C.W. Gardiner

Handbook of Stochastic Methods

for Physics, Chemistry and the Natural Sciences

Second Edition
With 29 Figures



Springer

Professor Dr. Crispin W. Gardiner, Ph. D.
Physics Department, Victoria University of Wellington
Wellington, New Zealand

Series Editor:

Professor Dr. Dr. h.c.mult. Hermann Haken
Institut für Theoretische Physik und Synergetik der Universität Stuttgart
D-70550 Stuttgart, Germany
and
Center for Complex Systems, Florida Atlantic University
Boca Raton, FL 33431, USA

Cip-Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme
Gardiner, Crispin W.:
Handbook of stochastic methods: for physics, chemistry and the natural sciences/C.W. Gardiner. - 2. ed., 4. printing. - Berlin; Heidelberg; New York; Barcelona; Budapest; Hong Kong; London; Milan; Paris; Santa Clara; Singapore; Tokyo: Springer, 1996
(Springer series in synergetics; Vol. 13)
ISBN 3-540-61634-9

NE: GT

2nd Edition 1985
4th Printing 1997

ISSN 0172-7389

ISBN 3-540-61634-9 Study Edition
Springer-Verlag Berlin Heidelberg New York

ISBN 3-540-15607-0 2nd Edition
Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1983, 1985
Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

SPIN 10547672 55/3144 - 5 4 3 2 1 0 - Printed on acid-free paper

Foreword

In the past, the Springer Series in Synergetics has consisted predominantly of *conference proceedings* on this new interdisciplinary field, a circumstance dictated by its rapid growth. As synergetics matures, it becomes more and more desirable to present the relevant experimental and theoretical results in a coherent fashion and to provide students and research workers with fundamental "know-how" by means of texts and monographs.

From the very beginning, we have stressed that the formation of spatial, temporal, or functional structures by complex systems can be adequately dealt with only if stochastic processes are properly taken into account. For this reason, I gave an introduction to these processes in my book *Synergetics. An Introduction*, Volume 1 of this series. But research workers and students wanting to penetrate the theory of these processes more deeply were quite clearly in need of a far more comprehensive text. This gap has been filled by the present book by Professor Crispin Gardiner. It provides a solid basis for forthcoming volumes in the series which draw heavily on the methods and concepts of stochastic processes. These include *Noise-Induced Transitions*, by W. Horsthemke and R. Lefever, *The Kinetic Theory of Electromagnetic Processes*, by Y. L. Klimontovich, and *Concepts and Models of a Quantitative Sociology*, by W. Weidlich and G. Haag.

Though synergetics provides us with rather general concepts, it is by no means "art pour l'art". On the contrary, the processes it deals with are of fundamental importance in self-organizing systems such as those of biology, and in the construction of devices, e.g., in electronics. Therefore I am particularly pleased that the present book has been written by a scientist who has himself applied – and even developed – such methods in the theory of random processes, for example in the fields of quantum optics and chemical reactions. Professor Gardiner's book will prove most useful not only to students and scientists working in synergetics, but also to a much wider audience interested in the theory of random processes and its important applications to a variety of fields.

H. Haken

Preface to the Corrected Printing

Since I started writing this book ten years ago, a great deal has happened. I have been gratified to find how popular my exposition has become, and of course continually bemused that errors still come to light. I am very grateful to all those who have pointed them out to me, in particular to Matthew Collett, Scott Parkins, and Andrew Smith, who, as students and colleagues, over the last five years have kept me aware of everything they noticed. As well, I must also thank Prof. Urbaan Titulaer and Mr. Alexander Kainz, of the Johannes Kepler University of Linz, who sent me a very full and careful list of corrections.

As a consequence a number of corrections have been made in this second printing of the second edition. The most significant of these is the removal of the converse result of Sect. 3.7.3 b, which was incorrectly derived, and which is probably not true.

At this time I must also express my thanks to my wife Helen May and my youngest daughter Nell, who have been of such support in the years since this book was written.

Pasadena, California
October 1989

C. W. Gardiner

Preface to the Second Edition

In this edition I have corrected a number of misprints, and made a few alterations of a more substantial kind. In particular, I have rewritten Sections 4.2.3 and 4.3.6, using a more correct definition of the Stratonovich stochastic integral; I have clarified a slightly confusing exposition on boundaries in Section 5.2.1e; and I have rewritten Sections 6.3.3 and 6.4.4c to take account of recent progress in these fields. I have also slightly augmented the bibliography and references.

Pasadena, California
March 1985

C. W. Gardiner

Preface to the First Edition

My intention in writing this book was to put down in relatively simple language and in a reasonably deductive form, all those formulae and methods which have been scattered throughout the scientific literature on stochastic methods throughout the eighty years that they have been in use. This might seem an unnecessary aim since there are scores of books entitled "Stochastic Processes", and similar titles, but careful perusal of these soon shows that their aim does not coincide with mine. There are purely theoretical and highly mathematical books, there are books related to electrical engineering or communication theory, and there are books for biologists – many of them very good, but none of them covering the kind of applications that appear nowadays so frequently in Statistical Physics, Physical Chemistry, Quantum Optics and Electronics, and a host of other theoretical subjects that form part of the subject area of Synergetics, to which series this book belongs.

The main new point of view here is the amount of space which deals with methods of approximating problems, or transforming them for the purpose of approximating them. I am fully aware that many workers will not see their methods here. But my criterion here has been whether an approximation is *systematic*. Many approximations are based on unjustifiable or uncontrollable assumptions, and are justified *a posteriori*. Such approximations are not the subject of a systematic book – at least, not until they are properly formulated, and their range of validity controlled. In some cases I have been able to put certain approximations on a systematic basis, and they appear here – in other cases I have not. Others have been excluded on the grounds of space and time, and I presume there will even be some that have simply escaped my attention.

A word on the background assumed. The reader must have a good knowledge of practical calculus including contour integration, matrix algebra, differential equations, both ordinary and partial, at the level expected of a first degree in applied mathematics, physics or theoretical chemistry. This is not a text book for a particular course, though it includes matter that has been used in the University of Waikato in a graduate course in physics. It contains material which I would expect any student completing a doctorate in our quantum optics and stochastic processes theory group to be familiar with. There is thus a certain bias towards my own interests, which is the prerogative of an author.

I expect the readership to consist mainly of theoretical physicists and chemists, and thus the general standard is that of these people. This is not a rigorous book in the mathematical sense, but it contains results, all of which I am confident are provable rigorously, and whose proofs can be developed out of the demonstrations given.

The organisation of the book is as in the following table, and might raise some eyebrows. For, after introducing the general properties of Markov processes, I have chosen to base the treatment on the conceptually difficult but intuitively appealing concept of the stochastic differential equation. I do this because of my own experience of the simplicity of stochastic differential equation methods, once one has become familiar with the Ito calculus, which I have presented in Chapter 4 in a rather straightforward manner, such as I have not seen in any previous text. It is true that there is nothing in a stochastic differential equation that is not in a Fokker-Planck equation, but the stochastic differential equation is so much easier to write down and manipulate that only an excessively zealous purist would try to eschew the technique. On the other hand, only similar purists of an opposing camp would try to develop the theory without the Fokker-Planck equation, so Chapter 5 introduces this as a complementary and sometimes overlapping method of handling the same problem. Chapter 6 completes what may be regarded as the "central core" of the book with a treatment of the two main analytical approximation techniques: small noise expansions and adiabatic elimination.

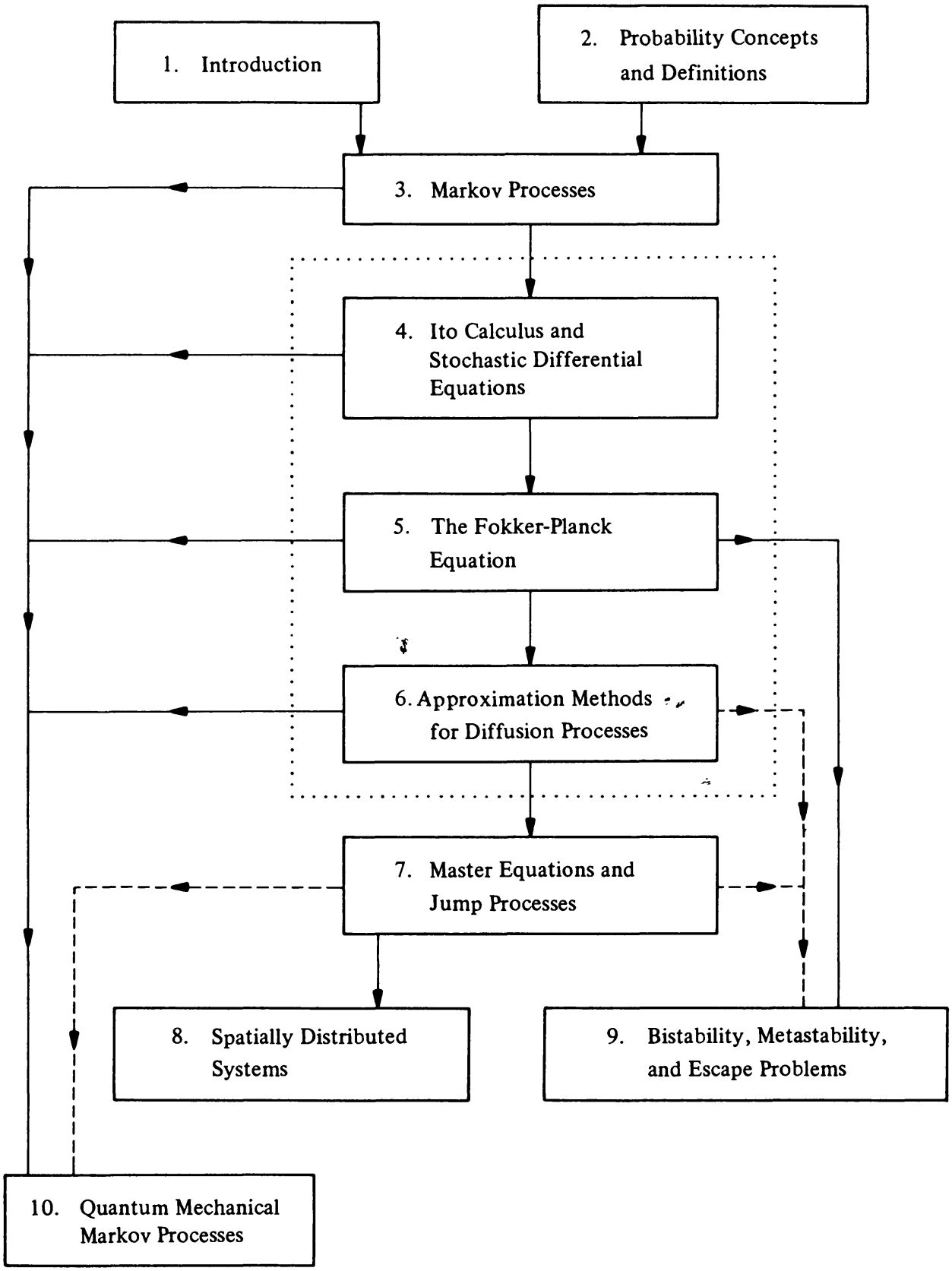
The remainder of the book is built around this core, since very many methods of treating the jump processes in Chapter 7 and the spatially distributed systems, themselves best treated as jump processes, depend on reducing the system to an approximating diffusion process. Thus, although *logically* the concept of a jump process is much simpler than that of a diffusion process, analytically, and in terms of computational methods, the reverse is true.

Chapter 9 is included because of the practical importance of bistability and, as indicated, it is almost independent of all but the first five chapters. Again, I have included only systematic methods, for there is a host of ad hoc methods in this field.

Chapter 10 requires some knowledge of quantum mechanics. I hope it will be of interest to mathematicians who study stochastic processes because there is still much to be done in this field which is of great practical importance and which naturally introduces new realms in stochastic processes – in particular, the rather fascinating field of stochastic processes in the complex plane which turn up as the only way of reducing quantum processes to ordinary stochastic processes. It is with some disappointment that I have noted a tendency among mathematicians to look the other way when quantum Markov processes are mentioned, for there is much to be done here. For example, I know of no treatment of escape problems in quantum Markov systems.

It is as well to give some idea of what is not here. I deal entirely with Markov processes, or systems that can be embedded in Markov processes. This means that no work on non-linear Markovian stochastic differential equations has been included, which I regret. However, van Kampen has covered this field rather well, and it is now well covered in his book on stochastic processes.

Other subjects have been omitted because I feel that they are not yet ready for a definitive formulation. For example, the theory of adiabatic elimination in spatially distributed systems, the theory of fluctuating hydrodynamics, renormalisation group methods in stochastic differential equations, and associated critical phenomena. There is a great body of literature on all of these, and a


1. Introduction
2. Probability Concepts and Definitions
3. Markov Processes
4. Ito Calculus and Stochastic Differential Equations
5. The Fokker-Planck Equation
6. Approximation Methods for Diffusion Processes
7. Master Equations and Jump Processes
8. Spatially Distributed Systems
9. Bistability, Metastability, and Escape Problems
10. Quantum Mechanical Markov Processes

Further, for the sake of compactness and simplicity I have normally presented only one way of formulating certain methods. For example, there are several different ways of formulating the adiabatic elimination results, though few have been used in this context. My formulation of quantum Markov processes and the use of P-representations is only one of many. To have given a survey of all formulations would have required an enormous and almost unreadable book. However, where appropriate I have included specific references, and further relevant matter can be found in the general bibliography.

Hamilton, New Zealand *C. W. Gardiner*
January, 1983

Acknowledgements

My warmest appreciation must go to Professor Hermann Haken for inviting me to write this book for the Springer Series in Synergetics, and for helping support a sabbatical leave in Stuttgart where I did most of the initial exploration of the subject and commenced writing the book.

The physical production of the manuscript would not have been possible without the thoroughness of Christine Coates, whose ability to produce a beautiful typescript, in spite of my handwriting and changes of mind, has never ceased to arouse my admiration. The thorough assistance of Moira Steyn-Ross in checking formulae and the consistency of the manuscript has been a service whose essential nature can only be appreciated by an author.

Many of the diagrams, and some computations, were prepared with the assistance of Craig Savage, for whose assistance I am very grateful.

Since I first became interested in stochastic phenomena, I have benefitted greatly from contact with a large number of people, and in particular I wish to thank L. Arnold, R. Graham, S. Grossman, F. Haake, P. Hohenberg, W. Horsthemke, N. G. van Kampen, R. Landauer, R. Lefever, M. Malek-Mansour, G. Nicolis, A. Nitzan, P. Ortoleva, J. Ross, F. Schlögl, and U. Titulaer.

To my colleagues, students and former students at the University of Waikato must go a considerable amount of credit for much of the work in this book; in particular to Bruce Liley, whose encouragement and provision of departmental support has been warmly appreciated. I want to express my appreciation to Dan Walls who first introduced me to this field, and with whom I have enjoyed a fruitful collaboration for many years; to Howard Carmichael, Peter Drummond, Ken McNeil, Gerard Milburn, Moira Steyn-Ross, and above all, to Subhash Chaturvedi, whose insights into and knowledge of this field have been of particular value.

It is my wish to dedicate this book to my wife, Heather, and my children, Simon and Amanda, who have been remarkably patient of my single minded application of my time to this project.

The extract from the Paper by A. Einstein which appears in Sect. 1.2.1 is reprinted with the permission of the Hebrew University, Jerusalem, Israel, who hold the copyright.

The diagram which appears as Fig. 1.3(b) is reprinted with permission of Princeton University Press.

Contents

1. A Historical Introduction

1.1 Motivation 1
1.2 Some Historical Examples 2
1.2.1 Brownian Motion 2
1.2.2 Langevin's Equation 6
1.3 Birth-Death Processes 8
1.4 Noise in Electronic Systems 11
1.4.1 Shot Noise 11
1.4.2 Autocorrelation Functions and Spectra 15
1.4.3 Fourier Analysis of Fluctuating Functions: Stationary Systems 17
1.4.4 Johnson Noise and Nyquist's Theorem 18

2. Probability Concepts

2.1 Events, and Sets of Events 21
2.2 Probabilities 22
2.2.1 Probability Axioms 22
2.2.2 The Meaning of $P(A)$ 23
2.2.3 The Meaning of the Axioms 23
2.2.4 Random Variables 24
2.3 Joint and Conditional Probabilities: Independence 25
2.3.1 Joint Probabilities 25
2.3.2 Conditional Probabilities 25
2.3.3 Relationship Between Joint Probabilities of Different Orders 26
2.3.4 Independence 27
2.4 Mean Values and Probability Density 28
2.4.1 Determination of Probability Density by Means of Arbitrary Functions 28
2.4.2 Sets of Probability Zero 29
2.5 Mean Values 29
2.5.1 Moments, Correlations, and Covariances 30
2.5.2 The Law of Large Numbers 30
2.6 Characteristic Function 32
2.7 Cumulant Generating Function: Correlation Functions and Cumulants 33
2.7.1. Example: Cumulant of Order 4: $\langle\langle X_1 X_2 X_3 X_4 \rangle\rangle$ 35
2.7.2 Significance of Cumulants 35

2.8 Gaussian and Poissonian Probability Distributions 36
2.8.1 The Gaussian Distribution 36
2.8.2 Central Limit Theorem 37
2.8.3 The Poisson Distribution 38
2.9 Limits of Sequences of Random Variables 39
2.9.1 Almost Certain Limit 40
2.9.2 Mean Square Limit (Limit in the Mean) 40
2.9.3 Stochastic Limit, or Limit in Probability 40
2.9.4 Limit in Distribution 41
2.9.5 Relationship Between Limits 41

3. Markov Processes
3.1 Stochastic Processes 42
3.2 Markov Process 43
3.2.1 Consistency – the Chapman-Kolmogorov Equation 43
3.2.2 Discrete State Spaces 44
3.2.3 More General Measures 44
3.3 Continuity in Stochastic Processes 45
3.3.1 Mathematical Definition of a Continuous Markov Process .. 46
3.4 Differential Chapman-Kolmogorov Equation 47
3.4.1 Derivation of the Differential Chapman-Kolmogorov Equation 48
3.4.2 Status of the Differential Chapman-Kolmogorov Equation . 51
3.5 Interpretation of Conditions and Results 51
3.5.1 Jump Processes: The Master Equation 52
3.5.2 Diffusion Processes – the Fokker-Planck Equation 52
3.5.3 Deterministic Processes – Liouville's Equation 53
3.5.4 General Processes 54
3.6 Equations for Time Development in Initial Time – Backward Equations 55
3.7 Stationary and Homogeneous Markov Processes 56
3.7.1 Ergodic Properties 57
3.7.2 Homogeneous Processes 60
3.7.3 Approach to a Stationary Process 61
3.7.4 Autocorrelation Function for Markov Processes 64
3.8 Examples of Markov Processes 66
3.8.1 The Wiener Process 66
3.8.2 The Random Walk in One Dimension 70
3.8.3 Poisson Process 73
3.8.4 The Ornstein-Uhlenbeck Process 74
3.8.5 Random Telegraph Process 77

4. The Ito Calculus and Stochastic Differential Equations
4.1 Motivation 80
4.2 Stochastic Integration 83

4.2.1 Definition of the Stochastic Integral 83
4.2.2 Example $\int_{t_0}^{t} W(t')dW(t')$ 84
4.2.3 The Stratonovich Integral 86
4.2.4 Nonanticipating Functions 86
4.2.5 Proof that $dW(t)^2 = dt$ and $dW(t)^{2+N} = 0$ 87
4.2.6 Properties of the Ito Stochastic Integral 88
4.3 Stochastic Differential Equations (SDE) 92
4.3.1 Ito Stochastic Differential Equation: Definition 93
4.3.2 Markov Property of the Solution of an Ito Stochastic Differential Equation 95
4.3.3 Change of Variables: Ito's Formula 95
4.3.4 Connection Between Fokker-Planck Equation and Stochastic Differential Equation 96
4.3.5 Multivariable Systems 97
4.3.6 Stratonovich's Stochastic Differential Equation 98
4.3.7 Dependence on Initial Conditions and Parameters 101
4.4 Some Examples and Solutions 102
4.4.1 Coefficients Without x Dependence 102
4.4.2 Multiplicative Linear White Noise Process 103
4.4.3 Complex Oscillator with Noisy Frequency 104
4.4.4 Ornstein-Uhlenbeck Process 106
4.4.5 Conversion from Cartesian to Polar Coordinates 107
4.4.6 Multivariate Ornstein-Uhlenbeck Process 109
4.4.7 The General Single Variable Linear Equation 112
4.4.8 Multivariable Linear Equations 114
4.4.9 Time-Dependent Ornstein-Uhlenbeck Process 115

5. The Fokker-Planck Equation
5.1 Background 117
5.2 Fokker-Planck Equation in One Dimension 118
5.2.1 Boundary Conditions 118
5.2.2 Stationary Solutions for Homogeneous Fokker-Planck Equations 124
5.2.3 Examples of Stationary Solutions 126
5.2.4 Boundary Conditions for the Backward Fokker-Planck Equation 128
5.2.5 Eigenfunction Methods (Homogeneous Processes) 129
5.2.6 Examples 132
5.2.7 First Passage Times for Homogeneous Processes 136
5.2.8 Probability of Exit Through a Particular End of the Interval 142
5.3 Fokker-Planck Equations in Several Dimensions 143
5.3.1 Change of Variables 144
5.3.2 Boundary Conditions 146
5.3.3 Stationary Solutions: Potential Conditions 146
5.3.4 Detailed Balance 148

5.3.5 Consequences of Detailed Balance 150
5.3.6 Examples of Detailed Balance in Fokker-Planck Equations . 155
5.3.7 Eigenfunction Methods in Many Variables – Homogeneous Processes 165
5.4 First Exit Time from a Region (Homogeneous Processes) 170
5.4.1 Solutions of Mean Exit Time Problems 171
5.4.2 Distribution of Exit Points 174

6. Approximation Methods for Diffusion Processes
6.1 Small Noise Perturbation Theories 177
6.2 Small Noise Expansions for Stochastic Differential Equations 180
6.2.1 Validity of the Expansion 182
6.2.2 Stationary Solutions (Homogeneous Processes) 183
6.2.3 Mean, Variance, and Time Correlation Function 184
6.2.4 Failure of Small Noise Perturbation Theories 185
6.3 Small Noise Expansion of the Fokker-Planck Equation 187
6.3.1 Equations for Moments and Autocorrelation Functions 189
6.3.2 Example 192
6.3.3 Asymptotic Method for Stationary Distributions 194
6.4 Adiabatic Elimination of Fast Variables 195
6.4.1 Abstract Formulation in Terms of Operators and Projectors 198
6.4.2 Solution Using Laplace Transform 200
6.4.3 Short-Time Behaviour 203
6.4.4 Boundary Conditions 205
6.4.5 Systematic Perturbative Analysis 206
6.5 White Noise Process as a Limit of Nonwhite Process 210
6.5.1 Generality of the Result 215
6.5.2 More General Fluctuation Equations 215
6.5.3 Time Nonhomogeneous Systems 216
6.5.4 Effect of Time Dependence in L_1 217
6.6 Adiabatic Elimination of Fast Variables: The General Case 218
6.6.1 Example: Elimination of Short-Lived Chemical Intermediates 218
6.6.2 Adiabatic Elimination in Haken's Model 223
6.6.3 Adiabatic Elimination of Fast Variables: A Nonlinear Case 227
6.6.4 An Example with Arbitrary Nonlinear Coupling 232

7. Master Equations and Jump Processes
7.1 Birth-Death Master Equations – One Variable 236
7.1.1 Stationary Solutions 236
7.1.2 Example: Chemical Reaction $X \rightleftharpoons A$ 238
7.1.3 A Chemical Bistable System 241
7.2 Approximation of Master Equations by Fokker-Planck Equations 246
7.2.1 Jump Process Approximation of a Diffusion Process 246

7.2.2 The Kramers-Moyal Expansion 249
7.2.3 Van Kampen's System Size Expansion 250
7.2.4 Kurtz's Theorem 254
7.2.5 Critical Fluctuations 255
7.3 Boundary Conditions for Birth-Death Processes 257
7.4 Mean First Passage Times 259
7.4.1 Probability of Absorption 261
7.4.2 Comparison with Fokker-Planck Equation 261
7.5 Birth-Death Systems with Many Variables 262
7.5.1 Stationary Solutions when Detailed Balance Holds 263
7.5.2 Stationary Solutions Without Detailed Balance (Kirchoff's Solution) 266
7.5.3 System Size Expansion and Related Expansions 266
7.6 Some Examples 267
7.6.1 $X + A \rightleftharpoons 2X$ 267
7.6.2 $X \underset{k}{\overset{\gamma}{\rightleftharpoons}} Y \underset{\gamma}{\overset{k}{\rightleftharpoons}} A$ 267
7.6.3 Prey-Predator System 268
7.6.4 Generating Function Equations 273
7.7 The Poisson Representation 277
7.7.1 Kinds of Poisson Representations 282
7.7.2 Real Poisson Representations 282
7.7.3 Complex Poisson Representations 282
7.7.4 The Positive Poisson Representation 285
7.7.5 Time Correlation Functions 289
7.7.6 Trimolecular Reaction 294
7.7.7. Third-Order Noise 299

8. Spatially Distributed Systems
8.1 Background 303
8.1.1 Functional Fokker-Planck Equations 305
8.2 Multivariate Master Equation Description 307
8.2.1 Diffusion 307
8.2.2 Continuum Form of Diffusion Master Equation 308
8.2.3 Reactions and Diffusion Combined 313
8.2.4 Poisson Representation Methods 314
8.3 Spatial and Temporal Correlation Structures 315
8.3.1 Reaction $X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} Y$ 315
8.3.2 Reactions $B + X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} C$, $A + X \underset{k_2}{\rightarrow} 2X$ 319
8.3.3 A Nonlinear Model with a Second-Order Phase Transition .. 324
8.4 Connection Between Local and Global Descriptions 328
8.4.1 Explicit Adiabatic Elimination of Inhomogeneous Modes ... 328
8.5 Phase-Space Master Equation 331
8.5.1 Treatment of Flow 331
8.5.2 Flow as a Birth-Death Process 332
8.5.3 Inclusion of Collisions – the Boltzmann Master Equation .. 336
8.5.4 Collisions and Flow Together 339

9. Bistability, Metastability, and Escape Problems
9.1 Diffusion in a Double-Well Potential (One Variable) 342
9.1.1 Behaviour for $D = 0$ 343
9.1.2 Behaviour if D is Very Small 343
9.1.3 Exit Time 345
9.1.4 Splitting Probability 345
9.1.5 Decay from an Unstable State 347
9.2 Equilibration of Populations in Each Well 348
9.2.1 Kramers' Method 349
9.2.2 Example: Reversible Denaturation of Chymotrypsinogen .. 352
9.2.3 Bistability with Birth-Death Master Equations (One Variable) 354
9.3 Bistability in Multivariable Systems 357
9.3.1 Distribution of Exit Points 357
9.3.2 Asymptotic Analysis of Mean Exit Time 362
9.3.3 Kramers' Method in Several Dimensions 363
9.3.4 Example: Brownian Motion in a Double Potential 366

10. Quantum Mechanical Markov Processes
10.1 Quantum Mechanics of the Harmonic Oscillator 373
10.1.1 Interaction with an External Field 375
10.1.2 Properties of Coherent States 376
10.2 Density Matrix and Probabilities 380
10.2.1 Von Neumann's Equation 382
10.2.2 Glauber-Sudarshan P-Representation 382
10.2.3 Operator Correspondences 383
10.2.4 Application to the Driven Harmonic Oscillator 384
10.2.5 Quantum Characteristic Function 386
10.3 Quantum Markov Processes 388
10.3.1 Heat Bath 388
10.3.2 Correlations of Smooth Functions of Bath Operators ... 389
10.3.3 Quantum Master Equation for a System Interacting with a Heat Bath 390
10.4 Examples and Applications of Quantum Markov Processes 395
10.4.1 Harmonic Oscillator 395
10.4.2 The Driven Two-Level Atom 399
10.5 Time Correlation Functions in Quantum Markov Processes 402
10.5.1 Quantum Regression Theorem 404
10.5.2 Application to Harmonic Oscillator in the P-Representation 405
10.5.3 Time Correlations for Two-Level Atom 408
10.6 Generalised P-Representations 408
10.6.1 Definition of Generalised P-Representation 409
10.6.2 Existence Theorems 411
10.6.3 Relation to Poisson Representation 413
10.6.4 Operator Identities 414

10.7 Application of Generalised *P*-Representations to Time-Development Equations 415
10.7.1 Complex *P*-Representation 416
10.7.2 Positive *P*-Representation 416
10.7.3 Example 418

References 421

Bibliography 427

Symbol Index 431

Author Index 435

Subject Index 437

1. A Historical Introduction

1.1 Motivation

Theoretical science up to the end of the nineteenth century can be viewed as the study of solutions of differential equations and the modelling of natural phenomena by deterministic solutions of these differential equations. It was at that time commonly thought that if all initial data could only be collected, one would be able to predict the future with certainty.

