci. URSS **18**, 151–154 (1938)

254. Stratonovich, R.L.: Topics in the Theory of Random Noise, vol. 1: General Theory of Random Processes. Nonlinear Transformations of Signals and Noise. Sovetskoe Radio, Moscow (1961). English transl.: Gordon & Breach, New York (1963)

255. Stratonovich, R.L.: A new form of representing stochastic integrals and equations. Vestnik Moskov. Univ. Ser. I Mat. Meh. **1**, 3–12 (1964) (Russian)

256. Stratonovich, R.L.: Conditional Markov Processes and Their Application to the Theory of Optimal Control. Izdat. Moskov. Univ., Moscow (1966). English transl.: American Elsevier, New York (1967)

257. Stratonovich, R.L., Romanovskii, Yu.M.: Parametric representation of linear and nonlinear oscillatory systems by a random force. Nauchn. Dokl. Vyssh. Shkoly Fiz.-Mat. Nauki **3**, 221–224 (1958) (Russian)

258. Stroock, D.: Probability Theory, an Analytic View. Cambridge University Press, Cambridge (1993)

259. Shur, M.G.: Linear differential equations with randomly perturbed parameters. Izv. Akad. Nauk USSR Ser. Mat. **29**, 783–806 (1965). English transl.: Amer. Math. Soc. Transl. (2) **72**, 251–276 (1968)

260. Tikhonov, V.I.: Effect of fluctuation action on the simplest parametric systems. Avtom. i Telemekh. **19**, 717–724 (1958). English transl.: Automat. Remote Control **19**, 705–711 (1958)

261. Titchmarsh, E.C.: The Theory of Functions. Oxford University Press, London (1950)

262. Tsar'kov, E.F.: Asymptotic exponential stability in the mean square of the trivial solution of stochastic functional differential equations. Teor. Veroyatnost. i Primenen. **21**, 871–875 (1976) (Russian)

263. Tutubalin, V.N.: Limit theorems for a product of random matrices. Teor. Veroyatnost. i Primenen. **10**, 19–32 (1965); Theor. Probability Appl. **10**, 15–27 (1965)

264. Ueno, T.: The diffusion satisfying Wentzell's boundary condition and Markov process on the boundary. I. Proc. Japan Acad. **36**, 533–538 (1960)

265. Ueno, T.: The diffusion satisfying Wentzell's boundary condition and Markov process on the boundary. II. Proc. Japan Acad. **36**, 625–629 (1960)

266. Ventsel, A.D.: Semigroups of operators that correspond to a generalized differential operator of second-order. Dokl. Akad. Nauk USSR **111**, 269–272 (1956) (Russian)

267. Ventsel, A.D.: On lateral conditions for multidimensional diffusion processes. Teor. Veroyatnost. i Primenen. **4**, 172–185 (1959). Theor. Probability Appl. **4**, 164–177 (1959)

268. Volkonskii, V.A., Rozanov, Yu.A.: Some limit theorems for random functions. Teor. Veroy-atn. I Primenen. **4**, 2 (1959); English translation: Theory Probab. Appl. **4**(2), 178–197 (1959).
269. Vorovich, I.I.: On stability of motion for random disturbances. Izv. Akad. Nauk USSR Ser. Mat. **20**, 17–32 (1956) (Russian)
270. Wang, P.K.C.: On the almost sure stability of linear time-lag systems with stochastic parameters. Internat. J. Control **2**(5), 433–440 (1965)
271. Watanabe, S.: Generalized arc-sine laws for one-dimensional diffusion processes and random walks. Proc. Symp. Pure Math. **57**, 157–174 (1995)
272. Watson, G.N.: A Treatise on the Theory of Bessel Functions, 2nd edn. Cambridge University Press/Macmillan, Cambridge/New York (1944). Russian transl.: IL, Moscow (1947)
273. Willems, J.L.: Mean-square stability criteria for linear white noise stochastic systems. Probl. Control Inf. Theory **2**, 3–4, 199–217 (1973)
274. Willems, J.L.: Lyapunov Functions and Global Frequency Domain Stability Criteria for a Class of Stochastic Feed Back Systems. Lecture Notes in Mathematics, vol. 294 (1972)
275. Wong, E., Zakai, M.: On the relation between ordinary and stochastic differential equations. Internat. J. Engrg. Sci. **3**, 213–229 (1965)
276. Wong, E., Zakai, M.: On the convergence of ordinary integrals to stochastic integrals. Ann. Math. Statist. **36**, 1560–1564 (1965)
277. Wonham, W.M.: Lyapunov criteria for weak stochastic stability. J. Differential Equations **2**, 195–207 (1966)
278. Wonham, W.M.: A Lyapunov method for the estimation of statistical averages. J. Differential Equations **2**, 365–377 (1966)
279. Wonham, W.M.: Lecture Notes on Stochastic Control. Part 1. Brown University, Providence (1967)
280. Wonham, W.M.: Optimal stationary control of a linear system with state-dependent noise. SIAM J. Control **5**, 486–500 (1967)
281. Wonham, W.M.: Random differential equations in control theory. Reprinted from Probabilistic Methods in Applied Mathematics, pp. 131–212. Academic Press, New York (1970)
282. Yakubovich, V.A., Levit, M.V.: Algebraic criteria for stochastic stability of linear systems with parametric control of the type of white noise. Prikl. Mat. Meh. **1**(36), 142–147 (1972) (Russian)
283. Yaglom, A.M.: On the statistical reversibility of Brownian motion. Mat. Sb. **24**(66), 457–492 (1949). German transl.: Abhandlungen aus der Sowjetischen Physik, Folge III, pp. 7–42. Verlag Kultur und Fortschritt, Berlin (1952)
284. Yaglom, A.M.: An introduction to the theory of stationary random functions. Uspekhi Mat. Nauk **7**(5(51)), 3–168 (1952). English transl.: rev. edn., Prentice-Hall, Englewood Cliffs (1962)
285. Yoshizawa, T.: Lyapunov's function and boundedness of solutions. Funkcial. Ekvac. **2**, 95–142 (1959)

Index