We now know this is not so, in at least two ways. Firstly, the advent of quantum mechanics within a quarter of a century gave rise to a new physics, and hence a new theoretical basis for all science, which had as an essential basis a purely statistical element. Secondly, more recently, the concept of chaos has arisen, in which even quite simple differential equation systems have the rather alarming property of giving rise to essentially unpredictable behaviour. To be sure, one can predict the future of such a system given its initial conditions, but any error in the initial conditions is so rapidly magnified that no practical predictability is left. In fact, the existence of chaos is really not surprising, since it agrees with more of our everyday experience than does pure predictability—but it is surprising perhaps that it has taken so long for the point to be made.

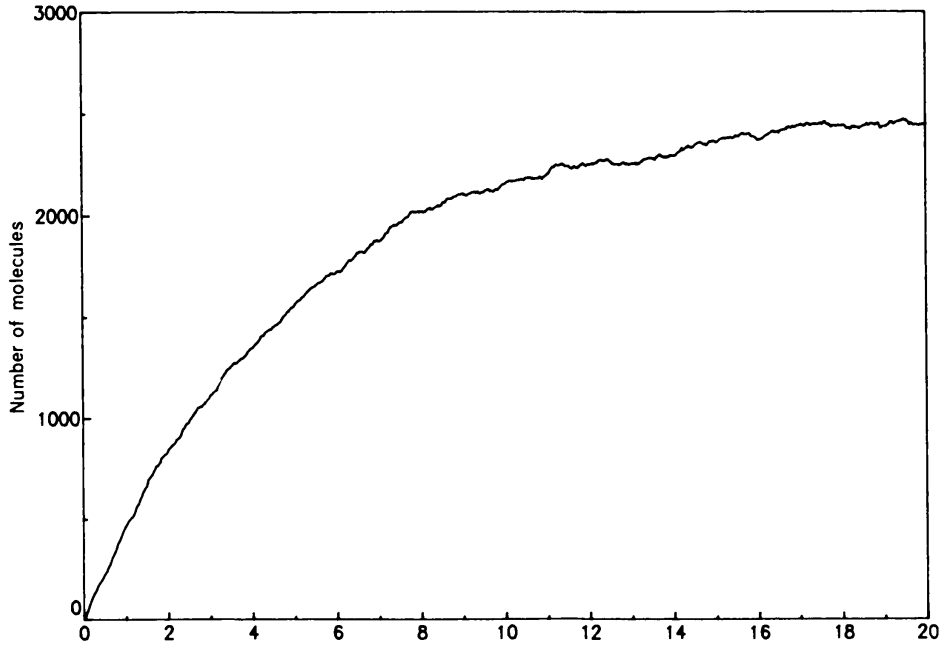


Fig. 1.1. Stochastic simulation of an isomerisation reaction $X \rightleftharpoons A$

Chaos and quantum mechanics are not the subject of this chapter. Here I wish to give a semihistorical outline of how a phenomenological theory of fluctuating phenomena arose and what its essential points are. The very usefulness of predictable models indicates that life is not entirely chaos. But there is a limit to predictability, and what we shall be most concerned with in this book are models of limited predictability. The experience of careful measurements in science normally gives us data like that of Fig. 1.1, representing the growth of the number of molecules of a substance X formed by a chemical reaction of the form $X \rightleftharpoons A$. A quite well defined deterministic motion is evident, and this is reproducible, unlike the fluctuations around this motion, which are not.

1.2 Some Historical Examples

1.2.1 Brownian Motion

The observation that, when suspended in water, small pollen grains are found to be in a very animated and irregular state of motion, was first systematically investigated by Robert Brown in 1827, and the observed phenomenon took the name *Brownian Motion* because of his fundamental pioneering work. Brown was a botanist—indeed a very famous botanist—and of course tested whether this motion was in some way a manifestation of life. By showing that the motion was present in any suspension of fine particles—glass, minerals and even a fragment of the sphinx—he ruled out any specifically organic origin of this motion. The motion is illustrated in Fig. 1.2.

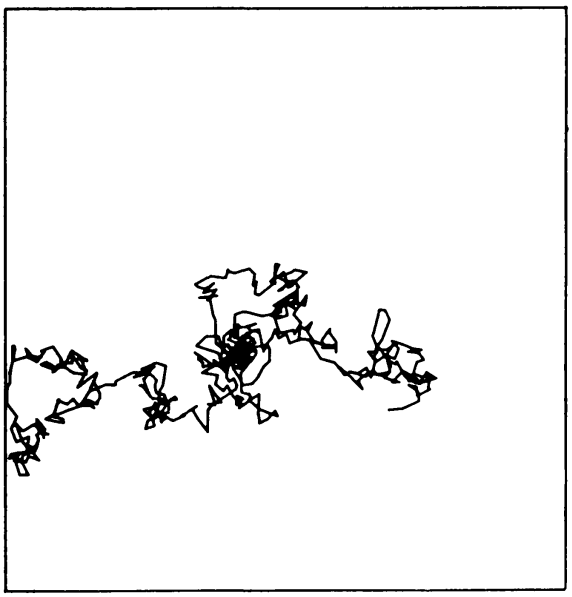


Fig. 1.2. Motion of a point undergoing Brownian motion

The riddle of Brownian motion was not quickly solved, and a satisfactory explanation did not come until 1905, when *Einstein* published an explanation under the rather modest title "über die von der molekular-kinetischen Theorie der

Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen" (concerning the motion, as required by the molecular-kinetic theory of heat, of particles suspended in liquids at rest) [1.2]. The same explanation was independently developed by *Smoluchowski* [1.3], who was responsible for much of the later systematic development and for much of the experimental verification of Brownian motion theory.

There were two major points in Einstein's solution to the problem of Brownian motion.

(i) The motion is caused by the exceedingly frequent impacts on the pollen grain of the incessantly moving molecules of liquid in which it is suspended.
(ii) The motion of these molecules is so complicated that its effect on the pollen grain can only be described probabilistically in terms of exceedingly frequent statistically independent impacts.

The existence of fluctuations like these ones calls out for a statistical explanation of this kind of phenomenon. Statistics had already been used by Maxwell and Boltzmann in their famous gas theories, but only as a description of possible states and the likelihood of their achievement and not as an intrinsic part of the time evolution of the system. *Rayleigh* [1.1] was in fact the first to consider a statistical description in this context, but for one reason or another, very little arose out of his work. For practical purposes, Einstein's explanation of the nature of Brownian motion must be regarded as the beginning of stochastic modelling of natural phenomena.

Einstein's reasoning is very clear and elegant. It contains all the basic concepts which will make up the subject matter of this book. Rather than paraphrase a classic piece of work, I shall simply give an extended excerpt from Einstein's paper (author's translation):

"It must clearly be assumed that each individual particle executes a motion which is independent of the motions of all other particles; it will also be considered that the movements of one and the same particle in different time intervals are independent processes, as long as these time intervals are not chosen too small.

"We introduce a time interval τ into consideration, which is very small compared to the observable time intervals, but nevertheless so large that in two successive time intervals τ, the motions executed by the particle can be thought of as events which are independent of each other.

"Now let there be a total of n particles suspended in a liquid. In a time interval τ, the X-coordinates of the individual particles will increase by an amount Δ, where for each particle Δ has a different (positive or negative) value. There will be a certain *frequency law* for Δ; the number dn of the particles which experience a shift which is between Δ and $\Delta + d\Delta$ will be expressible by an equation of the form

$$dn = n\phi(\Delta)d\Delta, \tag{1.2.1}$$

where

$$\int_{-\infty}^{\infty} \phi(\Delta)d\Delta = 1 \tag{1.2.2}$$

and ϕ is only different from zero for very small values of Δ, and satisifes the condition

$$\phi(\Delta) = \phi(-\Delta). \tag{1.2.3}$$

"We now investigate how the diffusion coefficient depends on ϕ. We shall once more restrict ourselves to the case where the number ν of particles per unit volume depends only on x and t.

"Let $\nu = f(x, t)$ be the number of particles per unit volume. We compute the distribution of particles at the time $t + \tau$ from the distribution at time t. From the definition of the function $\phi(\Delta)$, it is easy to find the number of particles which at time $t + \tau$ are found between two planes perpendicular to the x-axis and passing through points x and $x + dx$. One obtains

$$f(x, t + \tau)dx = dx \int_{-\infty}^{\infty} f(x + \Delta, t)\phi(\Delta)d\Delta . \tag{1.2.4}$$

But since τ is very small, we can set

$$f(x, t + \tau) = f(x, t) + \tau \frac{\partial f}{\partial t} . \tag{1.2.5}$$

Furthermore, we develop $f(x + \Delta, t)$ in powers of Δ:

$$f(x + \Delta, t) = f(x, t) + \Delta \frac{\partial f(x, t)}{\partial x} + \frac{\Delta^2}{2!} \frac{\partial^2 f(x, t)}{\partial x^2} + \cdots . \tag{1.2.6}$$

We can use this series under the integral, because only small values of Δ contribute to this equation. We obtain

$$f + \frac{\partial f}{\partial \tau}\tau = f \int_{-\infty}^{\infty} \phi(\Delta)d\Delta + \frac{\partial f}{\partial x} \int_{-\infty}^{\infty} \Delta\phi(\Delta)d\Delta + \frac{\partial^2 f}{\partial x^2} \int_{-\infty}^{\infty} \frac{\Delta^2}{2} \phi(\Delta)d\Delta . \tag{1.2.7}$$

Because $\phi(x) = \phi(-x)$, the second, fourth, etc., terms on the right-hand side vanish, while out of the 1st, 3rd, 5th, etc., terms, each one is very small compared with the previous. We obtain from this equation, by taking into consideration

$$\int_{-\infty}^{\infty} \phi(\Delta)d\Delta = 1 \tag{1.2.8}$$

and setting

$$\frac{1}{\tau} \int_{-\infty}^{\infty} \frac{\Delta^2}{2} \phi(\Delta)d\Delta = D , \tag{1.2.9}$$

and keeping only the 1st and third terms of the right-hand side,

$$\frac{\partial f}{\partial t} = D \frac{\partial^2 f}{\partial x^2} \cdots . \tag{1.2.10}$$

This is already known as the differential equation of diffusion and it can be seen that D is the diffusion coefficient. ...

"The problem, which corresponds to the problem of diffusion from a single point (neglecting the interaction between the diffusing particles), is now completely determined mathematically: its solution is

$$f(x, t) = \frac{n}{\sqrt{4\pi D}} \frac{e^{-x^2/4Dt}}{\sqrt{t}} \cdot \ldots \tag{1.2.11}$$

"We now calculate, with the help of this equation, the displacement λ_x in the direction of the X-axis that a particle experiences on the average or, more exactly, the square root of the arithmetic mean of the square of the displacement in the direction of the X-axis; it is

$$\lambda_x = \sqrt{\bar{x}^2} = \sqrt{2Dt}\,." \tag{1.2.12}$$

Einstein's derivation is really based on a discrete time assumption, that impacts happen only at times $0, \tau, 2\tau, 3\tau \ldots$, and his resulting equation (1.2.10) for the distribution function $f(x, t)$ and its solution (1.2.11) are to be regarded as approximations, in which τ is considered so small that t may be considered as being continuous. Nevertheless, his description contains very many of the major concepts which have been developed more and more generally and rigorously since then, and which will be central to this book. For example:

i) *The Chapman-Kolmogorov Equation* occurs as Einstein's equation (1.2.4). It states that the probability of the particle being at point x at time $t + \tau$ is given by the sum of the probability of all possible "pushes" Δ from positions $x + \Delta$, multiplied by the probability of being at $x + \Delta$ at time t. This assumption is based on the independence of the push Δ of any previous history of the motion: it is only necessary to know the initial position of the particle at time t—not at any previous time. This is the *Markov postulate* and the Chapman Kolmogorov equation, of which (1.2.4) is a special form, is the central dynamical equation to all Markov processes. These will be studied in detail in Chap. 3.

ii) *The Fokker-Planck Equation:* Eq. (1.2.10) is the diffusion equation, a special case of the Fokker-Planck equation, which describes a large class of very interesting stochastic processes in which the system has a continuous sample path. In this case, that means that the pollen grain's position, if thought of as obeying a probabilistic law given by solving the diffusion equation (1.2.10), in which time t is continuous (not discrete, as assumed by Einstein), can be written $x(t)$, where $x(t)$ is a *continuous function of time*-but a random function. This leads us to consider the possibility of describing the dynamics of the system in some direct probabilistic way, so that we would have a *random* or *stochastic differential equation for the path.* This procedure was initiated by Langevin with the famous equation that to this day bears his name. We will discuss this in detail in Chap. 4.

iii) *The Kramers-Moyal* and similar expansions are essentially the same as that used by Einstein to go from (1.2.4) (the Chapman-Kolmogorov equation) to the

diffusion equation (1.2.10). The use of this type of approximation, which effectively replaces a process whose sample paths need not be continuous with one whose paths are continuous, has been a topic of discussion in the last decade. Its use and validity will be discussed in Chap. 7.

1.2.2 Langevin's Equation

Some time after Einstein's original derivation, *Langevin* [1.4] presented a new method which was quite different from Einstein's and, according to him, "infinitely more simple." His reasoning was as follows.

From statistical mechanics, it was known that the mean kinetic energy of the Brownian particle should, in equilibrium, reach a value

$$\langle \tfrac{1}{2} m v^2 \rangle = \tfrac{1}{2} kT \tag{1.2.13}$$

(T; absolute temperature, k; Boltzmann's constant). (Both Einstein and Smoluchowski had used this fact). Acting on the particle, of mass m there should be two forces:

i) a viscous drag: assuming this is given by the same formula as in macroscopic hydrodynamics, this is $-6\pi\eta a\, dx/dt$, η being the viscosity and a the diameter of the particle, assumed spherical.

ii) another *fluctuating force* X which represents the incessant impacts of the molecules of the liquid on the Brownian particle. All that is known about it is that fact, and that it should be positive and negative with equal probability. Thus, the equation of motion for the position of the particle is given by Newton's law as

$$m \frac{d^2x}{dt^2} = -6\pi\eta a \frac{dx}{dt} + X \tag{1.2.14}$$

and multiplying by x, this can be written

$$\frac{m}{2} \frac{d^2}{dt^2}(x^2) - mv^2 = -3\pi\eta a \frac{d(x^2)}{dt} + Xx\,, \tag{1.2.15}$$

where $v = dx/dt$. We now average over a large number of different particles and use (1.2.13) to obtain an equation for $\langle x^2 \rangle$:

$$\frac{m}{2} \frac{d^2\langle x^2 \rangle}{dt^2} + 3\pi\eta a \frac{d\langle x^2 \rangle}{dt} = kT\,, \tag{1.2.16}$$

where the term $\langle xX \rangle$ has been set equal to zero because (to quote Langevin) "of the irregularity of the quantity X". One then finds the general solution

$$\frac{d\langle x^2 \rangle}{dt} = kT/(3\pi\eta a) + C \exp(-6\pi\eta a t/m)\,, \tag{1.2.17}$$

where C is an arbitrary constant. Langevin estimated that the decaying exponential approaches zero with a time constant of the order of 10^{-8} s, which for any practical observation at that time, was essentially immediately. Thus, for practical purposes, we can neglect this term and integrate once more to get

$$\langle x^2 \rangle - \langle x_0^2 \rangle = [kT/(3\pi\eta a)]t \,. \tag{1.2.18}$$

This corresponds to (1.2.12) as deduced by Einstein, provided we identify

$$D = kT/(6\pi\eta a) \,, \tag{1.2.19}$$

a result which Einstein derived in the same paper but by independent means.

Langevin's equation was the first example of the *stochastic differential equation*—a differential equation with a random term X and hence whose solution is, in some sense, a random function. Each solution of Langevin's equation represents a different random trajectory and, using only rather simple properties of X (his fluctuating force), measurable results can be derived.

One question arises: Einstein explicitly required that (on a sufficiently large time scale) the change Δ be completely independent of the preceding value of Δ. Langevin did not mention such a concept explicitly, but it is there, implicitly, when one sets $\langle Xx \rangle$ equal to zero. The concept that X is extremely irregular *and* (which is not mentioned by Langevin, but is implicit) that X and x are *independent* of each other—that the irregularities in x as a function of time, do not somehow conspire to be always in the same direction as those of X, so that the product could possibly not be set equal to zero; these are really equivalent to Einstein's independence assumption. The method of Langevin equations is clearly very much more direct, at least at first glance, and gives a very natural way of generalising a dynamical equation to a probabilistic equation. An adequate mathematical grounding for the approach of Langevin, however, was not available until more than 40 years later, when Ito formulated his concepts of stochastic differential equations. And in this formulation, a precise statement of the independence of X and x led to the calculus of stochastic differentials, which now bears his name, and which will be fully developed in Chap. 4.

As a physical subject, Brownian motion had its heyday in the first two decades of this century, when Smoluchowski in particular, and many others carried out extensive theoretical and experimental investigations, which showed complete agreement with the original formulation of the subject as initiated by himself and *Einstein*, see [1.5]. More recently, with the development of laser light scattering spectroscopy, Brownian motion has become very much more quantitatively measurable. The technique is to shine intense, coherent laser light into a small volume of liquid containing Brownian particles, and to study the fluctuations in the intensity of the scattered light, which are directly related to the motions of the Brownian particles. By these means it is possible to observe Brownian motion of much smaller particles than the traditional pollen, and to derive useful data about the sizes of viruses and macromolecules. With the preparation of more concentrated suspensions, interactions between the particles appear, generating interesting and quite complex problems related to macromolecular suspensions and colloids [1.6]

The general concept of fluctuations describable by such equations has developed very extensively in a very wide range of situations. The advantages of a continuous description turn out to be very significant, since only a very few parameters are required, i.e., essentially the coefficients of the derivatives in (1.2.7):

$$\int_{-\infty}^{\infty} \Delta\phi(\Delta)d\Delta, \quad \text{and} \quad \int_{\infty}^{\infty} \Delta^2\phi(\Delta)d\Delta\,. \tag{1.2.20}$$

It is rare to find a problem which cannot be specified, in at least some degree of approximation, by such a system, and for qualitative simple analysis of problems it is normally quite sufficient to consider an appropriate Fokker-Planck equation, of a form obtained by allowing both coefficients (1.2.20) to depend on x, and in a space of an appropriate number of dimensions.

1.3 Birth-Death Processes

A wide variety of phenomena can be modelled by a particular class of process called a birth-death process. The name obviously stems from the modelling of human or animal populations in which individuals are born, or die. One of the most entertaining models is that of the prey-predator system consisting of two kinds of animal, one of which preys on the other, which is itself supplied with an inexhaustible food supply. Thus letting X symbolise the prey, Y the predator, and A the food of the prey, the process under consideration might be

$$X + A \to 2X \tag{1.3.1a}$$

$$X + Y \to 2Y \tag{1.3.1b}$$

$$Y \to B \tag{1.3.1c}$$

which have the following naive, but charming interpretation. The first equation symbolises the prey eating one unit of food, and reproducing immediately. The second equation symbolises a predator consuming a prey (who thereby dies—this is the only death mechanism considered for the prey) and immediately reproducing. The final equation symbolises the death of the predator by natural causes. It is easy to guess model differential equations for x and y, the numbers of X and Y. One might assume that the first reaction symbolises a rate of production of X proportional to the product of x and the amount of food; the second equation a production of Y (and an equal rate of consumption of X) proportional to xy, and the last equation a death rate of Y, in which the rate of death of Y is simply proportional to y; thus we might write

$$\frac{dx}{dt} = k_1 ax - k_2 xy \tag{1.3.2a}$$

$$\frac{dy}{dt} = k_2 xy - k_3 y\,. \tag{1.3.2b}$$

The solutions of these equations, which were independently developed by *Lotka* [1.7] and *Volterra* [1.8] have very interesting oscillating solutions, as presented in Fig. 1.3a. These oscillations are qualitatively easily explicable. In the absence of significant numbers of predators, the prey population grows rapidly until the presence of so much prey for the predators to eat stimulates their rapid reproduction, at the same time reducing the number of prey which get eaten. Because a large number of prey have been eaten, there are no longer enough to maintain the population of predators, which then die out, returning us to our initial situation. The cycles repeat indefinitely and are indeed, at least qualitatively, a feature of many real prey-predator systems. An example is given in Fig. 1.3b.

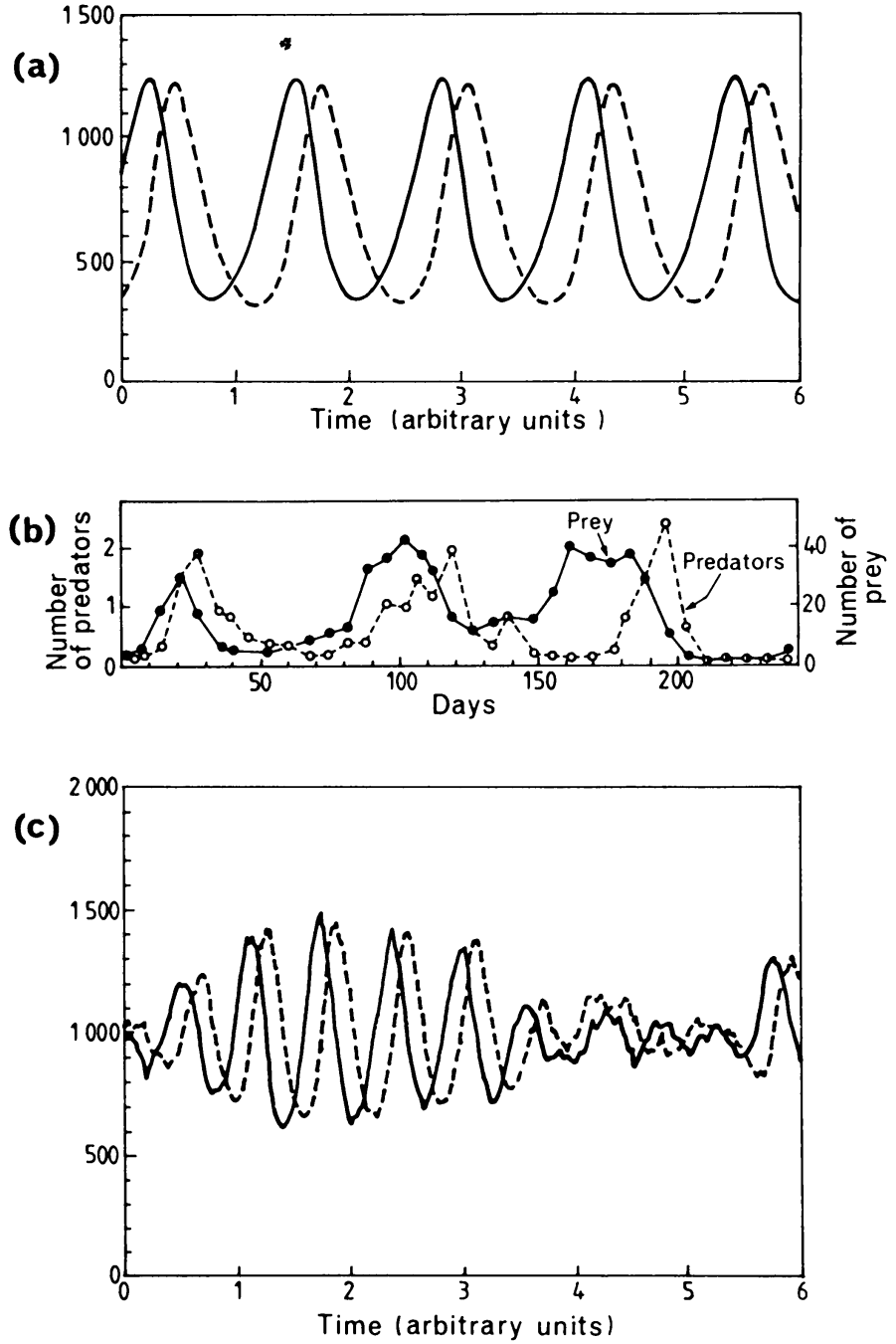


Fig. 1.3a-c. Time development in prey-predator systems. (**a**) Plot of solutions of the deterministic equations (1.3.2) (x = solid line, y = dashed line). (**b**) Data for a real prey-predator system. Here the predator is a mite (Eotetranychus sexmaculatus—dashed line) which feeds on oranges, and the prey is another mite (Typhlodromus occidentalis). Data from [1.16, 17]. (**c**) Simulation of stochastic equations (1.3.3)

Of course, the realistic systems do not follow the solutions of differential equations exactly—they fluctuate about such curves. One must include these fluctuations and the simplest way to do this is by means of a *birth-death master equation.* We assume a probability distribution, $P(x, y, t)$, for the number of individuals at a given time and ask for a probabilistic law corresponding to (1.3.2). This is done by assuming that in an infinitesimal time Δt, the following *transition probability* laws holds.

$$\text{Prob}\,(x \to x+1;\ y \to y) = k_1 a x \Delta t, \tag{1.3.3a}$$

$$\text{Prob}\,(x \to x-1;\ y \to y+1) = k_2 x y \Delta t, \tag{1.3.3b}$$

$$\text{Prob}\,(x \to x;\ y \to y-1) = k_3 y \Delta t, \tag{1.3.3c}$$

$$\text{Prob}\,(x \to x;\ y \to y) = 1-(k_1 a x + k_2 x y + k_3 y)\Delta t. \tag{1.3.3d}$$

Thus, we simply, for example, replace the simple rate laws by probability laws. We then employ what amounts to the same equation as Einstein and others used, i.e., the Chapman-Kolmogorov equation, namely, we write the probability at $t + \Delta t$ as a sum of terms, each of which represents the probability of a previous state multiplied by the probability of a transition to the state (x, y). Thus, we find

$$\frac{P(x, y, t + \Delta t) - P(x, y, t)}{\Delta t} = k_1 a(x - 1)P(x - 1, y, t) + k_2(x + 1)(y - 1) \times P(x + 1, y - 1, t) + k_3(y + 1)P(x, y + 1, t) - (k_1 a x + k_2 x y + k_3 y) \times P(x, y, t) \tag{1.3.4}$$

and letting $\Delta t \to 0$, $= \partial P(x, y, t)/\partial t$. In writing the assumed probability laws (1.3.3), we are assuming that the probability of each of the events occurring can be determined simply from the knowledge of x and y. This is again the Markov postulate which we mentioned in Sect. 1.2.1. In the case of Brownian motion, very convincing arguments can be made in favour of this Markov assumption. Here it is by no means clear. The concept of heredity, i.e., that the behaviour of progeny is related to that of parents, clearly contradicts this assumption. How to *include* heredity is another matter; by no means does a unique prescription exist.

The assumption of the Markov postulate in this context is valid to the extent that different individuals of the same species are similar; it is invalid to the extent that, nevertheless, perceptible inheritable differences do exist.

This type of model has a wide application—in fact to any system to which a population of indivuduals may be attributed, for example systems of molecules of various chemical compounds, of electrons, of photons and similar physical particles as well as biological systems. The particular choice of transition probabilities is made on various grounds determined by the degree to which details of the births and deaths involved are known. The simple multiplicative laws, as illustrated in (1.3.3), are the most elementary choice, ignoring, as they do, almost all details of

the processes involved. In some of the physical processes we can derive the transition probabilities in much greater detail and with greater precision.

Equation (1.3.4) has no simple solution, but one major property differentiates equations like it from an equation of Langevin's type, in which the fluctuation term is simply added to the differential equation. Solutions of (1.3.4) determine both the gross deterministic motion and the fluctuations; the fluctuations are typically of the same order of magnitude as the square roots of the *numbers* of individuals involved. It is not difficult to simulate a sample time development of the process as in Fig. 1.3c. The figure does show the correct general features, but the model is so obviously simplified that exact agreement can never be expected. Thus, in contrast to the situation in Brownian motion, we are not dealing here so much with a theory of a phenomenon, as with a class of mathematical models, which are simple enough to have a very wide range of approximate validity. We will see in Chap. 7 that a theory can be developed which can deal with a wide range of models in this category, and that there is indeed a close connection between this kind of theory and that of stochastic differential equations.

1.4 Noise in Electronic Systems

The early days of radio with low transmission powers and primitive receivers, made it evident to every ear that there were a great number of highly irregular electrical signals which occurred either in the atmosphere, the receiver, or the radio transmitter, and which were given the collective name of "noise", since this is certainly what they sounded like on a radio. Two principal sources of noise are shot noise and Johnson noise.

1.4.1 Shot Noise

In a vacuum tube (and in solid-state devices) we get a nonsteady electrical current, since it is generated by individual electrons, which are accelerated across a distance and deposit their charge one at a time on the anode. The electric current arising from such a process can be written

$$I(t) = \sum_{t_k} F(t - t_k) , \tag{1.4.1}$$

where $F(t\text{-}t_k)$ represents the contribution to the current of an electron which arrives at time t_k. Each electron is therefore assumed to give rise to the same shaped pulse, but with an appropriate delay, as in Fig. 1.4.

A statistical aspect arises immediately we consider what kind of choice must be made for t_k. The simplest choice is that each electron arrives independently of the previous one—that is, the times t_k are randomly distributed with a certain average number per unit time in the range $(-\infty, \infty)$, or whatever time is under consideration.

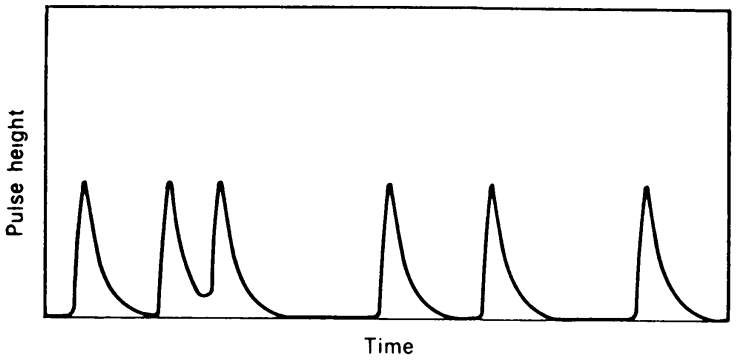


Fig. 1.4. Illustration of shot noise: identical electric pulses arrive at random times

The analysis of such noise was developed during the 1920's and 1930's and was summarised and largely completed by *Rice* [1.9]. It was first considered as early as 1918 by *Schottky* [1.10].

We shall find that there is a close connection between shot noise and processes described by birth-death master equations. For, if we consider n, the number of electrons which have arrived up to a time t, to be a statistical quantity described by a probability $P(n, t)$, then the assumption that the electrons arrive independently is clearly the Markov assumption. Then, assuming the probability that an electron will arrive in the time interval between t and $t + \Delta t$ is completely independent of t and n, its only dependence can be on Δt. By choosing an appropriate constant λ, we may write

$$\text{Prob}\,(n \to n + 1, \text{ in time } \Delta t) = \lambda \Delta t \tag{1.4.2}$$

so that

$$P(n, t + \Delta t) = P(n, t)\,(1 - \lambda \Delta t) + P(n - 1, t)\lambda \Delta t \tag{1.4.3}$$

and taking the limit $\Delta t \to 0$

$$\frac{\partial P(n, t)}{\partial t} = \lambda[P(n-1, t) - P(n, t)] \tag{1.4.4}$$

which is a pure birth process. By writing

$$G(s, t) = \sum s^n P(n, t) \tag{1.4.5}$$

[here, $G(s, t)$ is known as the generating function for $P(n, t)$, and the particular technique of solving (1.28) is very widely used], we find

$$\frac{\partial G(s, t)}{\partial t} = \lambda(s - 1)G(s, t) \tag{1.4.6}$$

so that

$$G(s, t) = \exp\,[\lambda(s-1)t]G(s, 0)\,. \tag{1.4.7}$$

By requiring at time $t = 0$ that no electrons had arrived, it is clear that $P(0, 0)$ is 1 and $P(n, 0)$ is zero for all $n \geqslant 1$, so that $G(s, 0) = 1$. Expanding the solution (1.4.7)

in powers of s, we find

$$P(n, t) = \exp(-\lambda t)(\lambda t)^n/n! \tag{1.4.8}$$

which is known as a *Poisson distribution* (Sect. 2.8.3). Let us introduce the variable $N(t)$, which is to be considered as the number of electrons which have arrived up to time t, and is a random quantity. Then,

$$P(n, t) = \text{Prob}\ \{N(t) = n\}, \tag{1.4.9}$$

and $N(t)$ can be called a *Poisson process variable*. Then clearly, the quantity $\mu(t)$, formally defined by

$$\mu(t) = dN(t)/dt, \tag{1.4.10}$$

is zero, except when $N(t)$ increases by 1; at that stage it is a Dirac delta function, i.e.,

$$\mu(t) = \sum_k \delta(t - t_k), \tag{1.4.11}$$

where the t_k are the times of arrival of the individual electrons. We may write

$$I(t) = \int_{-\infty}^{\infty} dt' F(t - t')\mu(t'). \tag{1.4.12}$$

A very reasonable restriction on $F(t - t')$ is that it vanishes if $t < t'$, and that for $t \to \infty$, it also vanishes. This simply means that no current arises from an electron before it arrives, and that the effect of its arrival eventually dies out. We assume then, for simplicity, the very commonly encountered form

$$\begin{aligned} F(t) &= q\,\mathrm{e}^{-\alpha t} && (t > 0) \\ &= 0 && (t < 0) \end{aligned} \tag{1.4.13}$$

so that (1.4.12) can be rewritten as

$$I(t) = \int_{-\infty}^{t} dt' q\,\mathrm{e}^{-\alpha(t-t')} \frac{dN(t')}{dt'}. \tag{1.4.14}$$

We can derive a simple differential equation. We differentiate $I(t)$ to obtain

$$\frac{dI(t)}{dt} = \left[q\,\mathrm{e}^{-\alpha(t-t')} \frac{dN(t')}{dt}\right]_{t'=t} + \int_{-\infty}^{t} dt'(-\alpha q)\mathrm{e}^{-\alpha(t-t')} \frac{dN(t')}{dt'} \tag{1.4.15}$$

so that

$$\frac{dI(t)}{dt} = -\alpha I(t) + q\mu(t). \tag{1.4.16}$$

This is a kind of stochastic differential equation, similar to Langevin's equation, in which, however, the fluctuating force is given by $q\mu(t)$, where $\mu(t)$ is the derivative of the Poisson process, as given by (1.4.11). However, the mean of $\mu(t)$ is nonzero, in fact, from (1.4.10)

$$\langle \mu(t)dt \rangle = \langle dN(t) \rangle = \lambda dt \tag{1.4.17}$$

$$\langle [dN(t) - \lambda dt]^2 \rangle = \lambda dt \tag{1.4.18}$$

from the properties of the Poisson distribution, for which the variance equals the mean. Defining, then, the fluctuation as the difference between the mean value and $dN(t)$, we write

$$d\eta(t) = dN(t) - \lambda dt\,, \tag{1.4.19}$$

so that the stochastic differential equation (1.4.16) takes the form

$$dI(t) = [\lambda q - \alpha I(t)]\, dt + q d\eta(t)\,. \tag{1.4.20}$$

Now how does one solve such an equation? In this case, we have an academic problem anyway since the solution is known, but one would like to have a technique. Suppose we try to follow the method used by Langevin—what will we get as an answer? The short reply to this question is: nonsense. For example, using ordinary calculus and assuming $\langle I(t)d\eta(t) \rangle = 0$, we can derive

$$\frac{d\langle I(t) \rangle}{dt} = \lambda q - \alpha \langle I(t) \rangle \qquad \text{and} \tag{1.4.21}$$

$$\frac{1}{2}\frac{d\langle I^2(t) \rangle}{dt} = \lambda q \langle I(t) \rangle - \alpha \langle I^2(t) \rangle \tag{1.4.22}$$

solving in the limit $t \to \infty$, where the mean values would reasonably be expected to be constant one finds

$$\langle I(\infty) \rangle = \lambda q/\alpha \qquad \text{and} \tag{1.4.23}$$

$$\langle I^2(\infty) \rangle = (\lambda q/\alpha)^2\,. \tag{1.4.24}$$

The first answer is reasonable—it merely gives the average current through the system in a reasonable equation, but the second implies that the mean square current is the same as the square of the mean, i.e., the current at $t \to \infty$ does not fluctuate! This is rather unreasonable, and the solution to the problem will show that stochastic differential equations are rather more subtle than we have so far presented.

Firstly, the notation in terms of differentials used in (1.4.17–20) has been chosen deliberately. In deriving (1.4.22), one uses ordinary clalculus, i.e., one writes

$$d(I^2) \equiv (I + dI)^2 - I^2 = 2IdI + (dI)^2 \tag{1.4.25}$$

and then one drops the $(dI)^2$ as being of second order in dI. But now look at (1.4.18): this is equivalent to

$$\langle d\eta(t)^2 \rangle = \lambda\, dt \tag{1.4.26}$$

so that a *quantity of second order in $d\eta$ is actually of first order in dt.* The reason is not difficult to find. Clearly,

$$d\eta(t) = dN(t) - \lambda\, dt, \tag{1.4.27}$$

but the curve of $N(t)$ is a step function, discontinuous, and certainly not differentiable, at the times of arrival of the individual electrons. In the ordinary sense, none of these calculus manipulations is permissible. But we can make sense out of them as follows. Let us simply calculate $\langle d(I^2) \rangle$ using (1.4.20, 25, 26):

$$\begin{aligned} \langle d(I)^2 \rangle &= 2\langle I\{[\lambda q - \alpha I]dt + q\, d\eta(t)\}\rangle \\ &\quad + \langle \{[\lambda q - \alpha I]dt + q\, d\eta(t)\}^2 \rangle\ . \end{aligned} \tag{1.4.28}$$

We now assume again that $\langle I(t)d\eta(t)\rangle = 0$ and expand, after taking averages using the fact that $\langle d\eta(t)^2\rangle = \lambda\, dt$, to 1*st order in dt*. We obtain

$$\frac{1}{2} d\langle I^2 \rangle = \left[\lambda q \langle I \rangle - \alpha \langle I^2 \rangle + \frac{q^2 \lambda}{2}\right] dt \tag{1.4.29}$$

and this gives

$$\langle I^2(\infty) \rangle - \langle I(\infty) \rangle^2 = \frac{q^2 \lambda}{2\alpha}\ . \tag{1.4.30}$$

Thus, there are fluctuations from this point of view, as $t \to \infty$. The extra term in (1.4.29) as compared to (1.4.22) arises directly out of the statistical considerations implicit in $N(t)$ being a discontinuous random function.

Thus we have discovered a somewhat deeper way of looking at Langevin's kind of equation—the treatment of which, from this point of view, now seems extremely naive. In Langevin's method the fluctuating force X is not specified, but it will become clear in this book that problems such as we have just considered are very widespread in this subject. The moral is that random functions cannot normally be differentiated according to the usual laws of calculus: special rules have to be developed, and a precise specification of what one means by differentiation becomes important. We will specify these problems and their solutions in Chap. 4 which will concern itself with situations in which the fluctuations are Gaussian.

1.4.2 Autocorrelation Functions and Spectra

The measurements which one can carry out on fluctuating systems such as electric circuits are, in practice, not of unlimited variety. So far, we have considered the

distribution functions, which tell us, at any time, what the probability distribution of the values of a stochastic quantity are. If we are considering a measureable quantity $x(t)$ which fluctuates with time, in practice we can sometimes determine the distribution of the values of x, though more usually, what is available at one time are the mean $\bar{x}(t)$ and the variance var $\{x(t).\}$

The mean and the variance do not tell a great deal about the underlying dynamics of what is happening. What would be of interest is some quantity which is a is a measure of the influence of a value of x at time t on the value at time $t + \tau$. Such a quantity is the *autocorrelation function,* which was apparently first introduced by *Taylor* [1.11] as

$$G(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt\, x(t)x(t+\tau)\,. \tag{1.4.31}$$

This is the time average of a two-time product over an arbitrary large time T, which is then allowed to become infinite.

Nowadays purpose built autocorrelators exist, which sample data and directly construct the autocorrelation function of a desired process, from laser light scattering signals to bacterial counts. It is also possible to construct autocorrelation programs for high speed on line experimental computers. Further, for very fast systems, there are clipped autocorrelators, which measure an approximation to the autocorrelation function given by defining a variable $c(t)$ such that

$$\begin{aligned} c(t) &= 0 \qquad x(t) < l \\ &= 1 \qquad x(t) > l \end{aligned} \tag{1.4.32}$$

and computing the autocorrelation function of that variable.

A more traditional approach is to compute the spectrum of the quantity $x(t)$. This is defined in two stages. First, define

$$y(\omega) = \int_0^T dt\, e^{-i\omega t} x(t) \tag{1.4.33}$$

then the spectrum is defined by

$$S(\omega) = \lim_{T\to\infty} \frac{1}{2\pi T} |y(\omega)|^2\,. \tag{1.4.34}$$

The autocorrelation function and the spectrum are closely connected. By a little manipulation one finds

$$S(\omega) = \lim_{T\to\infty} \left[\frac{1}{\pi} \int_0^T \cos(\omega\tau) d\tau \frac{1}{T} \int_0^{T-\tau} x(t)x(t+\tau) dt \right] \tag{1.4.35}$$

and taking the limit $T \to \infty$ (under suitable assumptions to ensure the validity of certain interchanges of order), one finds

$$S(\omega) = \frac{1}{\pi}\int_0^\infty \cos(\omega\tau)G(\tau)d\tau\,. \tag{1.4.36}$$

This is a fundamental result which relates the Fourier transform of the autocorrelation function to the spectrum. The result may be put in a slightly different form when one notices that

$$G(-\tau) = \lim_{T\to\infty}\frac{1}{T}\int_{-\tau}^{T-\tau} dt\; x(t+\tau)x(t) = G(\tau) \tag{1.4.37}$$

so we obtain

$$\boxed{S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau}G(\tau)d\tau} \tag{1.4.38}$$

with the corresponding inverse

$$\boxed{G(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau}S(\omega)d\omega\,.} \tag{1.4.39}$$

This result is known as the *Wiener-Khinchin theorem* [1.12,13] and has widespread application.

It means that one may either directly measure the autocorrelation function of a signal, or the spectrum, and convert back and forth, which by means of the fast Fourier transform and computer is relatively straightforward.

1.4.3 Fourier Analysis of Fluctuating Functions: Stationary Systems

The autocorrelation function has been defined so far as a time average of a signal, but we may also consider the *ensemble average,* in which we repeat the same measurement many times, and compute averages, denoted by $\langle\ \rangle$. It will be shown that for very many systems, the time average is equal to the ensemble average; such systems are termed *ergodic* (Sect. 3.7.1).

If we have such a fluctuating quantity $x(t)$, then we can consider the average

$$\langle x(t)x(t+\tau)\rangle = G(\tau), \tag{1.4.40}$$

this result being the consequence of our ergodic assumption.

Now it is very natural to write a Fourier transform for the stochastic quantity $x(t)$

$$x(t) = \int d\omega\; c(\omega)\, e^{i\omega t} \tag{1.4.41}$$

and consequently,

$$c(\omega) = \frac{1}{2\pi}\int dt\; x(t)\, e^{-i\omega t}\,. \tag{1.4.42}$$

Note that $x(t)$ real implies

$$c(\omega) = c^*(-\omega)\,. \tag{1.4.43}$$

If the system is ergodic, we must have a constant $\langle x(t)\rangle$, since the time average is clearly constant. The process is then *stationary* by which we mean that all time-dependent averages are functions only of time differences, i.e., averages of functions $x(t_1)$, $x(t_2)$, ... $x(t_n)$ are equal to those of $x(t_1 + \Delta)$, $x(t_2 + \Delta)$, ... $x(t_n + \Delta)$.

For convenience, in what follows we assume $\langle x\rangle = 0$. Hence,

$$\langle c(\omega)\rangle = \frac{1}{2\pi}\int dt\,\langle x\rangle e^{-i\omega t} = 0 \tag{1.4.44}$$

$$\begin{aligned}\langle c(\omega)c^*(\omega')\rangle &= \frac{1}{(2\pi)^2}\iint dt\,dt' e^{-i\omega t + i\omega' t'}\langle x(t)x(t')\rangle \\ &= \frac{1}{(2\pi)}\delta(\omega - \omega')\int d\tau\, e^{i\omega\tau}G(\tau) \\ &= \delta(\omega - \omega')S(\omega)\,. \end{aligned} \tag{1.4.45}$$

Here we find not only a relationship between the mean square $\langle |c(\omega)|^2\rangle$ and the spectrum, but also the result that stationarity alone implies that $c(\omega)$ and $c^*(\omega')$ are uncorrelated, since the term $\delta(\omega - \omega')$ arises because $\langle x(t)x(t')\rangle$ is a function only of $t - t'$.

1.4.4 Johnson Noise and Nyquist's Theorem

Two brief and elegant papers appeared in 1928 in which *Johnson* [1.14] demonstrated experimentally that an electric resistor automatically generated fluctuations of electric voltage, and *Nyquist* [1.15] demonstrated its theoretical derivation, in complete accordance with Johnson's experiment. The principle involved was already known by *Schottky* [1.10] and is the same as that used by Einstein and Langevin. This principle is that of thermal equilibrium. If a resistor R produces electric fluctuations, these will produce a current which will generate heat. The heat produced in the resistor must exactly balance the energy taken out of the fluctuations. The detailed working out of this principle is not the subject of this section, but we will find that such results are common throughout the physics and chemistry of stochastic processes, where the principles of statistical mechanics, whose basis is not essentially stochastic, are brought in to complement those of stochastic processes. The experimental result found was the following. We have an electric resistor of resistance R at absolute temperature T. Suppose by means of a suitable filter we measure $E(\omega)d\omega$, the voltage across the resistor with angular frequency in the range $(\omega, \omega + d\omega)$. Then, if k is Boltzmann's constant,

$$\langle E^2(\omega)\rangle = R\,kT/\pi\ . \tag{1.4.46}$$

This result is known nowadays as *Nyquist's theorem*. Johnson remarked. "The effect is one of the causes of what is called 'tube noise' in vacuum tube amplifiers. Indeed, it is often by far the larger part of the 'noise' of a good amplifier."

Johnson noise is easily described by the formalism of the previous subsection. The mean noise voltage is zero across a resistor, and the system is arranged so that it is in a steady state and is expected to be well represented by a stationary process. Johnson's quantity is, in practice, a limit of the kind (1.4.34) and may be summarised by saying that the voltage spectrum $S(\omega)$ is given by

$$S(\omega) = R\,kT/\pi, \tag{1.4.47}$$

that is, the spectrum is flat, i.e., a constant function of ω. In the case of light, the frequencies correspond to different colours of light. If we perceive light to be white, it is found that in practice all colours are present in equal proportions—the optical spectrum of white light is thus flat—at least within the visible range. In analogy, the term *white noise* is applied to a noise voltage (or any other fluctuating quantity) whose spectrum is flat.

White noise cannot actually exist. The simplest demonstration is to note that the mean power dissipated in the resistor in the frequency range (ω_1, ω_2) is given by

$$\int_{\omega_1}^{\omega_2} d\omega\ S(\omega)/R = kT(\omega_1 - \omega_2)/\pi \tag{1.4.48}$$

so that the total power dissipated in all frequencies is infinite! Nyquist realised this, and noted that, in practice, there would be quantum corrections which would, at room temperature, make the spectrum flat only up to 7×10^{13} Hz, which is not detectable in practice, in a radio situation. The actual power dissipated in the resistor would be somewhat less than infinite, 10^{-10} W in fact! And in practice there are other limiting factors such as the inductance of the system, which would limit the spectrum to even lower frequencies.

From the definition of the spectrum in terms of the autocorrelation function given in Sect. 1.4, we have

$$\langle E(t+\tau)E(t)\rangle = G(\tau) \tag{1.4.49}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, e^{-i\omega\tau} 2R\,kT \tag{1.4.50}$$

$$= 2R\,kT\delta(\tau)\,, \tag{1.4.51}$$

which implies that no matter how small the time difference τ, $E(t+\tau)$ and $E(\tau)$ are not correlated. This is, of course, a direct result of the flatness of the spectrum. A typical model of $S(\omega)$ that is almost flat is

$$S(\omega) = R\,kT/[\pi(\omega^2\tau_C^2 + 1)] \tag{1.4.52}$$

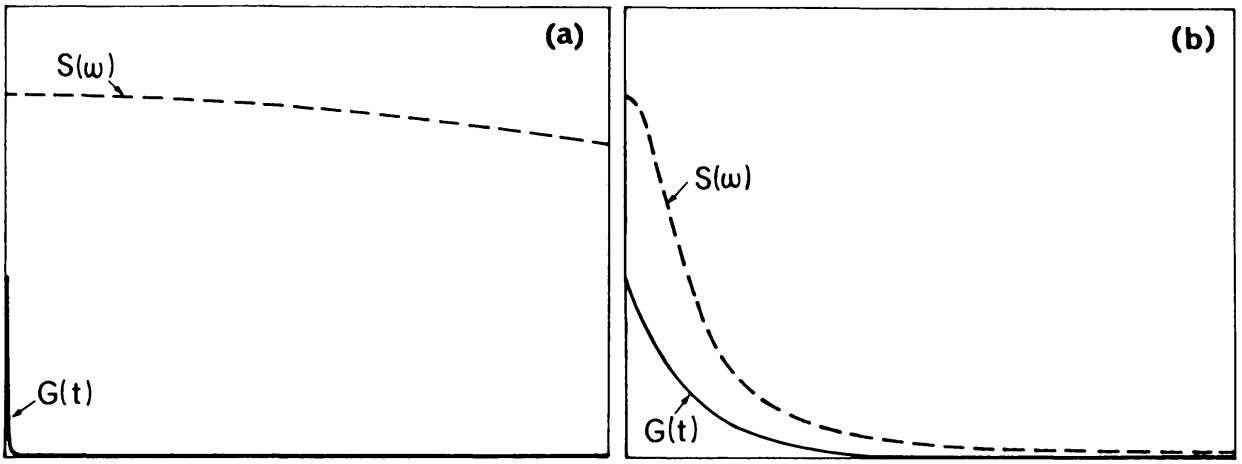


Fig. 1.5.a,b. Correlation Functions (———) and corresponding spectra (- - - - -) for (**a**) short correlation time corresponding to an almost flat spectrum; (**b**) long correlation time, giving a quite rapidly decreasing spectrum

This is flat provided $\omega \ll \tau_C^{-1}$. The Fourier transform can be explicitly evaluated in this case to give

$$\langle E(t+\tau)E(t)\rangle = (R\,kT/\tau_C)\exp(-\tau/\tau_C) \tag{1.4.53}$$

so that the autocorrelation function vanishes only for $\tau \gg \tau_C$, which is called the *correlation time* of the fluctuating voltage. Thus, the delta function correlation function appears as an idealisation, only valid on a sufficiently long time scale.

This is very reminiscent of Einstein's assumption regarding Brownian motion and of the behaviour of Langevin's fluctuating force. The idealised *white noise* will play a highly important role in this book but, in just the same way as the fluctuation term that arises in a stochastic differential equation is not the same as an ordinary differential, we will find that differential equations which include white noise as a driving term have to be handled with great care. Such equations arise very naturally in any fluctuating system and it is possible to arrange by means of *Stratonovich's rules* for ordinary calculus rules to apply, but at the cost of imprecise mathematical definition and some difficulties in stochastic manipulation. It turns out to be far better to abandon ordinary calculus and use the *Ito calculus*, which is not very different (it is, in fact, very similar to the calculus presented for shot noise) and to preserve tractable statistical properties. All these matters will be discussed thoroughly in Chap. 4.

White noise, as we have noted above, does not exist as a physically realisable process and the rather singular behaviour it exhibits does not arise in any realisable context. It is, however, fundamental in a mathematical, and indeed in a physical sense, in that it is an idealisation of very many processes that do occur. The slightly strange rules which we will develop for the calculus of white noise are not really very difficult and are very much easier to handle than any method which always deals with a real noise. Furthermore, situations in which white noise is not a good approximation can very often be indirectly expressed quite simply in terms of white noise. In this sense, white noise is the starting point from which a wide range of stochastic descriptions can be derived, and is therefore fundamental to the subject of this book.

2. Probability Concepts

In the preceding chapter, we introduced probability notions without any definitions. In order to formulate essential concepts more precisely, it is necessary to have some more precise expression of these concepts. The intention of this chapter is to provide some background, and to present a number of essential results. It is not a thorough outline of mathematical probability, for which the reader is referred to standard mathematical texts such as those by *Feller* [2.1] and *Papoulis* [2.2].

2.1 Events, and Sets of Events

It is convenient to use a notation which is as general as possible in order to describe those occurrences to which we might wish to assign probabilities. For example, we may wish to talk about a situation in which there are 6.4×10^{14} molecules in a certain region of space; or a situation in which a Brownian particle is at a certain point $\boldsymbol{x}$ in space; or possibly there are 10 mice and 3 owls in a certain region of a forest.

These occurrences are all examples of practical realisations of *events*. More abstractly, an event is simply a member of a certain space, which in the cases most practically occuring can be characterised by a vector of integers

$$\boldsymbol{n} = (n_1, n_2, n_3 \ldots) \tag{2.1.1}$$

or a vector of real numbers

$$\boldsymbol{x} = (x_1, x_2, x_3 \ldots). \tag{2.1.2}$$

The dimension of the vector is arbitary.

It is convenient to use the language of set theory, introduce the concept of a *set of events*, and use the notation

$$\omega \in A \tag{2.1.3}$$

to indicate that the event ω is one of events contained in A. For example, one may consider the set $A(25)$ of events in the ecological population in which there are no more than 25 animals present; clearly the event $\bar{\omega}$ that there are 3 mice, a tiger, and no other animals present satisfies

$$\bar{\omega} \in A(25). \tag{2.1.4}$$

More significantly, suppose we define the set of events $A(\boldsymbol{r}, \Delta V)$ that a molecule is within a volume element ΔV centred on a point $\boldsymbol{r}$. In this case, the practical significance of working in terms of sets of events becomes clear, because we should normally be able to determine whether or not a molecule is within a neighbourhood ΔV of $\boldsymbol{r}$, but to determine whether the particle is exactly at $\boldsymbol{r}$ is impossible. Thus, if we define the event $\omega(\boldsymbol{y})$ that the molecule is at point $\boldsymbol{y}$, it makes sense to ask whether

$$\omega(\boldsymbol{y}) \in A(\boldsymbol{r}, \Delta V) \tag{2.1.5}$$

and to assign a certain probability to the *set* $A(\boldsymbol{r}, \Delta V)$, which is to be interpreted as the probability of the occurrence of (2.1.5)

2.2 Probabilities

Most people have an intuitive conception of a probability, based on their own experience. However, a precise formulation of intuitive concepts is fraught with difficulties, and it has been found most convenient to axiomatise probability theory as an essentially abstract science, in which a probability measure $P(A)$ is *assigned* to every set A, in the space of events, including

the set of all events: Ω (2.2.1)

the set of no events: $\varnothing$; (2.2.2)

in order to define probability, we need our sets of events to form a closed system (known by mathematicians as a *σ-algebra*) under the set theoretic operations of union and intersection.

2.2.1 Probability Axioms

We introduce the probability of A, $P(A)$, as a function of A satisfying the following *probability axioms*:

(i) $P(A) \geqslant 0$ for all A, (2.2.3)

(ii) $P(\Omega) = 1$, (2.2.4)

(iii) if A_i $(i=1, 2, 3, \ldots)$ is a countable (but possibly infinite) collection of *nonoverlapping sets*, i.e., such that

$$A_i \cap A_j = \varnothing \quad \text{for all} \quad i \neq j, \tag{2.2.5}$$

then

$$P\left(\bigcup_i A\right) = \sum_i P(A_i)\,. \tag{2.2.6}$$

These are all the axioms needed. Consequentially, however, we have:

(iv) if $\bar{A}$ is the complement of A, i.e., the set of all events not contained in A, then

$$P(\bar{A}) = 1 - P(A), \tag{2.2.7}$$

(v) $$P(\varnothing) = 0. \tag{2.2.8}$$

2.2.2 The Meaning of $P(A)$

There is no way of making probability theory correspond to reality without requiring a certain degree of intuition. The probability $P(A)$, as axiomatised above, is the intuitive probability that an "*arbitrary*" event ω, i.e., an event ω "*chosen at random*", will satisfy $\omega \in A$. Or more explicitly, if we choose an event "*at random*" from Ω N times, the relative frequency that the particular event chosen will satisfy $\omega \in A$ approaches $P(A)$ as the number of times, N, we choose the event, approaches infinity. The number of choices N can be visualised as being done one after the other ("*independent*" tosses of one die) or at the same time (N dice are thrown at the same time "*independently*"). All definitions of this kind must be intuitive, as we can see by the way undefined terms ("*arbitrary*", "*at random*", "*independent*") keep turning up. By eliminating what we now think of as intuitive ideas and axiomatising probability, *Kolomogorov* [2.3] cleared the road for a rigorous development of mathematical probability. But the circular definition problems posed by wanting an intuitive understanding remain. The simplest way of looking at axiomatic probability is as a formal method of manipulating probabilities using the axioms. In order to apply the theory, the probability space must be defined *and* the probability measure P assigned. These are *a priori probabilities*, which are simply assumed. Examples of such a priori probabilities abound in applied disciplines. For example, in equilibrium statistical mechanics one assigns equal probabilities to equal volumes of phase space. Einstein's reasoning in Brownian motion assigned a probability $\phi(\Delta)$ to the probability of a "push" Δ from a position x at time t.

The task of applying probability is (i) to assume some set of *a priori* probabilities which seem reasonable and to deduce results from this and from the structure of the probability space, (ii) to measure experimental results with some apparatus which is constructed to measure quantities in accordance with these a priori probabilities.

The structure of the probability space is very important, especially when the space of events is compounded by the additional concept of time. This extension makes the effective probability space infinite-dimensional, since we can construct events such as "the particle was at points $\boldsymbol{x}_n$ at times t_n, $n = 0, 1, 2, \ldots, \infty$".

2.2.3 The Meaning of the Axioms

Any intuitive concept of probability gives rise to nonnegative probabilities, and the probability that an arbitrary event is contained in the set of all events must be 1 no matter what our definition of the word arbitrary. Hence, axioms (i) and (ii) are understandable. The heart of the matter lies in axiom (iii). Suppose we are dealing with only 2 sets A and B, and $A \cap B = \varnothing$. This means there are *no* events con-

tained in both A and B. Therefore, the probability that $\omega \in A \cup B$ is the probability that *either* $\omega \in A$ or $\omega \in B$. Intuitive considerations tell us this probability is the sum of the individual probabilities, i.e.,

$$P(A \cup B) \equiv P\{(\omega \in A) \text{ or } (\omega \in B)\} = P(A) + P(B) \tag{2.2.9}$$

(notice this is not a proof—merely an explanation).

The extension now to any finite number of nonoverlapping sets is obvious, but the extension only to any *countable* number of nonoverlapping sets requires some comment.

This extension must be made restrictive because of the existence of sets labelled by a continuous index, for example, $\boldsymbol{x}$, the position in space. The probability of a molecule being in the set whose only element in $\boldsymbol{x}$ is zero; but the probability of being in a region R of finite volume is nonzero. The region R is a union of sets of the form $\{\boldsymbol{x}\}$—but not a *countable* union. Thus axiom (iii) is not applicable and the probability of being in R is *not* equal to the sum of the probabilities of being in $\{\boldsymbol{x}\}$.

2.2.4 Random Variables

The concept of a random variable is a notational convenience which is central to this book. Suppose we have an abstract probability space whose events can be written $\boldsymbol{x}$. Then we can introduce the random variable $F(\boldsymbol{x})$ which is a function of $\boldsymbol{x}$, which takes on certain values for each $\boldsymbol{x}$. In particular, the identity function of $\boldsymbol{x}$, written $\boldsymbol{X}(\boldsymbol{x})$ is of interest; it is given by

$$\boldsymbol{X}(\boldsymbol{x}) = \boldsymbol{x}. \tag{2.2.10}$$

We shall normally use capitals in this book to denote random variables and small letters $\boldsymbol{x}$ to denote their values whenever it is necessary to make a distinction.

Very often, we have some quite different underlying probability space Ω with values ω, and talk about $\boldsymbol{X}(\omega)$ which is some function of ω, and then omit explicit mention of ω. This can be for either of two reasons:

i) we specify the events by the values of $\boldsymbol{x}$ anyway, i.e., we identify $\boldsymbol{x}$ and ω;
ii) the underlying events ω are too complicated to describe, or sometimes, even to know.

For example, in the case of the position of a molecule in a liquid, we really should interpret each ω as being capable of specifying all the positions, momenta, and orientations of each molecule in that volume of liquid; but this is simply too difficult to write down, and often unnecessary.

One great advantage of introducing the concept of a random variable is the simplicity with which one may handle functions of random variables, e.g., $\boldsymbol{X}^2$, $\sin(\boldsymbol{a} \cdot \boldsymbol{X})$, etc, and compute means and distributions of these. Further, by defining stochastic differential equations, one can also quite simply talk about time development of random variables in a way which is quite analogous to the classical description by means of differential equations of nonprobabilistic systems.

2.3 Joint and Conditional Probabilities: Independence

2.3.1 Joint Probabilities

We explained in Sect. 2.2.3 how the occurrence of mutually exclusive events is related to the concept of nonintersecting sets. We now consider the concept $P(A \cap B)$, where $A \cap B$ is nonempty. An event ω which satisfies $\omega \in A$ will only satisfy $\omega \in A \cap B$ if $\omega \in B$ as well.

$$\text{Thus, } P(A \cap B) = P\{(\omega \in A) \text{ and } (\omega \in B)\} \tag{2.3.1}$$

and $P(A \cap B)$ is called the *joint probability* that the event ω is contained in both classes, or, alternatively, that both the events $\omega \in A$ and $\omega \in B$ occur. Joint probabilities occur naturally in the context of this book in two ways:

i) When the event is specified by a vector, e.g., m mice and n tigers. The probability of this event is the joint probability of [m mice (and any number of tigers)] and [n tigers (and any number of mice)]. All vector specifications are implicitly joint probabilities in this sense.

ii) *When more than one time is considered:* what is the probability that (at time t_1 there are m_1 tigers and n_1 mice) and (at time t_2 there are m_2 tigers and n_2 mice). To consider such a probability, we have effectively created out of the events at time t_1 and events at time t_2, *joint events* involving one event at each time. In essence, there is no difference between these two cases except for the fundamental dynamical role of time.

2.3.2 Conditional Probabilities

We may specify conditions on the events we are interested in and consider only these, e.g., the probability of 21 buffaloes given that we know there are 100 lions. What does this mean? Clearly, we will be interested only in those events contained in the set $B =$ {all events where exactly 100 lions occur}. This means that we to define conditional probabilities, which are defined only on the collection of all sets contained in B. we define the conditional probability as

$$P(A|B) = P(A \cap B)/P(B) \tag{2.3.2}$$

and this satisfies our intuitive conception that the conditional probability that $\omega \in A$ (given that we know $\omega \in B$), is given by dividing the probability of joint occurrence by the probability ($\omega \in B$).

We can define in both directions, i.e., we have

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A). \tag{2.3.3}$$

There is no particular conceptual difference between, say, the probability of {(21 buffaloes) given (100 lions)} and the reversed concept. However, when two times

are involved, we do see a difference. For example, the probability that a particle is at position x_1 at time t_1, given that it was at x_2 at the *previous* time t_2, is a very natural thing to consider; indeed, it will turn out to be a central concept in this book. The converse sounds strange, i.e., the probability that a particle is at position x_1 at time t_1, given that it will be at position x_2 at a later time t_2. It smacks of clairvoyance—we cannot conceive of any natural way in which we would wish to consider it, although it is, in principle, a quantity very similar to the "natural" conditional probability, in which the condition precedes the events under consideration.

The natural definition has already occurred in this book, for example, the $\phi(\Delta)d\Delta$ of Einstein (Sect. 1.2.1.) is the probability that a particle at x at time t will be in the range $[x + \Delta, x + \Delta + d\Delta]$ at time $t + \tau$, and similarly in the other examples. Our intuition tells us as it told Einstein (as can be seen by reading the extract from his paper) that this kind of conditional probability is directly related to the time development of a probabilistic system.

2.3.3 Relationship Between Joint Probabilities of Different Orders

Suppose we have a collection of sets B_t such that

$$B_t \cap B_j = \varnothing \tag{2.3.4}$$

$$\bigcup_t B_t = \Omega \tag{2.3.5}$$

so that the sets divide up the space Ω into nonoverlapping subsets.
Then

$$\bigcup_t (A \cap B_t) = A \cap (\bigcup_t B_t) = A \cap \Omega = A \tag{2.3.6}$$

Using now the probability axiom (iii), we see that $A \cap B_t$ satisfy the conditions on the A_t used there, so that

$$\sum_t P(A \cap B_t) = P[\bigcup_t (A \cup B_t)] \tag{2.3.7}$$

$$= P(A) \tag{2.3.8}$$

and thus

$$\sum_t P(A \mid B_t)P(B_t) = P(A) \tag{2.3.9}$$

Thus, summing over all mutually exclusive possibilities of B in the joint probability eliminates that variable.

Hence, in general,

$$\sum_t P(A_t \cap B_j \cap C_k \ldots) = P(B_j \cap C_k \cap \ldots). \tag{2.3.10}$$

The result (2.3.9) has very significant consequences in the development of the theory of stochastic processes, which depends heavily on joint probabilities.

2.3.4 Independence

We need a probabilistic way of specifying what we mean by independent events. Two sets of events A and B should represent independent sets of events if the specification that a particular event is contained in B has no influence on the probability of that event belonging to A. Thus, the conditional probability $P(A \mid B)$ should be independent of B, and hence

$$P(A \cap B) = P(A)P(B) \tag{2.3.11}$$

In the case of several events, we need a somewhat stronger specification. The events $(\omega \in A_i)$ $(i = 1,2,\ldots, n)$ will be considered to be independent if for any subset $(i_1, i_2, \ldots, i_k)$ of the set $(1,2, \ldots, n)$,

$$P(A_{i_1} \cap A_{i_2} \ldots A_{i_k}) = P(A_{i_1})P(A_{i_2}) \ldots P(A_{i_k}) . \tag{2.3.12}$$

It is important to require factorisation for all possible combinations, as in (2.3.12). For example, for three sets A_i, it is quite conceivable that

$$P(A_i \cap A_j) = P(A_i)P(A_j) \tag{2.3.13}$$

for all different i and j, but also that

$$A_1 \cap A_2 = A_2 \cap A_3 = A_3 \cap A_1 . \qquad \text{(see Fig 2.1)}$$

This requires

$$P(A_1 \cap A_2 \cap A_3) = P(A_2 \cap A_3 \cap A_3) = P(A_2 \cap A_3) = P(A_2)P(A_3) \neq P(A_1)P(A_2)P(A_3). \tag{2.3.14}$$

We can see that the occurrence of $\omega \in A_2$ and $\omega \in A_3$ necessarily implies the occurrence of $\omega \in A_1$. In this sense the events are obviously not independent.

Random variables $X_1, X_2, X_3, \ldots$, will be said to be independent random variables, if for all sets of the form $A_i = (x$ such that $a_i \leqslant x \leqslant b_i)$ the events $X_1 \in A_1$,

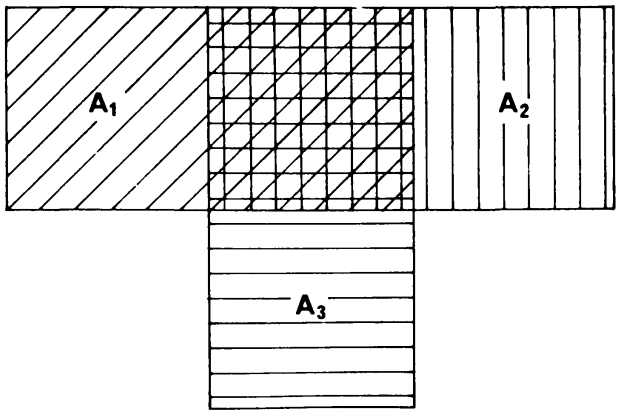


Fig. 2.1. Illustration of statistical independence in pairs, but not in threes. In the three sets $A_j \cap A_i$ is, in all cases, the central region. By appropriate choice of probabilities, we can arrange

$$P(A_i \cap A_j) = P(A_i)P(A_j)$$

$X_2 \in A_2$, $X_3 \in A_3$, ... are independent events. This will mean that all values of the X_t are assumed independently of those of the remaining X_t.

2.4 Mean Values and Probability Density

The mean value of a random variable $R(\omega)$ in which the basic events ω are countably specifiable is given by

$$\langle R \rangle = \sum_{\omega} P(\omega) R(\omega) , \tag{2.4.1}$$

where $P(\omega)$ means the probability of the set containing only the single event ω. In the case of a continuous variable, the probability axioms above enable us to define a probability density function $p(\omega)$ such that if $A(\omega_0, d\omega_0)$ is the set

$$(\omega_0 \leqslant \omega < \omega_0 + d\omega_0) , \tag{2.4.2}$$

then

$$p(\omega_0) d\omega_0 = P[A(\omega_0, d\omega_0)] \tag{2.4.3}$$

$$\equiv p(\omega_0, d\omega_0) . \tag{2.4.4}$$

The last is a notation often used by mathematicians. Details of how this is done have been nicely explained by *Feller* [2.1]. In this case,

$$\langle R \rangle = \int_{\omega \in \Omega} d\omega \, R(\omega) p(\omega) . \tag{2.4.5}$$

One can often (as mentioned in Sect. 2.2.4) use R itself to specify the event, so we will often write

$$\langle R \rangle = \int dR \, R p(R) . \tag{2.4.6}$$

Obviously, $p(R)$ is not the same function of R as $p(\omega)$ is of ω–more precisely

$$p(R_0) dR_0 = P[R_0 < R < R_0 + dR_0)] . \tag{2.4.7}$$

2.4.1 Determination of Probability Density by Means of Arbitrary Functions

Suppose for every function $f(R)$ we know

$$\langle f(R) \rangle = \int dR \, f(R) \, p(R) , \tag{2.4.8}$$

then we know $p(R)$. The proof follows by choosing

$$f(R) = 1 \qquad R_0 \leqslant R < R_0 + dR_0$$

$$= 0 \qquad \text{otherwise.}$$

Because the expectation of an arbitrary function is sometimes a little easier to work with than a density, this relation will be used occasionally in this book.

2.4.2 Sets of Probability Zero

If a density $p(R)$ exists, the probability that R is in the interval $(R_0, R_0 + dR)$ goes to zero with dR. Hence, the probability that R has *exactly* the value R_0 is zero; and similarly for any other value.

Thus, in such a case, there are sets $S(R_t)$, each containing only one point R_t, which have zero probability. From probability axiom (iii), any countable union of such sets, i.e., any set containing only a countable number of points (e.g., all rational numbers) has probability zero. In general, all equalities in probability theory are at best only "*almost certainly true*", i.e., they may be untrue on sets of probability zero. Alternatively, one says, for example,

$$X = Y \text{ (with probability 1)} \tag{2.4.9}$$

which is by no means the same as saying that

$$X(R) = Y(R) \text{ for all } R. \tag{2.4.10}$$

Of course, (if the theory is to have any connection with reality) events with probability zero do not occur.

In particular, notice that our previous result if inspected carefully, only implies that we know $p(R)$ only with probability 1, given that we know $\langle f(R) \rangle$ for all $f(R)$.

2.5 Mean Values

The question of what to measure in a probabilistic system is nontrivial. In practice, one measures either a set of individual values of a random variable (the number of animals of a certain kind in a certain region at certain points in time; the electric current passing through a given circuit element in each of a large number of replicas of that circuit, etc.) or alternatively, the measuring procedure may implicitly construct an average of some kind. For example, to measure an electric current, we may measure the electric charge transferred and divide by the time taken—this gives a measure of the average number of electrons transferred per unit time. It is important to note the essential difference in this case, that it will not normally be possible to measure anything other than a few selected averages and thus, higher moments (for example) will be unavailable.

In contrast, when we measure individual events (as in counting animals), we can then construct averages of the observables by the obvious method

$$\bar{X}_N = \frac{1}{N} \sum_{n=1}^{N} X(n) . \tag{2.5.1}$$

The quantities $X(n)$ are the individual observed values of the quantity X. We expect

that as the number of samples N becomes very large, the quantity $\bar{X}_N$ approaches the mean $\langle X \rangle$ and that, in fact,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f[X(n)] = \lim_{N\to\infty} \overline{f(X)}_N = \langle f(X) \rangle \tag{2.5.2}$$

and such a procedure will determine the probability density function $p(x)$ of X if we carry out this procedure for all functions f. The validity of this procedure depends on the degree of independence of the successive measurements and is dealt with in Sect. 2.5.2.

In the case where only averages themselves are directly determined by the measuring method, it will not normally be possible to measure $X(n)$ and therefore, it will not, in general, be possible to determine $\overline{f(X)}_N$. All that will be available will be $f(\bar{X}_N)$—quite a different thing unless f is linear. We can often find situations in which measurable quantities are related (by means of some theory) to mean values of certain functions, but to hope to measure, for example, the mean value of an arbitrary function of the number of electrons in a conductor is quite hopeless. The mean number—yes, and indeed even the mean square number, but the measuring methods available are not direct. We do *not* enumerate the individual numbers of electrons at different times and hence arbitrary functions are not attainable.

2.5.1 Moments, Correlations, and Covariances

Quantities of interest are given by the *moments* $\langle X^n \rangle$ since these are often easily calculated. However, probability densities must always vanish as $x \to \pm\infty$, so we see that higher moments tell us only about the properties of unlikely large values of X. In practice we find that the most important quantities are related to the first and second moments. In particular, for a single variable X, *the variance* defined by

$$\text{var}\{X\} \equiv \{\sigma[X]\}^2 \equiv \langle [X - \langle X \rangle]^2 \rangle \,, \tag{2.5.3}$$

and as is well known, the *variance* var $\{X\}$ or its square root the *standard deviation* $\sigma[X]$, is a measure of the degree to which the values of X deviate from the mean value $\langle X \rangle$.

In the case of several variables, we define the *covariance matrix* as

$$\langle X_i, X_j \rangle \equiv \langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle \equiv \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle \,. \tag{2.5.4}$$

Obviously,

$$\langle X_i, X_i \rangle = \text{var}\ \{X_i\} \,. \tag{2.5.5}$$

If the variables are independent *in pairs*, the covariance matrix is diagonal.

2.5.2 The Law of Large Numbers

As an application of the previous concepts, let us investigate the following model of measurement. We assume that we measure the same quantity N times, obtaining

sample values of the random variable $X(n)$ ($n = 1, 2, \ldots, N$). Since these are all measurements of the same quantity at successive times, we assume that for every n, $X(n)$ has the same probability distribution but we do not assume the $X(n)$ to be independent. However, provided the covariance matrix $\langle X(n), X(m)\rangle$ vanishes sufficiently rapidly as $|n - m| \to \infty$, then defining

$$\bar{X}_N = \frac{1}{N}\sum_{n=1}^{N} X(n)\,, \tag{2.5.6}$$

we shall show

$$\lim_{N\to\infty} \bar{X}_N = \langle X\rangle\,. \tag{2.5.7}$$

It is clear that

$$\langle \bar{X}_N\rangle = \langle X\rangle\,. \tag{2.5.8}$$

We now calculate the variance of $\bar{X}_N$ and show that as $N \to \infty$ it vanishes under certain conditions:

$$\langle \bar{X}_N\bar{X}_N\rangle - \langle \bar{X}_N\rangle^2 = \frac{1}{N^2}\sum_{n,m=1}^{N} \langle X_n, X_m\rangle\,. \tag{2.5.9}$$

Provided $\langle X_n, X_m\rangle$ falls off sufficiently rapidly as $|n - m| \to \infty$, we find

$$\lim_{N\to\infty} (\mathrm{var}\{\bar{X}_N\}) = 0 \tag{2.5.10}$$

so that $\lim_{N\to\infty} \bar{X}_N$ is a deterministic variable equal to $\langle X\rangle$.

Two models of $\langle X_n, X_m\rangle$ can be chosen.

a) $$\langle X_n, X_m\rangle \sim K\lambda^{|m-n|} \qquad (\lambda < 1) \tag{2.5.11}$$

for which one finds

$$\mathrm{var}\{\bar{X}_N\} = \frac{2K}{N^2}\left(\frac{\lambda^{N+2} - N(\lambda - 1) - \lambda}{(\lambda - 1)^2}\right) - \frac{K}{N} \to 0\,. \tag{2.5.12}$$

b) $$\langle X_n, X_m\rangle \sim |n - m|^{-1} \qquad (n \neq m) \tag{2.5.13}$$

and one finds approximately

$$\mathrm{var}\{\bar{X}_N\} \sim \frac{2}{N}\log N - \frac{1}{N} \to 0\,. \tag{2.5.14}$$

In both these cases, $\mathrm{var}\{X_N\} \to 0$. The rate of convergence is very different. In-

terpreting n, m as the times at which the measurement is carried out, one sees than even very slowly decaying correlations are permissible. The law of large numbers comes in many forms, which are nicely summarised by *Papoulis* [2.2]. The central limit theorem is an even more precise result in which the limiting distribution function of $\bar{X}_N - \langle X \rangle$ is determined (see Sect. 2.8.2).

2.6 Characteristic Function

One would like a condition where the variables are independent, not just in pairs. To this end (and others) we define the characteristic function.

If $\boldsymbol{s}$ is the vector $(s_1, s_2, \ldots, s_n)$, and $\boldsymbol{X}$ the vector of random variables $(X_1, X_2, \ldots, X_n)$, then the characteristic function (or moment generating function) is defined by

$$\phi(\boldsymbol{s}) = \langle \exp(\mathrm{i}\boldsymbol{s} \cdot \boldsymbol{X}) \rangle = \int d\boldsymbol{x}\, p(\boldsymbol{x}) \exp(\mathrm{i}\boldsymbol{s} \cdot \boldsymbol{x}) . \tag{2.6.1}$$

The characteristic function has the following properties [Ref. 2.1, Chap. XV]

i) $\phi(\boldsymbol{0}) = 1$

ii) $\phi|(\boldsymbol{s})| \leqslant 1$

iii) $\phi(\boldsymbol{s})$ is a uniformly continuous function of its arguments for all finite real $\boldsymbol{s}$ [2.5].

iv) If the *moments* $\langle \prod_i X_i^{m_i} \rangle$ exist, then

$$\langle \prod_i X_i^{m_i} \rangle = \left[\prod_i \left(-\mathrm{i} \frac{\partial}{\partial s_i} \right)^{m_i} \phi(\boldsymbol{s}) \right]_{\boldsymbol{s}=\boldsymbol{0}} . \tag{2.6.2}$$

v) A sequence of probability densities converges to limiting probability density if and only if the corresponding characteristic functions converge to the corresponding characteristic function of the limiting probability density.

vi) Fourier inversion formula

$$p(\boldsymbol{x}) = (2\pi)^{-n} \int d\boldsymbol{s}\, \phi(\boldsymbol{s}) \exp(-\mathrm{i}\boldsymbol{x} \cdot \boldsymbol{s}) \tag{2.6.3}$$

Because of this inversion formula, $\phi(\boldsymbol{s})$ determines $p(\boldsymbol{x})$ with probability 1. Hence, the characteristic function does truly *characterise* the probability density.

vii) Independent random variables: from the definition of independent random variables in Sect. 2.3.4, it follows that the variables $X_1, X_2 \ldots$ are independent if and only if

$$p(x_1, x_2, \ldots, x_n) = p_1(x_1) p_2(x_2) \ldots p_n(x_n), \tag{2.6.4}$$

in which case,

$$\phi(s_1, s_2, \ldots, s_n) = \phi_1(s_1)\phi_2(s_2) \ldots \phi_n(s_n) \tag{2.6.5}$$

(viii) Sum of independent random variables: if X_1, X_2, ..., are independent random variables and if

$$Y = \sum_{t=1}^{n} X_t , \tag{2.6.6}$$

and the characteristic function of Y is

$$\phi_y(s) = \langle \exp (\mathrm{i}sY) \rangle , \tag{2.6.7}$$

then

$$\phi_y(s) = \prod_{t=1}^{n} \phi_t(s) . \tag{2.6.8}$$

The characteristic function plays an important role in this book which arises from the convergence property (v), which allows us to perform limiting processes on the characteristic function rather than the probability distribution itself, and often makes proofs easier. Further, the fact that the characteristic function is truly characteristic, i.e., the inversion formula (vi), shows that different characteristic functions arise from different distributions. As well as this, the straightforward derivation of the moments by (2.6.2) makes any determination of the characteristic function directly relevant to measurable quantities.

2.7 Cumulant Generating Function: Correlation Functions and Cumulants

A further important property of the characteristic function arises by considering its logarithm

$$\Phi(s) = \log \phi(s) \tag{2.7.1}$$

which is called the *cumulant generating function*. Let us assume that all moments exist so that $\phi(s)$ and hence, $\Phi(s)$, is expandable in a power series which can be written as

$$\Phi(s) = \sum_{r=1}^{\infty} \mathrm{i}^r \sum_{\{m\}} \langle\!\langle X_1^{m_1} X_2^{m_2} \ldots X_n^{m_n} \rangle\!\rangle \frac{s_1^{m_1} s_2^{m_2} \ldots s_n^{m_n}}{m_1! \, m_2! \ldots m_n!} \delta(r, \sum_{t=1}^{n} m_t) , \tag{2.7.2}$$

where the quantities $\langle\!\langle X_1^{m_1} X_2^{m_2} \ldots X_n^{m_n} \rangle\!\rangle$ are called the *cumulants* of the variables X. The notation chosen should not be taken to mean that the cumulants are functions of the particular product of powers of the X; it rather indicates the moment of highest order which occurs in their expression in terms of moments. *Stratonovich*

[2.4] also uses the term *correlation functions*, a term which we shall reserve for cumulants which involve more than one X_t. For, if the X are all independent, the factorisation property (2.6.6) implies that $\Phi(s)$ (the cumulant generating function) is a sum of n terms, each of which is a function of only one s_t and hence the coefficient of mixed terms, i.e., the *correlation functions* (in our terminology) are all zero and the converse is also true. Thus, the magnitude of the correlation functions is a measure of the degree of correlation.

The cumulants and correlation functions can be evaluated in terms of moments by expanding the characteristic function as a power series:

$$\phi(s) = \sum_{r=1}^{\infty} \frac{i^r}{r!} \sum_{(m)} \langle X_1^{m_1} X_2^{m_2} \ldots X_n^{m_n} \rangle \frac{r!}{m_1! m_2! \ldots m_n!} \delta(r, \sum_{t=1}^{n} m_t)\, s_1^{m_1} s_2^{m_2} \ldots s_n^{m_n}, \quad (2.7.3)$$

expanding the logarithm in a power series, and comparing it with (2.7.2) for $\Phi(s)$ No *simple* formula can be given, but the first few cumulants can be exhibited: we find

$$\langle\langle X_t \rangle\rangle = \langle X_t \rangle \quad (2.7.4)$$

$$\langle\langle X_t X_j \rangle\rangle = \langle X_t X_j \rangle - \langle X_t \rangle \langle X_j \rangle \quad (2.7.5)$$

$$\langle\langle X_t X_j X_k \rangle\rangle = \langle X_t X_j X_k \rangle - \langle X_t X_j \rangle \langle X_k \rangle - \langle X_t \rangle \langle X_j X_k \rangle - \langle X_t X_k \rangle \langle X_j \rangle + 2\langle X_i \rangle \langle X_j \rangle \langle X_k \rangle \quad (2.7.6)$$

(here, all formulae are valid for any number of equal i,j,k,l). An explicit general formula can be given as follows. Suppose we wish to calculate the cumulant $\langle\langle X_1 X_2 X_3 \ldots X_n \rangle\rangle$. The procedure is the following:

i) write a sequence of n dots .....;

ii) divide into $p + 1$ subsets by inserting angle brackets

$$\langle \ldots \rangle \langle \ldots \rangle \langle \ldots\ldots \rangle \ldots \langle \ldots \rangle ;$$

iii) distribute the symbols $X_1 \ldots X_n$ in place of the dots in such a way that all *different* expressions of this kind occur, e.g.,

$$\langle X_1 \rangle \langle X_2 X_3 \rangle = \langle X_1 \rangle \langle X_3 X_2 \rangle \neq \langle X_3 \rangle \langle X_1 X_2 \rangle ;$$

iv) take the sum of all such terms for a given p. Call this $C_p(X_1, X_2, \ldots, X_n)$;

v) $$\langle\langle X_1\, X_2 \ldots X_n \rangle\rangle = \sum_{p=0}^{n-1} (-1)^p\, p!\, C_p(X_1, X_2, \ldots, X_n) . \quad (2.7.7)$$

A derivation of this formula was given by *Meeron* [2.6]. The particular procedure is due to *van Kampen* [2.7].

vi) *Cumulants in which there is one or more repeated element*, e.g., $\langle\langle X_1^2 X_3 X_2 \rangle\rangle$: simply evaluate $\langle\langle X_1 X_2 X_3 X_4 \rangle\rangle$ and set $X_4 = X_1$ in the resulting expression.

2.7.1 Example: Cumulant of Order 4: $\langle\langle X_1X_2X_3X_4\rangle\rangle$

a) $p = 0$

only term is $\langle X_1X_2X_3X_4\rangle \equiv C_0(X_1X_2X_3X_4)$.

b) $p = 1$

partition $\langle.\rangle\langle...\rangle$

Term $\{\langle X_1\rangle\langle X_2X_3X_4\rangle + \langle X_2\rangle\langle X_3X_4X_1\rangle + \langle X_3\rangle\langle X_4X_1X_2\rangle$
$+ \langle X_4\rangle\langle X_1X_2X_3\rangle\} \equiv D_1$

partition $\langle..\rangle\langle..\rangle$

Term $\langle X_1X_2\rangle\langle X_3X_4\rangle + \langle X_1X_3\rangle\langle X_2X_4\rangle + \langle X_1X_4\rangle\langle X_2X_3\rangle \equiv D_2$.

Hence,

$$D_1 + D_2 = C_1(X_1X_2X_3X_4)\,.$$

c) $p = 2$

partition $\langle.\rangle\langle.\rangle\langle..\rangle$

Term $\langle X_1\rangle\langle X_2\rangle\langle X_3X_4\rangle + \langle X_1\rangle\langle X_3\rangle\langle X_2X_4\rangle + \langle X_1\rangle\langle X_4\rangle\langle X_2X_3\rangle$
$+ \langle X_2\rangle\langle X_3\rangle\langle X_1X_4\rangle + \langle X_2\rangle\langle X_4\rangle\langle X_1X_3\rangle + \langle X_3\rangle\langle X_4\rangle\langle X_1X_2\rangle$
$= C_2(X_1X_2X_3X_4)$.

d) $p = 3$

partition $\langle.\rangle\langle.\rangle\langle.\rangle\langle.\rangle$

Term $\langle X_1\rangle\langle X_2\rangle\langle X_3\rangle\langle X_4\rangle = C_3(X_1X_2X_3X_4)$.

Hence,

$$\langle\langle X_1X_2X_3X_4\rangle\rangle = C_0 - C_1 + 2C_2 - 6C_3 \qquad (2.7.8$$

2.7.2 Significance of Cumulants

From (2.7.4, 5) we see that the first two cumulants are the means $\langle X_i\rangle$ and co variances $\langle X_i, X_j\rangle$. Higher-order cumulants contain information of decreasing significance, unlike higher-order moments. We cannot set all *moments* higher than a certain order equal to zero since $\langle X^{2n}\rangle \geqslant \langle X^n\rangle^2$ and thus, all moments contain information about lower moments.

For cumulants, however, we can consistently set

$$\langle\langle X\rangle\rangle = a$$
$$\langle\langle X^2\rangle\rangle = \sigma^2$$
$$\langle\langle X^n\rangle\rangle = 0 \ (n > 2),$$

and we can easily deduce by using the inversion formula for the characteristic function that

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp[-(x-a)^2/2\sigma^2], \tag{2.7.9}$$

a Gaussian probability distribution. It does not, however, seem possible to give more than this intuitive justification. Indeed, the theorem of *Marcienkiewicz* [2.8,9] shows that the cumulant generating function cannot be a polynomial of degree greater than 2, that is, either all but the first 2 cumulants vanish or there are an infinite number of nonvanishing cumulants. The greatest significance of cumulants lies in the definition of the correlation functions of different variables in terms of them; this leads further to important approximation methods.

2.8 Gaussian and Poissonian Probability Distributions

2.8.1 The Gaussian Distribution

By far the most important probability distribution is the Gaussian, or normal distribution. Here we collect together the most important facts about it.

If X is a vector of n Gaussian random variables, the corresponding multivariate probability density function can be written

$$p(\boldsymbol{x}) = [(2\pi)^n \det(\sigma)]^{-1/2} \exp[-\tfrac{1}{2}(\boldsymbol{x}-\bar{\boldsymbol{x}})^{\mathrm{T}} \sigma^{-1} (\boldsymbol{x}-\bar{\boldsymbol{x}})] \tag{2.8.1}$$

so that

$$\langle X \rangle = \int d\boldsymbol{x}\, \boldsymbol{x}\, p(\boldsymbol{x}) = \bar{\boldsymbol{x}} \tag{2.8.2}$$

$$\langle X X^{\mathrm{T}} \rangle = \int d\boldsymbol{x}\, \boldsymbol{x}\boldsymbol{x}^{\mathrm{T}} p(\boldsymbol{x}) = \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^{\mathrm{T}} + \sigma \tag{2.8.3}$$

and the characteristic function is given by

$$\phi(\boldsymbol{s}) = \langle \exp(\mathrm{i}\boldsymbol{s}^{\mathrm{T}} X) \rangle = \exp(\mathrm{i}\boldsymbol{s}^{\mathrm{T}} \bar{\boldsymbol{x}} - \tfrac{1}{2}\boldsymbol{s}^{\mathrm{T}} \sigma \boldsymbol{s}) . \tag{2.8.4}$$

This particularly simple characteristic function implies that all cumulants of higher order than 2 vanish, and hence means that all moments of order higher than 2 are expressible in terms of those of order 1 and 2. The relationship (2.8.3) means that σ is the covariance matrix (as defined in Sect. 2.5.1), i.e., the matrix whose elements are the second-order correlation functions. Of course, σ is symmetric.

The precise relationship between the higher moments and the covariance matrix σ can be written down straightforwardly by using the relationship between the moments and the characteristic function [Sect.2.6 (iv)]. The formula is only simple if $\bar{\boldsymbol{x}} = 0$, in which case the odd moments vanish and the even moments satisfy

$$\langle X_i X_j X_k \ldots \rangle = \frac{(2N)!}{N! 2^N} \{\sigma_{ij}\sigma_{kl}\sigma_{mn} \ldots\}_{\mathrm{sym}} , \tag{2.8.4}$$

where the subscript "sym" means the symmetrised form of the product of σ's, and $2N$ is the order of the moment. For example,

$$\langle X_1 X_2 X_3 X_4 \rangle = \frac{4!}{4.2!}\left\{\frac{1}{3}\left[\sigma_{12}\sigma_{34} + \sigma_{41}\sigma_{23} + \sigma_{13}\sigma_{24}\right]\right\}$$

$$= \sigma_{12}\sigma_{34} + \sigma_{41}\sigma_{23} + \sigma_{13}\sigma_{24} \tag{2.8.5}$$

$$\langle X_1^4 \rangle = \frac{4!}{4.2!}\{\sigma_{11}^2\} = 3\sigma_{11}^2 . \tag{2.8.6}$$

2.8.2 Central Limit Theorem

The Gaussian distribution is important for a variety of reasons. Many variables are in practice, empirically well approximated by Gaussians and the reason for this arises from the *central limit theorem*, which, roughly speaking, asserts that a random variable composed of the sum of many parts, each independent but arbitrarily distributed, is Gaussian. More precisely, let $X_1, X_2, X_3, \ldots, X_n$ be independent random variables such that

$$\langle X_i \rangle = 0, \qquad \text{var}\{X_i\} = b_i^2 \tag{2.8.7}$$

and let the distribution function of X_i be $p_i(x_i)$.

Define

$$S_n = \sum_{i=1}^{n} X_i , \tag{2.8.8}$$

and

$$\sigma_n^2 = \text{var}\{S_n\} = \sum_{i=1}^{n} b_i^2 . \tag{2.8.9}$$

We require further the fulfilment of the *Lindeberg condition*:

$$\lim_{n\to\infty}\left[\frac{1}{\sigma_n^2}\sum_{i=1}^{n}\int_{|x|>t\sigma_n} dx\, x^2\, p_i(x)\right] = 0 \tag{2.8.10}$$

for any fixed $t > 0$. Then, under these conditions, the distribution of the normalised sums S_n/σ_n tends to the Gaussian with zero mean and unit variance.

The proof of the theorem can be found in [2.1]. It is worthwhile commenting on the hypotheses, however. We first note that the summands X_i are required to be independent. This condition is not absolutely necessary; for example, choose

$$X_i = \sum_{r=i}^{i+j} Y_r \tag{2.8.11}$$

where the Y_j are independent. Since the sum of the X's can be rewritten as a sum of Y's (with certain finite coefficients), the theorem is still true.

Roughly speaking, as long as the correlation between X_i and X_j goes to zero sufficiently rapidly as $|i-j| \to \infty$, a central limit theorem will be expected. The Lindeberg condition (2.8.10) is not an obviously understandable condition but is the weakest condition which expresses the requirement that the probability for $|X_i|$ to be large is very small. For example, if all the b_i are infinite or greater than some constant C, it is clear that σ_n^2 diverges as $n \to \infty$. The sum of integrals in (2.8.10) is the sum of contributions to variances for all $|X_i| > t\sigma_n$, and it is clear that as $n \to \infty$, each contribution goes to zero. The Lindeberg condition requires the sum of all the contributions not to diverge as fast as σ_n^2. In practice, it is a rather weak requirement; satisfied if $|X_i| < C$ for all X_i, or if $p_i(x)$ go to zero sufficiently rapidly as $x \to \pm \infty$. An exception is

$$p_i(x) = a_i[\pi(x^2 + a_i^2)]^{-1} ; \tag{2.8.12}$$

the *Cauchy,* or *Lorentzian* distribution. The variance of this distribution is infinite and, in fact, the sum of all the X_i has a distribution of the same form as (2.8.12) with a_i replaced by $\sum_{i=1}^{n} a_i$. Obviously, the Lindeberg condition is not satisfied.

A related condition, also known as the Lindeberg condition, will arise in Sect. 3.3.1. where we discuss the replacement of a discrete process by one with continuous steps.

2.8.3 The Poisson Distribution

A distribution which plays a central role in the study of random variables which take on positive integer values is the Poisson distribution. If X is the relevant variable the Poisson distribution is defined by

$$P(X = x) \equiv P(x) = \mathrm{e}^{-\alpha}\alpha^x/x! \tag{2.8.13}$$

and clearly, the *factorial moments,* defined by

$$\langle X^r \rangle_{\mathrm{f}} = \langle x(x-1) \ldots (x-r+1) \rangle , \tag{2.8.14}$$

are given by

$$\langle X^r \rangle_{\mathrm{f}} = \alpha^r , \tag{2.8.15}$$

For variables whose range is nonnegative integral, we can very naturally define the *generating function*

$$G(s) = \sum_{x=0}^{\infty} s^x P(x) = \langle s^x \rangle \tag{2.8.16}$$

which is related to the characteristic function by

$$G(s) = \phi(-\mathrm{i} \log s) . \tag{2.8.17}$$

The generating function has the useful property that

$$\langle X^r \rangle_{\mathrm{f}} = \left[\left(\frac{\partial}{\partial s} \right)^r G(s) \right]_{s=1} . \tag{2.8.18}$$

For the Poisson distribution we have

$$G(s) = \sum_{x=0}^{\infty} \frac{\mathrm{e}^{-\alpha}(s\alpha)^x}{x!} = \exp[\alpha(s-1)] . \tag{2.8.19}$$

We may also define the factorial cumulant generating function $g(s)$ by

$$g(s) = \log G(s) \tag{2.8.20}$$

and the *factorial cumulants* $\langle\langle X^r \rangle\rangle_{\mathrm{f}}$ by

$$g(s) = \sum_{x=1}^{\infty} \langle\langle X^r \rangle\rangle_{\mathrm{f}} \frac{(s-1)^r}{r!} .$$

We see that the Poisson distribution has all but the first factorial cumulant zero.

The Poisson distribution arises naturally in very many contexts, for example, we have already met it in Sect.1.2.3 as the solution of a simple master equation. It plays a similar central role in the study of random variables which take on integer values to that occupied by the Gaussian distribution in the study of variables with a continuous range. However, the only simple multivariate generalisation of the Poisson is simply a product of Poissons, i.e., of the form

$$P(x_1, x_2, x_3, \ldots) = \prod_{i=1}^{n} \frac{\mathrm{e}^{-\alpha_i}(\alpha_i)^{x_i}}{x_i!} . \tag{2.8.21}$$

There is no logical concept of a correlated multipoissonian distribution, similar to that of a correlated multivariate Gaussian distribution.

2.9 Limits of Sequences of Random Variables

Much of computational work consists of determining *approximations* to random variables, in which the concept of a *limit of a sequence of random variables* naturally arises. However, there is no unique way of defining such a limit.

For, suppose we have a probability space Ω, and a sequence of random variables X_n defined on Ω. Then by the limit of the sequence as $n \to \infty$

$$X = \lim_{n \to \infty} X_n , \tag{2.9.1}$$

we mean a random variable X which, in some sense, is approached by the sequence

of random variables X_n. The various possibilities arise when one considers that the probability space Ω has elements ω which have a probability density $p(\omega)$. Then we can choose the following definitions.

2.9.1 Almost Certain Limit

X_n converges *almost certainly* to X if, for all ω except a set of probability zero

$$\lim_{n\to\infty} X_n(\omega) = X(\omega). \tag{2.9.2}$$

Thus each realisation of X_n converges to X and we write

$$\text{ac-lim}_{n\to\infty} X_n = X \tag{2.9.3}$$

2.9.2 Mean Square Limit (Limit in the Mean)

Another possibility is to regard the $X_n(\omega)$ as functions of ω, and look for the mean square deviation of $X_n(\omega)$ from $X(\omega)$. Thus, we say that X_n converges to X in the *mean square* if

$$\lim_{n\to\infty} \int d\omega\, p(\omega)[X_n(\omega) - X(\omega)]^2 \equiv \lim_{n\to\infty} \langle (X_n - X)^2 \rangle = 0 \tag{2.9.4}$$

This is the kind of limit which is well known in Hilbert space theory. We write

$$\text{ms-lim}_{n\to\infty} X_n = X \tag{2.9.5}$$

2.9.3 Stochastic Limit, or Limit in Probability

We can consider the possibility that $X_n(\omega)$ approaches X because the probability of deviation from X approaches zero: precisely, this means that if for any $\varepsilon > 0$

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0 \tag{2.9.6}$$

then the *stochastic limit* of X_n is X.

Note that the probability can be written as follows. Suppose $\chi_\varepsilon(z)$ a function such that

$$\begin{aligned} \chi_\varepsilon(z) &= 1 \qquad |z| > \varepsilon \\ &= 0 \qquad |z| \leqslant \varepsilon \end{aligned} \tag{2.9.7}$$

Then

$$P(|X_n - X| > \varepsilon) = \int d\omega\, p(\omega)\chi_\varepsilon(|X_n - X|)\,. \tag{2.9.8}$$

In this case, we write

$$\text{st-}\lim_{n\to\infty} X_n = X \tag{2.9.9}$$

2.9.4 Limit in Distribution

An even weaker form of convergence occurs if, for any continuous bounded function $f(x)$

$$\lim_{n\to\infty} \langle f(X_n)\rangle = \langle f(X)\rangle \,. \tag{2.9.10}$$

In this case the convergence of the limit is said to be *in distribution.* In particular, using $\exp(ixs)$ for $f(x)$, we find that the characteristic functions approach each other, and hence the probability density of X_n approaches that of X.

2.9.5 Relationship Between Limits

The following relations can be shown.

Almost certain convergence $\Longrightarrow$ stochastic convergence.
Convergence in mean square $\Longrightarrow$ stochastic convergence.
Stochastic convergence $\Longrightarrow$ convergence in distribution.

All of these limits have uses in applications.

3. Markov Processes

3.1 Stochastic Processes

All of the examples given in Chap. 1 can be mathematically described as *stochastic processes* by which we mean, in a loose sense, systems which evolve probabilistically in time or more precisely, systems in which a certain time-dependent random variable $\boldsymbol{X}(t)$ exists. We can measure values $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots$, etc., of $\boldsymbol{X}(t)$ at times t_1, $t_2, t_3, \ldots$ and we assume that a set of joint probability densities exists

$$p(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2; \boldsymbol{x}_3, t_3; \ldots) \tag{3.1.1}$$

which describe the system completely.

In terms of these joint probability density functions, one can also define conditional probability densities:

$$\begin{aligned} &p(\boldsymbol{x}_1, t_1,; \boldsymbol{x}_2, t_2; \ldots | \boldsymbol{y}_1, \tau_1; \boldsymbol{y}_2, \tau_2; \ldots) \\ &\quad = p(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2; \ldots; \boldsymbol{y}_1, \tau_1; \boldsymbol{y}_2, \tau_2; \ldots)/p(\boldsymbol{y}_1, \tau_1; \boldsymbol{y}_2, \tau_2; \ldots). \end{aligned} \tag{3.1.2}$$

These definitions are valid independently of the ordering of the times, although it is usual to consider only times which increase from right to left i.e.,

$$t_1 \geqslant t_2 \geqslant t_3 \geqslant \ldots \geqslant \tau_1 \geqslant \tau_2 \geqslant \ldots . \tag{3.1.3}$$

The concept of an evolution equation leads us to consider the conditional probabilities as predictions of the future values of $\boldsymbol{X}(t)$ (i.e., $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ at times $t_1, t_2, \ldots$), given the knowledge of the past (values $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots$, at times $\tau_1, \tau_2, \ldots$).

The concept of a general stochastic process is very loose. To define the process we need to know at least all possible joint probabilities of the kind in (3.1.1). If such knowledge does define the process, it is known as a *separable stochastic process.* All the processes considered in this book will be assumed to be separable.

The most simple kind of stochastic process is that of complete independence:

$$p(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2; \boldsymbol{x}_3, t_3; \ldots) = \prod_i p(\boldsymbol{x}_i, t_i) \tag{3.1 4}$$

which means that the value of $\boldsymbol{X}$ at time t is completely independent of its values in the past (or future). An even more special case occurs when the $p(\boldsymbol{x}_i, t_i)$ are independent of t_i, so that the same probability law governs the process at all times. We then have the *Bernoulli trials*, in which a probabilistic process is repeated at successive times.

The next most simple idea is that of the *Markov process* in which knowledge o only the present determines the future.

3.2 Markov Process

The *Markov assumption* is formulated in terms of the conditional probabilities. W require that if the times satisfy the ordering (3.1.3), the conditional probability i determined entirely by the knowledge of the most recent condition, i.e.,

$$p(x_1, t_1; x_2, t_2; \ldots \,|\, y_1, \tau_1; y_2, \tau_2; \ldots) = p(x_1, t_1; x_2, t_2; \ldots \,|\, y_1, \tau_1). \tag{3.2.1}$$

This is simply a more precise statement of the assumptions made by Einstein Smoluchowski and others. It is, even by itself, extremely powerful. For it mean that we can define everything in terms of the simple conditional probabilitie $p(x_1, t_1 \,|\, y_1, \tau_1)$. For example, by definition of the conditional probability densit $p(x_1, t_1; x_2, t_2 \,|\, y_1, \tau_1) = p(x_1, t_1 \,|\, x_2, t_2; y_1, \tau_1) p(x_2, t_2 \,|\, y_1, \tau_1)$ and using the Marko assumption (3.2.1), we find

$$p(x_1, t_1; x_2, t_2; y_1, \tau_1) = p(x_1, t_1 \,|\, x_2, t_2) p(x_2, t_2 \,|\, y_1, \tau_1) \tag{3.2.2}$$

and it is not difficult to see that an arbitrary joint probability can be expressed sim ply as

$$\begin{aligned} &p(x_1, t_1; x_2, t_2; x_3, t_3; \ldots x_n, t_n) \\ &\quad = p(x_1, t_1 \,|\, x_2, t_2) p(x_2, t_2 \,|\, x_3, t_3) p(x_3, t_3 \,|\, x_4, t_4) \ldots \\ &\qquad \ldots p(x_{n-1}, t_{n-1} \,|\, x_n, t_n) p(x_n, t_n) \end{aligned} \tag{3.2.3}$$

provided

$$t_1 \geqslant t_2 \geqslant t_3 \geqslant \ldots \geqslant t_{n-1} \geqslant t_n \,. \tag{3.2.4}$$

3.2.1 Consistency—the Chapman-Kolmogorov Equation

From Sect.2.3.3 we require that summing over all mutually exclusive events c one kind in a joint probability eliminates that variable, i.e.,

$$\sum_B P(A \cap B \cap C \ldots) = P(A \cap C \ldots) \,; \tag{3.2.5}$$

and when this is applied to stochastic processes, we get two deceptively simil equations:

$$\begin{aligned} p(x_1, t_1) &= \int dx_2 \, p(x_1, t_1; x_2, t_2) \\ &= \int dx_2 \, p(x_1, t_1 \,|\, x_2, t_2) p(x_2, t_2) \,. \end{aligned} \tag{3.2.6}$$

This equation is an identity valid for all stochastic processes and is the first in a hierarchy of equations, the second of which is

$$p(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_3, t_3) = \int d\boldsymbol{x}_2\, p(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2 | \boldsymbol{x}_3, t_3)$$
$$= \int d\boldsymbol{x}_2\, p(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_2, t_2; \boldsymbol{x}_3, t_3) p(\boldsymbol{x}_2, t_2 | \boldsymbol{x}_3, t_3). \tag{3.2.7}$$

This equation is also always valid. We now introduce the Markov assumption. If $t_1 \geqslant t_2 \geqslant t_3$, we can drop the t_3 dependence in the doubly conditioned probability and write

$$p(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_3, t_3) = \int d\boldsymbol{x}_2\, p(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_2, t_2) p(\boldsymbol{x}_2, t_2 | \boldsymbol{x}_3, t_3) \tag{3.2.8}$$

which is the *Chapman-Kolmogorov equation.*

What is the essential difference between (3.2.8) and (3.2.6)? The obvious answer is that (3.2.6) is for unconditioned probabilities, whereas (3.2.7) is for conditional probabilities. Equation (3.2.8) is a rather complex nonlinear functional equation relating all conditional probabilities $p(\boldsymbol{x}_i, t_i | \boldsymbol{x}_j, t_j)$ to each other, whereas (3.2.6) simply constructs the one time probabilities in the future t_1 of t_2, given the conditional probability $p(\boldsymbol{x}_1, t_1 | \boldsymbol{x}_2, t_2)$.

The Chapman-Kolmogorov equation has many solutions. These are best understood by deriving the differential form which is done in Sect. 3.4.1 under certain rather mild conditions.

3.2.2 Discrete State Spaces

In the case where we have a discrete variable, we will use the symbol $\boldsymbol{N} = (N_1, N_2, N_3 \ldots)$, where the N_i are random variables which take on integral values. Clearly, we now replace

$$\int d\boldsymbol{x} \leftrightarrow \sum_{\boldsymbol{n}} \tag{3.2.9}$$

and we can now write the Chapman-Kolmogorov equation for such a process as

$$P(\boldsymbol{n}_1, t_1 | \boldsymbol{n}_3, t_3) = \sum_{\boldsymbol{n}_2} P(\boldsymbol{n}_1, t_1 | \boldsymbol{n}_2, t_2)\, P(\boldsymbol{n}_2, t_2 | \boldsymbol{n}_3, t_3)\,. \tag{3.2.10}$$

This is now a matrix multiplication, with possibly infinite matrices.

3.2.3 More General Measures

A more general formulation would assume a measure $d\mu(\boldsymbol{x})$ instead of $d\boldsymbol{x}$ where a variety of choices can be made. For example, if $\mu(\boldsymbol{x})$ is a step function with steps at integral values of $\boldsymbol{x}$, we recover the discrete state space form. Most mathematical works attempt to be as general as possible. For applications, such generality can lead to lack of clarity so, where possible, we will favour a more specific notation.

3.3 Continuity in Stochastic Processes

Whether or not the random variable $X(t)$ has a continuous range of possible values is a completely different question from whether the sample path of $X(t)$ is a continuous function of t. For example, in a gas composed of molecules with velocities $V(t)$, it is clear that all possible values of $V(t)$ are in principle realisable, so that the *range* of $V(t)$ is continuous. However, a model of collisions in a gas of hard spheres as occurring instantaneously is often considered, and in such a model the velocity before the collision, v_i, will change instantaneously at the time of impact to another value v_f, so the sample path of $V(t)$ is not continuous. Nevertheless, in such a model, the *position* of a gas molecule $X(t)$ would be expected to be continuous.

A major question now arises. Do *Markov* processes with *continuous sample paths* actually exist in reality? Notice the combination of *Markov* and *continuous*. It is almost certainly the case that in a classical picture (i.e., not quantum mechanical), all variables with a continuous range have continuous sample paths. Even the hard sphere gas mentioned above is an idealisation and more realistically, one should allow some potential to act which would continuously deflect the molecules during a collision. But it would also be the case that, if we observe on such a fine time scale, the process will probably not be Markovian. The immediate history of the whole system will almost certainly be required to predict even the probabilistic future. This is certainly born out in all attempts to derive Markovian probabilistic equations from mechanics. Equations which are derived are rarely truly Markovian—rather there is a certain characteristic memory time during which the previous history is important (*Haake* [3.1]).

This means that there is really no such thing as a Markov process; rather, there may be systems whose memory time is so small that, on the time scale on which we carry out observations, it is fair to regard them as being well approximated by a Markov process. But in this case, the question of whether the sample paths are actually continuous is not relevant. The sample paths of the approximating Markov process certainly need not be continuous. Even if collisions of molecules are not accurately modelled by hard spheres, during the time taken for a collision, a finite change of velocity takes place and this will appear in the approximating Markov process as a discrete step. On this time scale, even the position may change discontinuously, thus giving the picture of Brownian motion as modelled by Einstein.

In chemical reactions, for example, the time taken for an individual reaction to proceed to completion—roughly of the same order of magnitude as the collision time for molecules—provides yet another minimum time, since during this time, states which cannot be described in terms of individual molecules exist. Here, therefore, the very description of the state in terms of individual molecules requires a certain minimum time scale to be considered.

However, Markov processes with continuous sample paths do exist mathematically and are useful in describing reality. The model of the gas mentioned above provides a useful example. The position of the molecule is indeed probably best

modelled as changing discontinuously by discrete jumps. Compared to the distances travelled, however, these jumps are infinitesimal and a continuous curve provides a good approximation to the sample path. On the other hand, the velocities can change by amounts which are of the same order of magnitude as typical values attained in practice. The average velocity of a molecule in a gas is about 1000 m/s and during a collision can easily reverse its sign. The velocities simply cannot reach (with any significant probability) values for which the changes of velocity can be regarded as very small. Hence, there is no sense in a continuous path description of velocities in a gas.

3.3.1 Mathematical Definition of a Continuous Markov Process

For a *Markov process*, it can be shown [3.2] that with probability one, the sample paths are continuous functions of t, if for any $\varepsilon > 0$ we have

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|>\varepsilon} dx\, p(x, t + \Delta t | z, t) = 0 \tag{3.3.1}$$

uniformly in z, t and Δt.

This means that the probability for the final position x to be finitely different from z goes to zero *faster* than Δt, as Δt goes to zero. [Equation (3.3.1) is sometimes called the Lindeberg condition.]

Examples

i) Einstein's solution for his $f(x, t)$ (Sect. 1.2.1) is really the conditional probability $p(x, t|0, 0)$. Following his method we would find

$$p(x, t + \Delta t | z, t) = (4\pi D \Delta t)^{-1/2} \exp\left[-(x - z)^2/4D\Delta t\right] \tag{3.3.2}$$

and it is easy to check that (3.3.1) is satisfied in this case. Thus, Brownian motion in Einstein's formulation has continuous sample paths.

ii) Cauchy Process: Suppose

$$p(x, t + \Delta t | z, t) = \frac{\Delta t}{\pi} / [(x - z)^2 + \Delta t^2]\,. \tag{3.3.3}$$

Then this does not satisfy (3.3.1) so the sample paths are discontinuous.

However, in both cases, we have as required for consistency

$$\lim_{\Delta t \to 0} p(x, t + \Delta t | z, t) = \delta(x - z)\,, \tag{3.3.4}$$

and it is easy to show that in both cases, the Chapman-Kolomogorov equation is satisfied.

The difference between the two processes just described is illustrated in Fig. 3.1 in which simulations of both processes are given. The difference between the two is

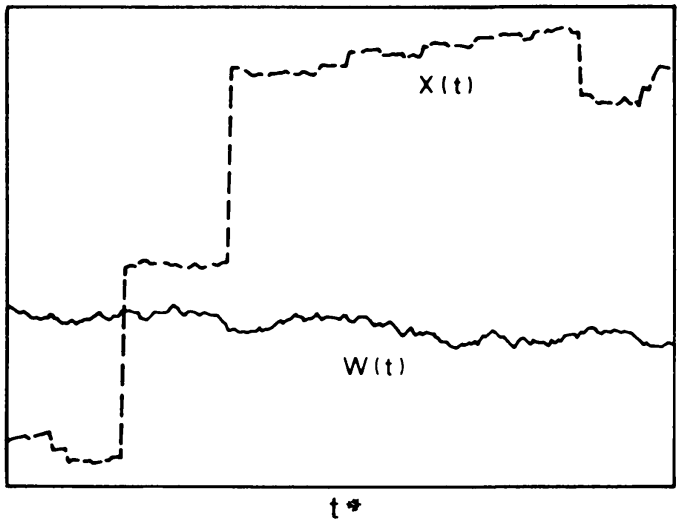


Fig. 3.1. Illustration of sample paths of the Cauchy process $X(t)$ (- - - - -) and Brownian motion $W(t)$ (———)

striking. Notice, however, that even the Brownian motion curve is extremely irregular, even though continuous—in fact it is nowhere differentiable. The Cauchy-process curve is, of course, wildly discontinuous.

3.4 Differential Chapman-Kolmogorov Equation

Under appropriate assumptions, the Chapman-Kolmogorov equation can be reduced to a differential equation. The assumptions made are closely connected with the continuity properties of the process under consideration. Because of the form of the continuity condition (3.3.1), one is led to consider a method of dividing the differentiability conditions into parts, one corresponding to continuous motion of a representative point and the other to discontinuous motion.

We require the following conditions *for all* $\varepsilon > 0$:

$$\text{i)} \lim_{\Delta t \to 0} p(\boldsymbol{x}, t + \Delta t \,|\, \boldsymbol{z}, t)/\Delta t = W(\boldsymbol{x} \,|\, \boldsymbol{z}, t) \tag{3.4.1}$$

uniformly in $\boldsymbol{x}$, $\boldsymbol{z}$, and t for $|\boldsymbol{x} - \boldsymbol{z}| \geqslant \varepsilon$;

$$\text{ii)} \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|\boldsymbol{x}-\boldsymbol{z}|<\varepsilon} d\boldsymbol{x}\,(x_i - z_i) p(\boldsymbol{x}, t + \Delta t \,|\, \boldsymbol{z}, t) = A_i(\boldsymbol{z}, t) + O(\varepsilon)\ ; \tag{3.4.2}$$

$$\text{iii)} \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|\boldsymbol{x}-\boldsymbol{z}|<\varepsilon} d\boldsymbol{x}\,(x_i - z_i)(x_j - z_j) p(\boldsymbol{x}, t + \Delta t \,|\, \boldsymbol{z}, t) = B_{ij}(\boldsymbol{z}, t) + O(\varepsilon); \tag{3.4.3}$$

the last two being uniform in $\boldsymbol{z}$, ε, and t.

Notice that all higher-order coefficients of the form in (3.4.2,3) must vanish. For example, consider the third-order quantity defined by

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|\boldsymbol{x}-\boldsymbol{z}|<\varepsilon} dx\,(x_i - z_i)(x_j - z_j)(x_k - z_k)\, p(\boldsymbol{x}, t + \Delta t \,|\, \boldsymbol{z}, t)$$
$$\equiv C_{ijk}(\boldsymbol{z}, t) + O(\varepsilon)\,. \tag{3.4.4}$$

Since C_{ijk} is symmetric in i, j, k, consider

$$\sum_{i,j,k} \alpha_i \alpha_j \alpha_k C_{ijk}(z, t) \equiv \bar{C}(\alpha, z, t) \tag{3.4.5}$$

so that

$$C_{ijk}(z, t) = \frac{1}{3!} \frac{\partial^3}{\partial \alpha_i \partial \alpha_j \partial \alpha_k} \bar{C}(\alpha, z, t) . \tag{3.4.6}$$

Then,

$$\begin{aligned} |\bar{C}(\alpha, z, t)| &\leqslant \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} |\alpha \cdot (x - z)| [\alpha \cdot (x - z)]^2 \, p(x, t + \Delta t | z, t) \, dx \\ &\quad + O(\varepsilon) \\ &\leqslant |\alpha| \varepsilon \lim_{\Delta t \to 0} \int [\alpha \cdot (x - z)]^2 p(x, t + \Delta t | z, t) dx + O(\varepsilon) \\ &= \varepsilon |\alpha| [\alpha_i \alpha_j B_{ij}(z, t) + O(\varepsilon)] + O(\varepsilon) \\ &= O(\varepsilon) \end{aligned} \tag{3.4.7}$$

so that C is zero. Similary, we can show that all corresponding higher-order quantities also vanish.

According to the condition for continuity (3.3.1), the process can only have continuous paths if $W(x | z, t)$ vanishes for all $x \neq z$. Thus, this function must in some way describe discontinuous motion, while the quantities A_i and B_{ij} must be connected with continuous motion.

3.4.1 Derivation of the Differential Chapman-Kolmogorov Equation

We consider the time evolution of the expectation of a function $f(z)$ which is twice continuously differentiable.

Thus,

$$\begin{aligned} &\partial_t \int dx \, f(x) p(x, t | y, t') \\ &\quad = \lim_{\Delta t \to 0} \{ \int dx \, f(x) [p(x, t + \Delta t | y, t') - p(x, t | y, t')] \} / \Delta t \qquad (3.4.8) \\ &\quad = \lim_{\Delta t \to 0} \{ \int dx \int dz \, f(x) p(x, t + \Delta t | z, t) p(z, t | y, t') \\ &\qquad - \int dz f(z) p(z, t | y, t') \} / \Delta t , \qquad (3.4.9) \end{aligned}$$

where we have used the Chapman-Kolmogorov equation in the positive term of (3.4.8) to produce the corresponding term in (3.4.9).

We now divide the integral over x into two regions $|x - z| \geqslant \varepsilon$ and $|x - z| < \varepsilon$. When $|x - z| < \varepsilon$, since $f(z)$ is, by assumption, twice continuously differentiable, we may write

$$f(\boldsymbol{x}) = f(\boldsymbol{z}) + \sum_i \frac{\partial f(\boldsymbol{z})}{\partial z_i}(x_i - z_i) + \sum_{i,j} \frac{1}{2}\frac{\partial^2 f(\boldsymbol{z})}{\partial z_i \partial z_j}(x_i - z_i)(x_j - z_j) + |\boldsymbol{x} - \boldsymbol{z}|^2 R(\boldsymbol{x}, \boldsymbol{z}), \tag{3.4.10}$$

where we have (again by the twice continuous differentiability)

$$|R(\boldsymbol{x}, \boldsymbol{z})| \to 0 \quad \text{as} \quad |\boldsymbol{x} - \boldsymbol{z}| \to 0. \tag{3.4.11}$$

Now substitute in (3.4.9):

$$\begin{aligned}(3.4.9) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Bigg\{ & \iint_{|x-z|<\varepsilon} d\boldsymbol{x} d\boldsymbol{z} \left[\sum_i (x_i - z_i)\frac{\partial f}{\partial z_i} + \sum_{i,j} \frac{1}{2}(x_i - z_i)(x_j - z_j)\frac{\partial^2 f}{\partial z_i \partial z_j} \right] \\ & \times p(\boldsymbol{x}, t + \Delta t | \boldsymbol{z}, t)\, p(\boldsymbol{z}, t | \boldsymbol{y}, t') \\ & + \iint_{|x-z|<\varepsilon} d\boldsymbol{x} d\boldsymbol{z}\, |\boldsymbol{x} - \boldsymbol{z}|^2 R(\boldsymbol{x}, \boldsymbol{z}) p(\boldsymbol{x}, t + \Delta t | \boldsymbol{z}, t) p(\boldsymbol{z}, t | \boldsymbol{y}, t') \\ & + \iint_{|x-z|\geqslant\varepsilon} d\boldsymbol{x} d\boldsymbol{z}\, f(\boldsymbol{x}) p(\boldsymbol{x}, t + \Delta t | \boldsymbol{z}, t) p(\boldsymbol{z}, t | \boldsymbol{y}, t') \\ & + \iint_{|x-z|<\varepsilon} d\boldsymbol{x} d\boldsymbol{z}\, f(\boldsymbol{z}) p(\boldsymbol{x}, t + \Delta t | \boldsymbol{z}, t) p(\boldsymbol{z}, t | \boldsymbol{y}, t') \\ & - \iint d\boldsymbol{x} d\boldsymbol{z}\, f(\boldsymbol{z}) p(\boldsymbol{x}, t + \Delta t | \boldsymbol{z}, t) p(\boldsymbol{z}, t | \boldsymbol{y}, t') \Bigg\}\end{aligned} \tag{3.4.12}$$

[notice that since $p(\boldsymbol{x}, t + \Delta t | \boldsymbol{z}, t)$ is a probability, the integral over $\boldsymbol{x}$ in the last term gives 1—this is simply the last term in (3.4.9)].

We now consider these line by line.

Lines 1,2: by the assumed uniform convergence, we take the limit inside the integral to obtain [using conditions (ii) and (iii) of Sect. 3.4]

$$\int d\boldsymbol{z} \left[\sum_i A_i(\boldsymbol{z}) \frac{\partial f}{\partial z_i} + \frac{1}{2} \sum_{i,j} B_{ij}(\boldsymbol{z}) \frac{\partial^2 f}{\partial z_i \partial z_j} \right] p(\boldsymbol{z}, t | \boldsymbol{y}, t') + O(\varepsilon). \tag{3.4.13}$$

Line 3: this is a remainder term and vanishes as $\varepsilon \to 0$. For

$$\begin{aligned} & \left| \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} d\boldsymbol{x}\, |\boldsymbol{x} - \boldsymbol{z}|^2 R(\boldsymbol{x}, \boldsymbol{z}) p(\boldsymbol{x}, t + \Delta t | \boldsymbol{z}, t) \right| \\ & \leqslant \left[\frac{1}{\Delta t} \int_{|x-z|<\varepsilon} d\boldsymbol{x}\, |\boldsymbol{x} - \boldsymbol{z}|^2 p(\boldsymbol{x}, t + \Delta t | \boldsymbol{z}, t) \right] \operatorname*{Max}_{|x-z|<\varepsilon} |R(\boldsymbol{x}, \boldsymbol{z})| \\ & \to [\textstyle\sum_i B_{ii}(\boldsymbol{z}, t) + O(\varepsilon)] \ \{\operatorname*{Max}_{|x-z|<\varepsilon} |R(\boldsymbol{x}, \boldsymbol{z})|\}. \end{aligned} \tag{3.4.14}$$

From (3.4.11) we can see that as $\varepsilon \to 0$, the factor in curly brackets vanishes.

Lines 4–6: We can put these all together to obtain

$$\iint_{|x-z|\geqslant\varepsilon} d\boldsymbol{x} d\boldsymbol{z}\, f(\boldsymbol{z}) [W(\boldsymbol{z} | \boldsymbol{x}, t) p(\boldsymbol{x}, t | \boldsymbol{y}, t') - W(\boldsymbol{x} | \boldsymbol{z}, t) p(\boldsymbol{z}, t | \boldsymbol{y}, t')]. \tag{3.4.15}$$

The whole right-hand side of (3.4.12) is independent of ε. Hence, taking the limit $\varepsilon \to 0$, we find

$$\partial_t \int dz\, f(z)p(z,t|y,t') = \int dz \left[\sum_i A_i(z,t)\frac{\partial f(z)}{\partial z_i} + \frac{1}{2}\sum_{i,j} B_{ij}(z)\frac{\partial^2 f(z)}{\partial z_i \partial z_j}\right] p(z,t|y,t')$$
$$+ \int dz\, f(z)\{\int dx[W(x|z,t)p(x,t|y,t') - W(x|z,t)p(z,t|y,t')]\}\ . \quad (3.4.16)$$

Notice, however, that we use the definition

$$\lim_{\varepsilon\to 0} \int_{|x-z|>\varepsilon} dx\, F(x,z) \equiv \int dx\, F(x,z) \quad (3.4.17)$$

for a principal value integral of a function $F(x,z)$. For (3.4.16) to have any meaning, this integral should exist. Equation (3.4.1) defines $W(x|z,t)$ only for $x \neq z$ and hence leaves open the possibility that it is infinite at $x = z$, as is indeed the case for the Cauchy process, discussed in Sect. 3.3.1, for which

$$W(x|z,t) = 1/[\pi(x-z)^2]\ . \quad (3.4.18)$$

However, if $p(x,t|y,t')$ is continuous and once differentiable, then the principal value integral exists. In the remainder of the book we shall not write this integral explicitly as a principal value integral since one rarely considers the singular cases for which it is necessary.

The final step now is to integrate by parts. We find

$$\int dz\, f(z)\partial_t p(z,t|y,t') = \int dz\, f(z)\Big\{-\sum_i \frac{\partial}{\partial z_i} A_i(z,t)p(z,t|y,t')$$
$$+ \sum_{i,j}\frac{1}{2}\frac{\partial^2}{\partial z_i \partial z_j} B_{ij}(z,t)p(z,t|y,t')$$
$$+ \int dx[W(z|x,t)p(x,t|y,t') - W(x|z,t)p(z,t|y,t')]\Big\}$$
$$+ \text{surface terms.} \quad (3.4.19)$$

We have not specified the range of the integrals. Suppose the process is confined to a region R with surface S. Then clearly,

$$p(x,t|z,t') = 0 \text{ unless both } x \text{ and } z \in R. \quad (3.4.20)$$

It is clear that by definition we have

$$W(x|z,t) = 0 \text{ unless both } x \text{ and } y \in R. \quad (3.4.21)$$

But the conditions on $A_i(z,t)$ and $B_{ij}(z,t)$ can result in discontinuities in these functions as defined by (3.4.2.3) since the conditional probability $p(x,t+\Delta t|z,t')$ can very reasonably change discontinuously as z crosses the boundary of R, reflecting the fact that no transitions are allowed from outside R to inside R.

In integrating by parts, we are forced to differentiate both A_t and B_{tj} and by our reasoning above, one cannot assume that this is possible on the boundary of the region. Hence, let us choose $f(\mathbf{z})$ to be arbitrary but nonvanishing only in an arbitrary region R' entirely contained in R. We can then deduce that for all $\mathbf{z}$ in *the interior* of R,

$$\begin{aligned}\partial_t p(\mathbf{z}, t|\mathbf{y}, t') = &-\sum_t \frac{\partial}{\partial z_t}[A_t(\mathbf{z}, t)p(\mathbf{z}, t|\mathbf{y}, t')] \\ &+ \sum_{t,j} \frac{1}{2}\frac{\partial^2}{\partial z_t \partial z_j}[B_{tj}(\mathbf{z}, t)p(\mathbf{z}, t|\mathbf{y}, t')] \\ &+ \int d\mathbf{x}\,[W(\mathbf{z}|\mathbf{x}, t)p(\mathbf{x}, t|\mathbf{y}, t') - W(\mathbf{x}|\mathbf{z}, t)p(\mathbf{z}, t|\mathbf{y}, t')]\,.\end{aligned} \tag{3.4.22}$$

Surface terms do not arise, since they necessarily vanish.

This equation does not seem to have any agreed name in the literature. Since it is purely a differential form of the Chapman-Kolmogorov equation, I propose to call it the *differential Chapman-Kolmogorov* equation.

3.4.2 Status of the Differential Chapman-Kolmogorov Equation

From our derivation it is not clear to what extent solutions of the differential Chapman-Kolmogorov equation are solutions of the Chapman-Kolmogorov equation itself or indeed, to what extent solutions exist. It is certainly true, however, that a set of conditional probabilities which obey the Chapman-Kolmogorov equation does generate a Markov process, in the sense that the joint probabilities so generated satisfy all probability axioms.

It can be shown [3.3] that, under certain conditions, if we specify $\boldsymbol{A}(\mathbf{x}, t)$, $\underline{\boldsymbol{B}}(\mathbf{x}, t)$ (which must be positive semi-definite), and $W(\mathbf{x}|\mathbf{y}, t)$ (which must be non-negative), that a non-negative solution to the differential Chapman-Kolmogorov equation exists, and this solution also satisfies the Chapman-Kolmogorov equation. The conditions to be satisfied are the *initial condition,*

$$p(\mathbf{z}, t|\mathbf{y}, t) = \delta(\mathbf{y} - \mathbf{z})$$

which follows from the definition of the conditional probability density, and any appropriate boundary conditions. These are very difficult to specify in the full equation, but in the case of the *Fokker-Planck equation* (Sect. 3.5.2) are given in Chap. 5.

3.5 Interpretation of Conditions and Results

Each of the conditions (i), (ii), (iii) of Sect. 3.4 can now be seen to give rise to a distinctive part of the equation, whose interpretation is rather straightforward. We can identify three processes taking place, which are known as jumps, drift and diffusion.

3.5.1 Jump Processes: The Master Equation

We consider a case in which

$$A_i(z, t) = B_{ij}(z, t) = 0 \tag{3.5.1}$$

so that we now have the *Master equation:*

$$\partial_t p(z, t|y, t') = \int dx\, [W(z|x, t)p(x, t|y, t') - W(x|z, t)p(z, t|y, t')]\,. \tag{3.5.2}$$

To first order in Δt we solve approximately, as follows. Notice that

$$p(z, t|y, t) = \delta(y - z)\,. \tag{3.5.3}$$

Hence,

$$p(z, t + \Delta t|y, t) = \delta(y - z)[1 - \int dx\, W(x|y, t)\Delta t] + W(z|y, t)\Delta t\,. \tag{3.5.4}$$

We see that for any Δt there is a finite probability, given by the coefficient of the $\delta(y - z)$ in (3.5.4), for the particle to stay at the original position y. The distribution of those particles which do not remain at y is given by $W(z|y, t)$ after appropriate normalisation. Thus, a typical path $X(t)$ will consist of sections of straight lines $X(t) =$ constant, interspersed with discontinuous jumps whose distribution is given by $W(z|y, t)$. For this reason, the process is known as a jump process. The paths are discontinuous at discrete points.

In the case where the state space consists of integers only, the Master equation takes the form

$$\partial_t P(n, t|n', t') = \sum_m [W(n|m, t)P(m, t|n', t') - W(m|n, t)P(n, t|n', t')]. \tag{3.5.5}$$

There is no longer any question that only jumps can occur, since only discrete values of the state variable $N(t)$ are allowed. It is most important, however, to be aware that a pure jump process can occur even though the variable $X(t)$ can take on a continuous range of variables.

3.5.2 Diffusion Processes—the Fokker-Planck Equation

If we assume the quantities $W(z|x, t)$ to be zero, the differential Chapman-Kolmogorov equation reduces to the *Fokker-Planck equation:*

$$\frac{\partial p(z, t|y, t')}{\partial t} = -\sum_i \frac{\partial}{\partial z_i}[A_i(z, t)p(z, t|y, t')] + \frac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial z_i \partial z_j}[B_{ij}(z, t)p\,(z, t|y, t')] \tag{3.5.6}$$

and the corresponding process is known mathematically as a *diffusion process.* The vector $A(z, t)$ is known as the drift vector and the matrix $\underline{B}(z, t)$ as the diffusion matrix. The diffusion matrix is positive semidefinite and symmetric as a result of its

definition in (3.4.3). It is easy to see that from the definition of $W(\boldsymbol{x}|\boldsymbol{z}, t)$ (3.4.1), the requirement (3.3.1) for continuity of the sample paths is satisfied if $W(\boldsymbol{x}|\boldsymbol{z}, t)$ is zero. Hence, the Fokker-Planck equation describes a process in which $\boldsymbol{X}(t)$ has continuous sample paths. In fact, we can heuristically give a much more definite description of the process. Let us consider computing $p(\boldsymbol{z}, t + \Delta t|\boldsymbol{y}, t)$, given that

$$p(\boldsymbol{z}, t|\boldsymbol{y}, t) = \delta(\boldsymbol{z} - \boldsymbol{y}) . \tag{3.5.7}$$

For a small Δt, the solution of the Fokker-Planck equation will still be on the whole sharply peaked, and hence derivatives of $A_i(\boldsymbol{z}, t)$ and $B_{ij}(\boldsymbol{z}, t)$ will be negligible compared to those of p. We are thus reduced to solving, approximately

$$\frac{\partial p(\boldsymbol{z}, t|\boldsymbol{y}, t')}{\partial t} = -\sum_i A_i(\boldsymbol{y}, t)\frac{\partial p(\boldsymbol{z}, t|\boldsymbol{y}, t')}{\partial z_i} + \sum_{i,j}\frac{1}{2} B_{ij}(\boldsymbol{y}, t)\frac{\partial^2 p(\boldsymbol{z}, t|\boldsymbol{y}, t')}{\partial z_i \partial z_j} , \tag{3.5.8}$$

where we have also neglected the time dependence of A_i and B_{ij} for small $t - t'$. Equation (3.5.8) can now be solved, subject to the initial condition (3.5.7), and we get

$$p(\boldsymbol{z}, t + \Delta t|\boldsymbol{y}, t) = (2\pi)^{-N/2}\{\det[\underset{\sim}{\boldsymbol{B}}(\boldsymbol{y}, t)]\}^{1/2}[\Delta t]^{-1/2}$$
$$\times \exp\left\{-\frac{1}{2}\frac{[\boldsymbol{z} - \boldsymbol{y} - \boldsymbol{A}(\boldsymbol{y}, t)\Delta t]^{\mathrm{T}}[\underset{\sim}{\boldsymbol{B}}(\boldsymbol{y}, t)]^{-1}[\boldsymbol{z} - \boldsymbol{y} - \boldsymbol{A}(\boldsymbol{y}, t)\Delta t]}{\Delta t}\right\}, \tag{3.5.9}$$

that is, a Gaussian distribution with variance matrix $\underset{\sim}{\boldsymbol{B}}(\boldsymbol{y}, t)$ and mean $\boldsymbol{y} + \boldsymbol{A}(\boldsymbol{y}, t)\Delta t$. We get the picture of the system moving with a systematic drift, whose velocity is $\boldsymbol{A}(\boldsymbol{y}, t)$, on which is superimposed a Gaussian fluctuation with covariance matrix $\underset{\sim}{\boldsymbol{B}}(\boldsymbol{y}, t)\Delta t$, that is, we can write

$$\boldsymbol{y}(t + \Delta t) = \boldsymbol{y}(t) + \boldsymbol{A}(\boldsymbol{y}(t), t)\Delta t + \boldsymbol{\eta}(t)\Delta t^{1/2} , \tag{3.5.10}$$

where $\langle \boldsymbol{\eta}(t) \rangle = 0$ (3.5.11)

$$\langle \boldsymbol{\eta}(t)\boldsymbol{\eta}(t)^{\mathrm{T}} \rangle = \underset{\sim}{\boldsymbol{B}}(\boldsymbol{y}, t) . \tag{3.5.12}$$

It is easy to see that this picture gives

i) sample paths which are always continuous — for, clearly, as $\Delta t \to 0$, $\boldsymbol{y}(t + \Delta t) \to \boldsymbol{y}(t)$;

ii) sample paths which are nowhere diffierentiable, because of the $\Delta t^{1/2}$ occurring in (3.5.10).

We shall see later, in Chap. 4 that the heuristic picture of (3.5.10) can be made much more precise and leads to the concept of the *stochastic differential equation.*

3.5.3 Deterministic Processes—Liouville's Equation

It is possible that in the differential Chapman-Kolmogorov equation (3.4.22) only the first term is nonzero, so we are led to the special case of a *Liouville equation:*

$$\frac{\partial p(z, t \mid y, t')}{\partial t} = -\sum_i \frac{\partial}{\partial z_i} [A_i(z, t) p(z, t \mid y, t')] \tag{3.5.13}$$

which occurs in classical mechanics. This equation describes a completely deterministic motion, i.e., if $x(y, t)$ is the solution of the ordinary differential equation

$$\frac{dx(t)}{dt} = A[x(t), t] \tag{3.5.14}$$

with $x(y, t') = y$, (3.5.15)

then the solution to (3.5.13) with initial condition

$$p(z, t' \mid y, t') = \delta(z - y) \tag{3.5.16}$$

is

$$p(z, t \mid y, t') = \delta[z - x(y, t)] . \tag{3.5.17}$$

The proof of this assertion is best obtained by direct substituion. For

$$-\sum_i \frac{\partial}{\partial z_i} \{A_i(z, t) \delta[z - x(y, t)]\} \tag{3.5.18}$$

$$= -\sum_i \frac{\partial}{\partial z_i} \{A_i[x(y, t), t] \delta[z - x(y, t)]\} \tag{3.5.19}$$

$$= -\sum_i \left[A_i[x(y, t), t] \frac{\partial}{\partial z_i} \delta[z - x(y, t)]\right] \tag{3.5.20}$$

and

$$\frac{\partial}{\partial t} \delta[z - x(y, t)] = -\sum_i \frac{\partial}{\partial z_i} \delta[z - x(y, t)] \frac{dx_i(y, t)}{dt} \tag{3.5.21}$$

and by use of (3.5.14), we see that (3.5.20,21) are equal. Thus, if the particle is in a well-defined initial position y at time t', it stays on the trajectory obtained by solving the ordinary differential equation (3.5.14).

Hence, deterministic motion, as defined by a first-order differential equation of the form (3.5.14), is an elementary form of Markov process. The solution (3.5.17) is, of course, merely a special case of the kind of process approximated by equations like (3.5.9) in which the Gaussian part is zero.

3.5.4 General Processes

In general, none of the quantities in $A(z, t)$, $B(z, t)$ and $W(x \mid z, t)$ need vanish, and in this case we obtain a process whose sample paths are as illustrated in Fig. 3.2, i.e., a piecewise continuous path made up of pieces which correspond to a diffusion process with a nonzero drift, onto which is superimposed a fluctuating part.



Fig. 3.2. Illustration of a sample path of a general Markov process, in which drift, diffusion and jumps exist

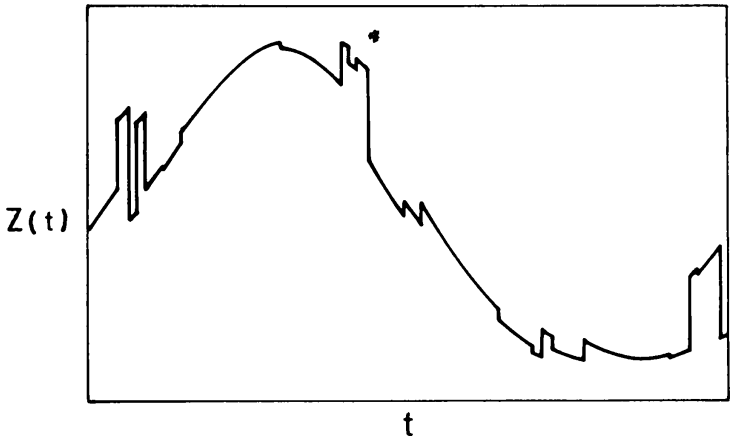


Fig. 3.3. Sample path of a Markov process with only drift and jumps

It is also possible that $A(z, t)$ is nonzero, but $B(z, t)$ is zero and here the sample paths are, as in Fig. 3.3, composed of pieces of smooth curve [solutions of (3.5.14)] with discontinuities superimposed. This is very like the picture one would expect in a dilute gas where the particles move freely between collisions which cause an instantaneous change in momentum, though not position.

3.6 Equations for Time Development in Initial Time—Backward Equations

We can derive much more simply than in Sect. 3.4, some equations which give the time development with respect to the initial variables y, t' of $p(x, t|y, t')$.

We consider

$$\lim_{\Delta t' \to 0} \frac{1}{\Delta t'} [p(x, t|y, t' + \Delta t') - p(x, t|y, t')] \tag{3.6.1}$$

$$= \lim_{\Delta t' \to 0} \frac{1}{\Delta t'} \int dz\, p(z, t' + \Delta t'|y, t')[p(x, t|y, t' + \Delta t') - p(x, t|z, t' + \Delta t')] \tag{3.6.2}$$

by use of the Chapman-Kolmogrov equation in the second term and by noting that the first term gives $1 \times p(x, t|y, t' + \Delta t')$.

The assumptions that are necessary are now the existence of all relevant deriva-

tives, and that $p(x, t|y, t')$ is continuous and bounded in x,t,t' for some range $t - t' > \delta > 0$. We may then write

$$= \lim_{\Delta t' \to 0} \frac{1}{\Delta t'} \int dz\, p(z, t' + \Delta t'|y, t')\,[p(x, t|y, t') - p(x, t|z, t')] \tag{3.6.3}$$

We now proceed using similar techniques to those used in Sect. 3.4.1 and finally derive

$$\frac{\partial p(x, t|y, t')}{\partial t'} = -\sum_i A_i(y, t') \frac{\partial p(x, t|y, t')}{\partial y_i} - \frac{1}{2} \sum_{ij} B_{ij}(y, t') \frac{\partial^2 p(x, t|y, t')}{\partial y_i \partial y_j}$$
$$+ \int dz\, W(z|y, t')\,[p(x, t|y, t') - p(x, t|z, t')] \tag{3.6.4}$$

which will be called the backward differential Chapman-Kolmogorov equation. In a mathematical sense, it is better defined than the corresponding forward equation (3.4.22). The appropriate initial condition for both equation is

$$p(x, t|y, t) = \delta(x - y) \text{ for all } t, \tag{3.6.5}$$

representing the obvious fact that if the particle is at y at time t, the probability density for finding it at x at the same time is $\delta(x-y)$.

The forward and the backward equations are equivalent to each other. For, solutions of the forward equation, subject to the initial condition (3.6.5) [or 3.5.4], and any appropriate boundary conditions, yield solutions of the Chapman-Kolmogorov equation, as noted in Sect. 3.4.2. But these have just been shown to yield the backward equation. (The relation between appropriate boundary conditions for the Fokker-Planck equations is dealt with in Sect. 5.2.1,4). The basic difference is which set of variables is held fixed. In the case of the forward equation, we hold (y, t') fixed, and solutions exist for $t \geqslant t'$, so that (3.6.5) is an *initial condition* for the forward equation. For the backward equation, solutions exist for $t' \leqslant t$, so that since the backward equation expresses development in t', (3.6.5) is really better termed *final condition* in this case.

Since they are equivalent, the forward and backward equations are both useful. The forward equation gives more directly the values of measurable quantities as a function of the observed time, t, and tends to be used more commonly in applications. The backward equation finds most application in the study of *first passage time* or *exit problems*, in which we find the probability that a particle leaves a region in a given time.

3.7 Stationary and Homogeneous Markov Processes

In Sect. 1.4.3 we met the concept of a stationary process, which represents the stochastic motion of a system which has settled down to a steady state, and whose stochastic properties are independent of when they are measured. Stationarity can be defined in various degrees, but we shall reserve the term "*stationary process*"

for a strict definition, namely, a stochastic process $X(t)$ is stationary if $X(t)$ and the process $X(t+\varepsilon)$ have the same statistics for any ε. This is equivalent to saying that all joint probability densities satisfy time translation invariance, i.e.,

$$p(x_1, t_1; x_2, t_2; x_3, t_3; \ldots; x_n, t_n) = p(x_1, t_1 + \varepsilon; x_2, t_2 + \varepsilon; x_3, t_3 + \varepsilon; \ldots; x_n, t_n + \varepsilon) \tag{3.7.1}$$

and hence such probabilities are only functions of the time differences, $t_i - t_j$. In particular, the one-time probability is independent of time and can be simply written as

$$p_s(x) \tag{3.7.2}$$

and the two-time joint probability as

$$p_s(x_1, t_1 - t_2; x_2, 0). \tag{3.7.3}$$

Finally, the conditional probability can also be written as

$$p_s(x_1, t_1 - t_2 | x_2, 0). \tag{3.7.4}$$

For a *Markov process*, since all joint probabilities can be written as products of the two-time conditional probability and the one-time probability, a necessary and sufficient condition for stationarity is the ability to write the one and two-time probabilities in the forms given in (3.7.1–3).

3.7.1 Ergodic Properties

If we have a stationary process, it is reasonable to expect that average measurements could be constructed by taking values of the variable x at successive times, and averaging various functions of these. This is effectively a belief that the law of large numbers (as explained in Sect. 2.5.2) applies to the variables defined by successive measurements in a stochastic process.

Let us define the variable $\bar{X}(T)$ by

$$\bar{X}(T) = \frac{1}{2T} \int_{-T}^{T} dt\, x(t), \tag{3.7.5}$$

where $x(t)$ is a stationary process, and consider the limit $T \to \infty$. This represents a possible model of measurement of the mean by averaging over all times. Clearly

$$\langle \bar{X}(t) \rangle = \langle x \rangle_s . \tag{3.7.6}$$

We now calculate the variance of $\bar{X}(T)$. Thus,

$$\langle \bar{X}(T)^2 \rangle = \frac{1}{4T^2} \int_{-T}^{T} \int_{-T}^{T} dt_1 dt_2 \, \langle x(t_1) x(t_2) \rangle \tag{3.7.7}$$

and if the process is stationary,

$$\langle x(t_1)\, x(t_2)\rangle \equiv R(t_1 - t_2) + \langle x\rangle^2 , \tag{3.7.8}$$

where R is the two-time correlation function. Hence,

$$\langle \bar{X}(T)^2\rangle - \langle x\rangle^2 = \frac{1}{4T^2} \int_{-2T}^{2T} d\tau\; R(\tau)(2T - |\tau|) \tag{3.7.8}$$

where the last factor follows by changing variables to

$$\tau = t_1 - t_2 \tag{3.7.9}$$
$$t = t_1$$

and integrating t.

The left-hand side is now the variance of $\bar{X}(T)$ and one can show that under certain conditions, this vanishes as $T \to \infty$. Most straightforwardly, all we require is that

$$\lim_{T\to\infty} \frac{1}{T} \int_{-2T}^{2T} d\tau \left(1 - \frac{|\tau|}{2T}\right) R(\tau) = 0 \tag{3.7.10}$$

which is a little obscure. However, it is clear that a sufficient condition for this limit to be zero is for

$$\int_0^\infty d\tau\; |R(\tau)| < \infty , \tag{3.7.11}$$

in which case, we simply require that the correlation function $\langle x(t_1), x(t_2)\rangle$ should tend to zero sufficiently rapidly as $|t_1 - t_2| \to \infty$. In cases of interest it is frequently found that the asymptotic behavior of $R(\tau)$ is

$$R(\tau) \sim \mathrm{Re}\{A \exp(-\tau/\tau_c)\} , \tag{3.7.12}$$

where τ_c is a (possibly complex) parameter known as the *correlation time.* Clearly the criterion of (3.7.11) is satisfied, and we find in this case that the variance in $\bar{X}(T)$ approaches zero so that using (3.7.6) and (2.9.4), we may write

$$\underset{T\to\infty}{\text{ms-lim}}\; \bar{X}(T) = \langle x\rangle_s . \tag{3.7.13}$$

This means that the averaging procedure (3.7.5) is indeed valid. It is not difficult to extend the result to an average of an infinite set of measurements at discrete times $t_n = t_0 + n\Delta t$.

Other ergodic hypotheses can easily be stated, and the two quantities that are of most interest are the autocorrelation function and the distribution function. As already mentioned in Sect. 1.4.2, the most natural way of measuring an autocorrelation function is through the definition

$$G(\tau, T) = \frac{1}{T}\int_0^T dt\, x(t)x(t+\tau) \tag{3.7.14}$$

and we can rather easily carry through similar reasoning to show that

$$\underset{T\to\infty}{\text{ms-lim}}\, G(\tau, T) = \langle x(t)x(t+\tau)\rangle_s\,, \tag{3.7.15}$$

provided the following condition is satisfied. Namely, define $\rho(\tau, \lambda)$ by

$$\langle x(t+\lambda+\tau)x(t+\lambda)x(t+\tau)x(t)\rangle_s = \rho(\tau, \lambda) + \langle x(t+\tau)x(t)\rangle_s^2\,. \tag{3.7.16}$$

Then we require

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-2T}^{2T}\left(1-\frac{|\lambda|}{2T}\right)\rho(\tau, \lambda)d\lambda = 0\,. \tag{3.7.17}$$

We can see that this means that for sufficiently large λ, the four-time average (3.7.16) factorises into a product of two-time averages, and that the "error term" $\rho(\tau, \lambda)$ must vanish sufficiently rapidly for $\lambda \to \infty$. Exponential behaviour, such as given in (3.7.12) is sufficient, and usually found.

We similarly find that the spectrum, given by the Fourier transform

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau}G(\tau)d\tau \tag{3.7.18}$$

as in Sect. 1.4, is also given by the procedure

$$S(\omega) = \lim_{T\to\infty}\frac{1}{2\pi T}\left|\int_0^T dt\, e^{-i\omega t}x(t)\right|^2\,. \tag{3.7.19}$$

Finally, the practical method of measuring the distribution function is to consider an interval (x_1, x_2) and measure $x(t)$ repeatedly to determine whether it is in this range or not. This gives a measure of $\int_{x_1}^{x_2} dx\, p_s(x)$. Essentially, we are then measuring the time average value of the function $\chi(x)$ defined by

$$\begin{aligned}\chi(x) &= 1 \qquad x_1 < x < x_2 \\ &= 0 \qquad \text{otherwise,}\end{aligned} \tag{3.7.20}$$

and we adapt the method of proving the ergodicity of $\langle x\rangle$ to find that the distribution is ergodic provided

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-2T}^{2T}d\tau\left(1-\frac{|\tau|}{2T}\right)\int_{x_1}^{x_2}dx' p_s(x')\left\{\int_{x_1}^{x_2}dx\,[p(x, \tau\,|\,x', 0) - p_s(x)]\right\} = 0\,. \tag{3.7.21}$$

The most obvious sufficient condition here is that

$$\lim_{\tau \to \infty} p(x, \tau \mid x', 0) = p_s(x) \tag{3.7.22}$$

and that this limit is approached sufficiently rapidly. In practice, an exponential approach is frequently found and this is, as in the case of the mean, quite sufficiently rapid.

This condition is, in fact, sufficient for ergodicity of the mean and autocorrelation function for a Markov process, since all means can be expressed in terms of conditional probabilities and the sufficiently rapid achievement of the limit (3.7.22) can be readily seen to be sufficient to guarantee both (3.7.17) and (3.7.10). We will call a Markov process simply "*ergodic*" if this rather strong condition is satisfied.

3.7.2 Homogeneous Processes

If the condition (3.7.22) is satisfied for a stationary Markov process, then we clearly have a way of constructing from the stationary Markov process a nonstationary process whose limit as time becomes large is the stationary process. We simply define the process for

$$t, t' > t_0 \tag{3.7.23}$$

by

$$p(\boldsymbol{x}, t) = p_s(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0) \qquad \text{and} \tag{3.7.24}$$

$$p(\boldsymbol{x}, t \mid \boldsymbol{x}', t') = p_s(\boldsymbol{x}, t \mid \boldsymbol{x}', t') \tag{3.7.25}$$

and all other joint probabilities are obtained from these in the usual manner for a Markov process. Clearly, if (3.7.22) is satisfied, we find that as $t \to \infty$ or as $t_0 \to -\infty$,

$$p(\boldsymbol{x}, t) \to p_s(\boldsymbol{x})$$

and all other probabilities become stationary because the conditional probability is stationary. Such a process is known as a *homogeneous* process.

The physical interpretation is rather obvious. We have a stochastic system whose variable $\boldsymbol{x}$ is by some external agency fixed to have a value $\boldsymbol{x}_0$ at time t_0. It then evolves back to a stationary system with the passage of time. This is how many stationary systems are created in practice.

From the point of view of the differential Chapman-Kolmogorov equation, we will find that the stationary distribution function $p_s(x)$ is a solution of the stationary differential Chapman-Kolmogorov equation, which takes the form

$$\begin{aligned} 0 = &-\sum_i \frac{\partial}{\partial z_i}[A_i(\boldsymbol{z})p(\boldsymbol{z}, t \mid \boldsymbol{y}, t')] + \frac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial z_i \partial z_j}[B_{ij}(\boldsymbol{z})p(\boldsymbol{z}, t \mid \boldsymbol{y}, t')] \\ &+ \int d\boldsymbol{x}\,[W(\boldsymbol{z} \mid \boldsymbol{x})p(\boldsymbol{x}, t \mid \boldsymbol{y}, t') - W(\boldsymbol{x} \mid \boldsymbol{z})p(\boldsymbol{z}, t \mid \boldsymbol{y}, t')], \end{aligned} \tag{3.7.26}$$

where we have used the fact that the process is homogeneous to note that $\boldsymbol{A}$, $\underline{\boldsymbol{B}}$ and W, as defined in (3.4.1–3), are independent of t. This is an alternative definition of a homogeneous process.

3.7.3 Approach to a Stationary Process

A converse problem also exists. Suppose A, $\underset{\sim}{B}$ and W are independent of time and $p_s(z)$ satisfies (3.7.26). Under what conditions does a solution of the differential Chapman-Kolmogorov equation approach the stationary solution $p_s(z)$?

There does not appear to be a complete answer to this problem. However, we can give a reasonably good picture as follows. We define a Lyapunov functional K of any two solutions p_1 and p_2 of the differential Chapman-Kolgomorov equation by

$$K = \int dx\, p_1(x, t) \log [p_1(x, t)/p_2(x, t)] \tag{3.7.27}$$

and assume for the moment that neither p_1 nor p_2 are zero anywhere. We will now show that K is always positive and dK/dt is always negative.

Firstly, noting that both $p_2(x, t)$ and $p_1(x, t)$ are normalised to one, we write

$$\begin{aligned} K[p_1, p_2, t] = \int dx\, p_1(x, t)\, \{&\log [p_1(x, t)/p_2(x, t)] \\ &+ p_2(x, t)/p_1(x, t) - 1\} \end{aligned} \tag{3.7.28}$$

and use the inequality valid for all $z > 0$,

$$-\log z + z - 1 \geqslant 0\,, \tag{3.7.29}$$

to show that $K \geqslant 0$.

Let us now show that $dK/dt \leqslant 0$. We can write (using an abbreviated notation)

$$\frac{dK}{dt} = \int dx \left\{ \frac{\partial p_1}{\partial t} [\log p_1 + 1 - \log p_2] - \frac{\partial p_2}{\partial t} [p_1/p_2] \right\}. \tag{3.7.30}$$

We now calculate one by one the contribution to dK/dt from drift, diffusion, and jump terms in the differential Chapman-Kolmogorov equation:

$$\begin{aligned} \left(\frac{dK}{dt}\right)_{\text{drift}} = \sum_i \int dx \Big\{ &- [\log (p_1/p_2) + 1] \frac{\partial}{\partial x_i} (A_i p_1) \\ &+ (p_1/p_2) \frac{\partial}{\partial x_i} (A_i p_2) \Big\} \end{aligned} \tag{3.7.31}$$

which can be rearranged to give

$$\left(\frac{dK}{dt}\right)_{\text{drift}} = \sum_i \int dx \frac{\partial}{\partial x_i} [- A_i p_1 \log (p_1/p_2)]. \tag{3.7.32}$$

Similarly, we may calculate

$$\begin{aligned} \left(\frac{dK}{dt}\right)_{\text{diff}} = -\frac{1}{2} \sum_{i,j} \int dx \Big\{ &[\log (p_1/p_2) + 1] \frac{\partial^2}{\partial x_i \partial x_j} (B_{ij} p_1) \\ &- (p_1/p_2) \frac{\partial^2}{\partial x_i \partial x_j} (B_{ij} p_2) \Big\} \end{aligned} \tag{3.7.33}$$

and after some rearranging we may write

$$\left(\frac{dK}{dt}\right)_{\text{diff}} = -\frac{1}{2}\sum_{i,j}\int dx\, p_1 B_{ij}\left\{\frac{\partial}{\partial x_i}[\log(p_1/p_2)]\right\}\left\{\frac{\partial}{\partial x_j}[\log(p_1/p_2)]\right\}$$
$$+\frac{1}{2}\sum_{i,j}\int dx\,\frac{\partial^2}{\partial x_i \partial x_j}\left[p_1 B_{ij}\log(p_1/p_2)\right]. \tag{3.7.34}$$

Finally, we may calculate the jump contribution similarly:

$$\left(\frac{dK}{dt}\right)_{\text{jump}} = \int dx dx'\Big\{[W(x\,|\,x')p_1(x',t) - W(x'\,|\,x)p_1(x,t)]$$
$$\times\{\log[p_1(x,t)/p_2(x,t)]+1\}$$
$$-[W(x\,|\,x')p_2(x',t) - W(x'\,|\,x)p_2(x,t)]p_1(x,t)/p_2(x,t)\Big\} \tag{3.7.35}$$

and after some rearrangement,

$$\left(\frac{dK}{dt}\right)_{\text{jump}} = \int dx dx'\, W(x\,|\,x')\{p_2(x',t)[\phi'\log[\phi/\phi'] - \phi + \phi']\}, \tag{3.7.36}$$

where

$$\phi = p_1(x,t)/p_2(x,t) \tag{3.7.37}$$

and ϕ' is similarly defined in terms of x'.

We now consider the simplest case. Suppose a stationary solution $p_s(x)$ exists which is nonzero everywhere, except at infinity, where it and its first derivative vanish. Then we may choose $p_2(x,t) = p_s(x)$. The contributions to dK/dt from (3.7.32) and the second term in (3.7.34) can be integrated to give surface terms which vanish at infinity so that we find

$$\left(\frac{dK}{dt}\right)_{\text{drift}} = 0 \tag{3.7.38a}$$

$$\left(\frac{dK}{dt}\right)_{\text{diff}} \leqslant 0 \tag{3.7.38b}$$

$$\left(\frac{dK}{dt}\right)_{\text{jump}} \leqslant 0, \tag{3.7.38c}$$

where the last inequality comes by setting $z = \phi'/\phi'$ in (3.7.29).

We must now consider under what situations the *equalities* in (3.7.38) are actually achieved. Inspection of (3.7.36) shows that this term will be zero if and only if $\phi = \phi'$ for almost all x and x' which are such that $W(x\,|\,x') \neq 0$. Thus, if $W(x\,|\,x')$ is never zero, i.e., if transitions can take place in *both directions* between any pair of states, the vanishing of the jump contribution implies that $\phi(x)$ is independent of x, i.e.,

$$p_1(\boldsymbol{x}, t)/p_s(\boldsymbol{x}) = \text{constant}. \tag{3.7.39}$$

The constant must equal one since $p_1(\boldsymbol{x}, t)$ and $p_s(\boldsymbol{x})$ are both normalised.

The term arising from diffusion will be strictly negative if B_{ij} is almost everywhere positive definite. Hence, we have now shown that under rather strong conditions, namely,

$$\begin{aligned} &p_s(\boldsymbol{x}) \neq 0 \text{ with probability 1} \\ &W(\boldsymbol{x}\,|\,\boldsymbol{x}') \neq 0 \text{ with probability 1,} \\ &B_{ij}(\boldsymbol{x}) \text{ positive definite with probability 1,} \end{aligned} \tag{3.7.40}$$

that any solution of the differential Chapman-Kolmogorov equation approaches the stationary solution $p_s(\boldsymbol{x})$ at $t \to \infty$.

The result fails in two basic kinds of systems.

a) Disconnected State Space

The result is best illustrated when A_i and B_{ij} vanish, so we have a pure jump system. Suppose the space divides into two regions R_1 and R_2 such that transitions from R_1 to R_2 and back are impossible; hence, $W(\boldsymbol{x}\,|\,\boldsymbol{x}') = 0$ if $\boldsymbol{x}$ and $\boldsymbol{x}'$ are not both in R_1 or R_2. Then it is possible to have $dK/dt = 0$ if

$$\begin{aligned} p_1(\boldsymbol{x}, t) &= \lambda_1 p_s(\boldsymbol{x}) \qquad x \in R_1 \\ &= \lambda_2 p_s(\boldsymbol{x}) \qquad x \in R_2 \end{aligned} \tag{3.7.41}$$

so that there is no unique stationary distribution. The two regions are disconnected and separate stochastic processes take place in each, and in each of these, there is a unique stationary solution. The relative probability of being R_1 or R_2 is not changed by the process.

A similar result holds, in general, if as well we have B_{ij} and A_i vanishing on the boundary between R_1 and R_2.

b) $p_s(\boldsymbol{x})$ Vanishes in Some Definite Region

If we have

$$\begin{aligned} p_s(\boldsymbol{x}) &= 0 \qquad \boldsymbol{x} \in R_1 \\ &\neq 0 \qquad \boldsymbol{x} \in R_2 \end{aligned} \tag{3.7.42}$$

and again A_i and B_{ij} vanish, then it follows that, since $p_s(\boldsymbol{x})$ satisfies the stationary equation (3.7.26),

$$W(\boldsymbol{x}\,|\,\boldsymbol{y}) = 0 \qquad \boldsymbol{x} \in R_1, \boldsymbol{y} \in R_2. \tag{3.7.43}$$

In other words, no transitions are possible from the region R_2 where the stationary distribution is positive to R_1, where the stationary distribution vanishes.

3.7.4 Autocorrelation Function for Markov Processes

For any Markov process, we can write a very elegant formula for the autocorrelation function. We define

$$\langle \boldsymbol{X}(t) | [\boldsymbol{x}_0, t_0] \rangle = \int d\boldsymbol{x}\, \boldsymbol{x}\, p(\boldsymbol{x}, t | \boldsymbol{x}_0, t_0)\,, \tag{3.7.44}$$

then the autocorrelation matrix

$$\langle \boldsymbol{X}(t)\boldsymbol{X}(t_0)^{\mathrm{T}} \rangle = \int d\boldsymbol{x}\, d\boldsymbol{x}_0\, \boldsymbol{x}\boldsymbol{x}_0^{\mathrm{T}} p(\boldsymbol{x}, t; \boldsymbol{x}_0, t_0) \tag{3.7.45}$$

$$= \int d\boldsymbol{x}_0 \langle \boldsymbol{X}(t) | [\boldsymbol{x}_0, t_0] \rangle \boldsymbol{x}_0^{\mathrm{T}} p(\boldsymbol{x}_0, t_0)\,. \tag{3.7.46}$$

Thus we see that (3.7.44) defines the mean of $\boldsymbol{X}(t)$ under the condition that $\boldsymbol{X}$ had the value $\boldsymbol{x}_0$ at time t_0, and (3.7.46) tells us that the autocorrelation matrix is obtained by averaging this conditional average (multiplied by $\boldsymbol{x}_0^{\mathrm{T}}$) at time t_0. These results are true by definition for any stochastic process.

In a Markov process we have, however, a unique conditional probability which determines the whole process. Thus, for a Markov process, we can state that $\langle \boldsymbol{X}(t) | [\boldsymbol{x}_0, t_0] \rangle$ is a uniquely defined quantity, since the knowledge of $\boldsymbol{x}_0$ at time t_0 completely determines the future of the process. The most notable use of this property is in the computation of the stationary autocorrelation function. To illustrate how this uniqueness is important, let us consider a non-Markov stationary process with joint probabilities

$$p_s(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2; \ldots \boldsymbol{x}_n, t_n), \tag{3.7.47}$$

which, of course, depend only on time differences. Let us now create a corresponding nonstationary process by selecting only sample paths which pass through the point $x = a$ at time $t = 0$. Thus, we define

$$p_a(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2; \ldots \boldsymbol{x}_n, t_n) = p_s(\boldsymbol{x}_1, t_1, \boldsymbol{x}_2, t_2; \ldots \boldsymbol{x}_n, t_n | \boldsymbol{a}, 0). \tag{3.7.48}$$

Then for this process we note that

$$\langle \boldsymbol{X}(t) | [\boldsymbol{x}_0, t_0] \rangle_a = \int d\boldsymbol{x}\, \boldsymbol{x}\, p_s(\boldsymbol{x}, t | \boldsymbol{x}_0, t_0; \boldsymbol{a}, 0) \tag{3.7.49}$$

which contains a dependence on $\boldsymbol{a}$ symbolised by the subscript $\boldsymbol{a}$ on the average bracket. If the original stationary process possesses appropriate ergodic properties, then

$$\lim_{\tau \to \infty} p_s(\boldsymbol{x}, t + \tau | \boldsymbol{x}_0, t_0 + \tau; \boldsymbol{a}, 0) = p_s(\boldsymbol{x}, t - t_0 | \boldsymbol{x}_0, 0) \tag{3.7.50}$$

so that we will also have a stationary conditional average of $\boldsymbol{x}$

$$\langle \boldsymbol{X}(t) | [\boldsymbol{x}_0, t_0] \rangle_s = \lim_{\tau \to \infty} \langle \boldsymbol{X}(t + \tau) | [\boldsymbol{x}_0, t_0 + \tau] \rangle_a \tag{3.7.51}$$

and the stationary autocorrelation matrix is given by

$$\langle X(t)X(t_0)^{\mathrm{T}}\rangle_s = \int dx_0\, x_0^{\mathrm{T}}\langle X(t)|[x_0, t_0]\rangle_s p_s(x_0) \tag{3.7.52}$$

$$= \lim_{\tau\to\infty}\langle X(t+\tau)X(t_0+\tau)^{\mathrm{T}}\rangle_a$$

$$= \lim_{\tau\to\infty}\int dx_0\, x_0^{\mathrm{T}}\langle x(t+\tau)|[x_0, t_0+\tau]\rangle_a p_a(x_0, t_0+\tau)\,. \tag{3.7.53}$$

However, when the process is Markovian, this cumbersome limiting procedure is not necessary since

$$\text{Markov} \Longrightarrow \langle X(t)|[x_0, t_0]\rangle_s = \langle X(t)|[x_0, t_0]\rangle_a$$

$$= \langle X(t)|[x_0, t_0]\rangle\,. \tag{3.7.54}$$

Equation (3.7.46) is a *regression theorem* when applied to a Markov process and is the basis of a more powerful regression theorem for *linear systems*. By this we mean systems such that a linear equation of motion exists for the means, i.e.,

$$d\langle X(t)|[x_0, t_0]\rangle/dt = -A\langle X(t)|[x_0, t_0]\rangle \tag{3.7.55}$$

which is very often the case in systems of practical interest, either as an exact result or as an approximation. The initial conditions for (3.7.55) are clearly

$$\langle X(t_0)|[x_0, t_0]\rangle = x_0\,. \tag{3.7.56}$$

Then from (3.7.50, 59)

$$\frac{d}{dt}\langle X(t)X(t_0)^{\mathrm{T}}\rangle = -A\langle X(t)X(t_0)^{\mathrm{T}}\rangle \tag{3.7.57}$$

with initial conditions $\langle X(t_0)X(t_0)^{\mathrm{T}}\rangle$. The time correlation matrix

$$\langle X(t)X(t_0)^{\mathrm{T}}\rangle - \langle X(t)\rangle\langle X(t_0)^{\mathrm{T}}\rangle = \langle X(t), X(t_0)^{\mathrm{T}}\rangle \tag{3.7.58}$$

obviously obeys the same equation, with the initial condition given by the covariance matrix at time t_0. In a stationary system, we have the result that if $G(t)$ is the stationary time correlation function and σ the stationary covariance matrix, then

$$dG(t)/dt = -A\,G(t) \tag{3.7.59}$$

and

$$G(0) = \sigma \tag{3.7.60}$$

or

$$G(t) = \exp[-At]\sigma \tag{3.7.61}$$

which is the *regression theorem* in its simplest form. We again stress that it is valid for the *Markov processes* in which the mean values obey *linear* evolution equations like (3.7.55).

For non-Markov processes there is no simple procedure. We must carry out the complicated procedure implicit in (3.7.53).

3.8 Examples of Markov Processes

We present here for reference some fundamental solutions of certain cases of the differential Chapman-Kolmogorov equation. These will have a wide application throughout the remainder of this book.

3.8.1 The Wiener Process

This takes its name from N. Wiener who studied it extensively. From the point of view of this chapter, it is the solution of the Fokker-Planck equation as discussed in Sect.3.5.2, in which there is only one variable $W(t)$, the drift coefficient is zero and the diffusion coefficient is 1. Thus, the Fokker-Planck equation for this case is

$$\frac{\partial}{\partial t} p(w, t \mid w_0, t_0) = \frac{1}{2} \frac{\partial^2}{\partial w^2} p(w, t \mid w_0, t_0) \,. \tag{3.8.1}$$

Utilising the initial condition

$$p(w, t_0 \mid w_0, t_0) = \delta(w - w_0) \tag{3.8.2}$$

on the conditional probability, we solve (3.8.1) by use of the characteristic function

$$\phi(s, t) = \int dw \, p(w, t \mid w_0, t_0) \exp(\mathrm{i}sw) \tag{3.8.3}$$

which satisfies

$$\frac{\partial \phi}{\partial t} = -\tfrac{1}{2} s^2 \phi \tag{3.8.4}$$

so that

$$\phi(s, t) = \exp\left[-\tfrac{1}{2} s^2 (t - t_0)\right] \phi(s, t_0) \,. \tag{3.8.5}$$

From (3.8.2), the initial condition is

$$\phi(s, t_0) = \exp(\mathrm{i}sw_0)$$

so that

$$\phi(s, t) = \exp\left[\mathrm{i}sw_0 - \tfrac{1}{2} s^2 (t - t_0)\right]. \tag{3.8.6}$$

Performing the Fourier inversion, we have the solution to (3.8.1):

$$p(w, t \mid w_0, t_0) = [2\pi(t - t_0)]^{-1/2} \exp[-(w - w_0)^2/2(t - t_0)] . \tag{3.8.7}$$

This represents a Gaussian, with

$$\langle W(t) \rangle = w_0 \tag{3.8.8}$$

$$\langle [W(t) - w_0]^2 \rangle = t - t_0 , \tag{3.8.9}$$

so that an initially sharp distribution spreads in time, as graphed in Fig.3.4.

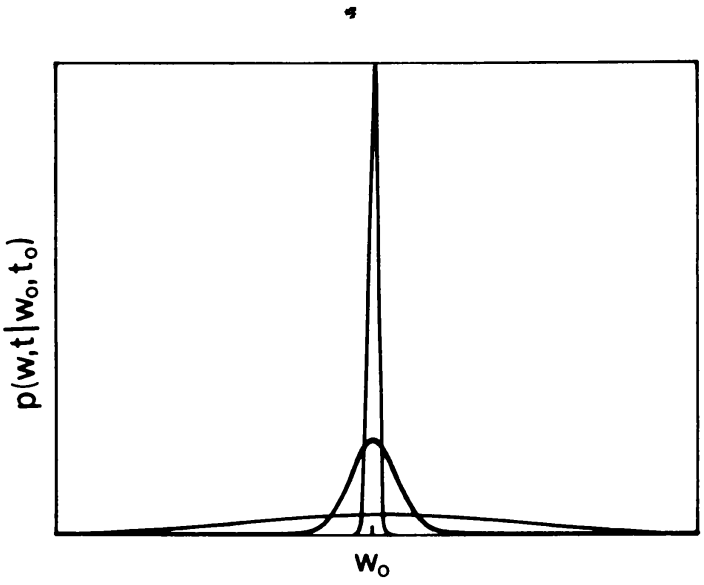


Fig. 3.4. Wiener process: spreading of an initially sharp distribution $p(w, t \mid w_0, t_0)$ with increasing time $t - t_0$

A multivariate Wiener process can be defined as

$$\boldsymbol{W}(t) = [W_1(t), W_2(t), \ldots, W_n(t)] \tag{3.8.10}$$

which satisfies the multivariable Fokker-Planck equation

$$\frac{\partial}{\partial t} p(\boldsymbol{w}, t \mid \boldsymbol{w}_0, t_0) = \frac{1}{2} \sum_i \frac{\partial^2}{\partial w_i^2} p(\boldsymbol{w}, t \mid \boldsymbol{w}_0, t_0) \tag{3.8.11}$$

whose solution is

$$p(\boldsymbol{w}, t \mid \boldsymbol{w}_0 , t_0) = [2\pi(t - t_0)]^{-n/2} \exp[-(\boldsymbol{w} - \boldsymbol{w}_0)^2/2(t - t_0)] , \tag{3.8.12}$$

a multivariate Gaussian with

$$\langle \boldsymbol{W}(t) \rangle = \boldsymbol{w}_0 \tag{3.8.13}$$

and

$$\langle [W_i(t) - w_{0i}] [W_j(t) - w_{0j}] \rangle = (t - t_0)\delta_{ij} . \tag{3.8.14}$$

The one-variable Wiener process is often simply called Brownian motion, since the Wiener process equation (3.8.1) is exactly the same as the differential equation of diffusion, shown by Einstein to be obeyed by Brownian motion, as we noted in Sect. 1.2. The terminology is, however, not universal.

Points of note concerning the Wiener process are:

a) Irregularity of Sample Paths

Although the mean value of $W(t)$ is zero, the mean square becomes infinite as $t \to \infty$. This means that the sample paths of $W(t)$ are *very* variable, indeed surprisingly so. In Fig. 3.5, we have given a few different sample paths with the same initial point to illustrate the extreme non-reproducibility of the paths.

b) Non-differentiability of Sample Paths

The Wiener process is a diffusion process and hence the sample paths of $W(t)$ are continuous. However, they are *not* differentiable. Consider

$$\text{Prob}\{|[W(t+h) - W(t)]/h| > k\} \,. \tag{3.8.15}$$

From the solution for the conditional probability, this probability is

$$2 \int_{kh}^{\infty} dw (2\pi h)^{-1/2} \exp(-w^2/2h) \tag{3.8.16}$$

and in the limit as $h \to 0$ this is one. This means that no matter what value of k choose, $|[W(t+h) - W(t)]/h|$ is almost certain to be greater than this, i.e., the derivative at any point is almost certainly infinite. This is in agreement with the similar intuitive picture presented in Sect.3.5.2 and the simulated paths given in Fig. 3.5 illustrate in point dramatically. This corresponds, of course, to the well-known experimental fact that the Brownian particles have an exceedingly irregular motion. However, this clearly an idealisation, since if $W(t)$ represents the position

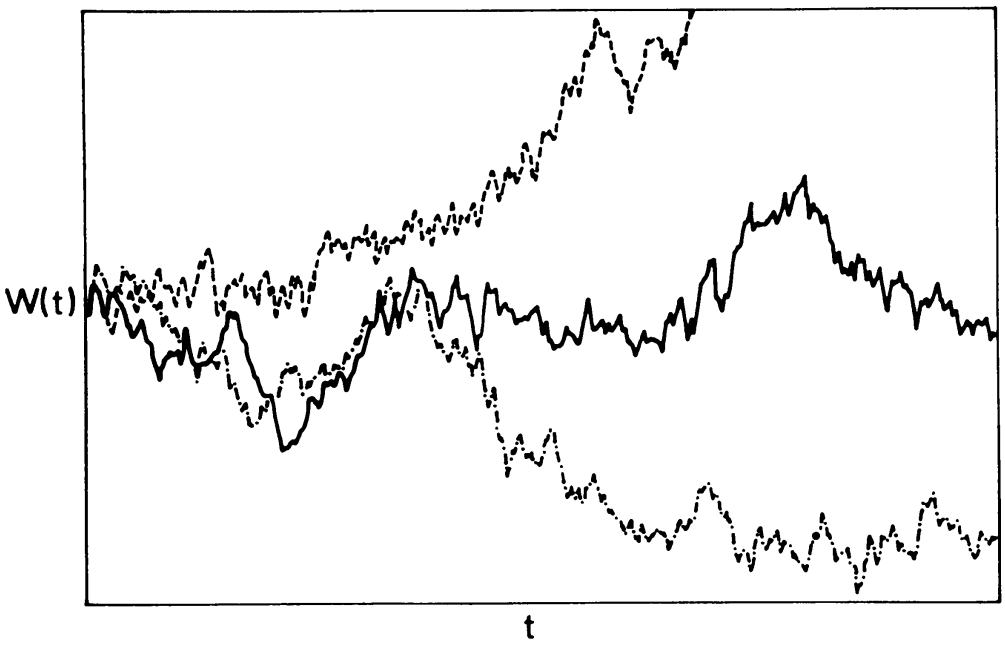


Fig. 3.5. Three simulated sample paths of the Wiener process, illustrating their great variability

of the Brownian particle, this means that its speed is almost certainly infinite. The Ornstein-Uhlenbeck process is a more realistic model of Brownian motion (Sect. 3.8.4).

c) Independence of Increment

The Wiener process is fundamental to the study of diffusion processes, and by means of stochastic differential equations, we can express any diffusion process in terms of the Wiener process.

Of particular importance is the statistical independence of the increments of $W(t)$. More precisely, since the Wiener process is a Markov process, the joint probability density can be written

$$p(w_n, t_n; w_{n-1}, t_{n-1}; w_{n-2}, t_{n-2}; \ldots ; w_0, t_0)$$
$$= \prod_{i=0}^{n-1} p(w_{i+1}, t_{i+1} \mid w_i, t_i) p(w_0, t_0) , \tag{3.8.17}$$

and using the explicit form of the conditional probabilities (3.8.7), we see that

$$p(w_n, t_n; w_{n-1}, t_{n-1}; w_{n-2}, t_{n-2}; \ldots; w_0, t_0)$$
$$= \prod_{i=0}^{n-1} \{[2\pi(t_{i+1} - t_i)]^{-1/2} \exp [-(w_{i+1} - w_i)^2/2(t_{i+1} - t_i)]\} p(w_0, t_0) . \tag{3.8.18}$$

If we define the variables

$$\Delta W_i \equiv W(t_i) - W(t_{i-1}) \tag{3.8.19}$$

$$\Delta t_i \equiv t_i - t_{i-1} , \tag{3.8.20}$$

then the joint probability density for the ΔW_i is

$$p(\Delta w_n; \Delta w_{n-1}; \Delta w_{n-2}; \ldots \Delta w_1; w_0)$$
$$= \prod_{i=1}^{n} \{(2\pi\Delta t_i)^{-1/2} \exp (-\Delta w_i^2/2\Delta t_i)\} p(w_0, t_0) \tag{3.8.21}$$

which shows from the definition of statistical independence given in Sect.2.3.4, that the variables ΔW_i are independent of each other and of $W(t_0)$.

The aspect of having independent increments ΔW_i is very important in the definition of stochastic integration which is carried out in Sect. 4.2.

d) Autocorrelation Functions

A quantity of great interest is the autocorrelation function, already discussed in Sects. 1.4.2 and 3.7.4. The formal definition is

$$\langle W(t)W(s) \mid [w_0, t_0] \rangle = \int dw_1 dw_2 \, w_1 w_2 p(w_1, t; w_2, s \mid w_0, t_0), \tag{3.8.22}$$

which is the mean product of $W(t)$ and $W(s)$ on the condition that the initial value is $W(t_0) = w_0$, and we can see, assuming $t > s$, that

$$\langle W(t)W(s)|[w_0, t_0]\rangle = \langle [W(t) - W(s)]W(s)\rangle + \langle [W(s)]^2\rangle \,. \tag{3.8.23}$$

Using the independence of increments, the first average is zero and the second is given by (3.8.9) so that we have, in general,

$$\langle W(t)W(s)|[w_0, t_0]\rangle = \min(t - t_0, s - t_0) + w_0^2 \tag{3.8.24}$$

which is correct for $t > s$ and $t < s$.

3.8.2 The Random Walk in One Dimension

This is a very famous problem, which is now considered classical. A man moves along a line, taking, at random, steps to the left or the right with equal probability. The steps are of length l so that his position can take on only the value nl, where n is integral. We want to know the probability that he reaches a given point a distance nl from the origin after a given elapsed time.

The problem can be defined in two ways. The first, which is more traditional, is to allow the walker to take steps at times $N\tau$ (N integral) at which times he *must* step either left or right, with equal probability. The second is to allow the walker to take steps left or right with a probability per unit time d which means that the walker waits at each point for a *variable* time. The second method is describable by a Master equation.

To do a Master equation treatment of the problem, we consider that the transition probability per unit time is given by the form

$$W(n + 1|n,t) = W(n - 1|n, t) = d; \tag{3.8.25}$$

otherwise, $W(n|m, t) = 0$ so that, according to Sect.3.5.1, the Master equation for the man to be at the position nl, given that he started at $n'l$, is

$$\begin{aligned}\partial_t P(n, t|n', t') = d[&P(n + 1, t|n', t') + P(n - 1, t|n', t') \\ &- 2P(n, t|n't')].\end{aligned} \tag{3.8.26}$$

The more classical form of the random walk does not assume that the man makes his jump to the left or right according to a Master equaton, but that he jumps left or right with equal probability at times $N\tau$, so that time is a discrete variable. In this case, we can write

$$\begin{aligned}P(n, (N + 1)\tau|n', N'\tau) = \tfrac{1}{2}\,&[P(n + 1, N\tau|n', N'\tau) \\ &+ P(n - 1, N\tau|n', N'\tau)]\,.\end{aligned} \tag{3.8.27}$$

If τ is small, we can view (3.8.26, 27) as approximations to each other by writing

$$P(n, (N + 1)\tau|n', N'\tau) \simeq P(n, N\tau|n', N'\tau) + \tau\partial_t P(n, t|n', t') \tag{3.8.28}$$

with $t = N\tau$, $t' = N'\tau$ and $d = \frac{1}{2}\tau^{-1}$, so that the transition probability per unit time in the Master equation model corresponds to half of the inverse waiting time τ in the discrete time model.

Both systems can be easily solved by introducing the characteristic function

$$G(s, t) = \langle e^{ins} \rangle = \sum_n P(n, t \mid n', t') e^{ins} \tag{3.8.29}$$

in which case the Master equation gives

$$\partial_t G(s, t) = d(e^{is} + e^{-is} - 2)G(s, t) \tag{3.8.30}$$

and the discrete time equation becomes

$$G(s, (N+1)\tau) = \tfrac{1}{2}(e^{is} + e^{-is})G(s, N\tau) . \tag{3.8.31}$$

Assuming the man starts at the origin $n' = 0$ at time $t' = 0$, we find

$$G(s, 0) = 1 \tag{3.8.32}$$

in both cases, so that the solution to (3.8.30) is

$$G_1(s, t) = \exp[(e^{is} + e^{-is} - 2)td] , \tag{3.8.33}$$

and to (3.8.31)

$$G_2(s, N\tau) = [\tfrac{1}{2}(e^{is} + e^{-is})]^N \tag{3.8.34}$$

which can be written

$$G_2(s, t) = \left[1 + \frac{dt}{N}(e^{is} + e^{-is} - 2)\right]^N . \tag{3.8.35}$$

Using the usual exponential limit

$$\lim_{N\to\infty} \left(1 + \frac{a}{N}\right)^N = e^a , \tag{3.8.36}$$

we see that, provided s is sufficently small

$$\lim_{N\to\infty} G_2(s, t) = G_1(s, t) \tag{3.8.37}$$

which, by the properties of the characteristic function (v) in Sect.2.6, means the probability distributions approach each other.

The appropriate probability distributions can be obtained by expanding $G_1(s, N\tau)$ and $G_2(s, t)$ in powers of $\exp(is)$; we find

$$P_1(n, t \mid 0, 0) = e^{-2td} I_n(4td) \tag{3.8.38}$$

$$P_2(n, N\tau \mid 0, 0) = (\tfrac{1}{2})^N N! \left[\left(\frac{N-n}{2} \right)! \left(\frac{N+n}{2} \right)! \right]^{-1}. \tag{3.8.39}$$

The discrete time distribution is also known as the Bernoulli distribution; it gives the probability of a total of n heads in tossing an unbiased coin N times.

The limit of continuous space is also of interest. If we set the distance travelled as

$$x = nl \tag{3.8.40}$$

so that the characteristic function of the distribution of x is

$$\phi_1(s, t) = \langle e^{isx} \rangle = G_1(ls, t) \exp[(e^{ils} + e^{-ils} - 2)td] . \tag{3.8.41}$$

then the limit of infinitesimally small steps $l \to 0$ is

$$\phi_1(s, t) \to \exp(-s^2 tD) , \tag{3.8.42}$$

where $$D = \lim_{l \to 0} (l^2 d). \tag{3.8.43}$$

This is the characteristic function of a Gaussian (Sect.2.8.1) of the form

$$p(x, t \mid 0, 0) = (4\pi Dt)^{-1/2} \exp(-x^2/4Dt) \tag{3.8.44}$$

and is of course the distribution for the Wiener process (Sect.3.8.1) or Brownian motion, as mentioned in Sect.1.2. Thus, the Wiener process can be regarded as the limit of a continuous time random walk in the limit of infinitesimally small step size.

The limit

$$l \to 0, \ \tau \to 0, \quad \text{with} \quad D = \lim_{l \to 0} (l^2/\tau) \tag{3.8.45}$$

of the discrete time random walk gives the same result. From this form, we see clearly the expression of D as the mean square distance travelled per unit time.

We can also see more directly that expanding the right-hand side of (3.8.26) as a function of x up to second order in l gives

$$\partial_t p(x, t \mid 0, 0) = (l^2 d) \partial_x^2 p(x, t \mid 0, 0) . \tag{3.8.46}$$

The three processes are thus intimately connected with each other at two levels, namely, under the limits considered, the stochastic equations approach each other and under those same limits, the solutions to these equations approach each other. These limits are exactly those used by Einstein. Comparison with Sect.1.2 shows that he modelled Brownian motion by a discrete time and space random walk, but nevertheless, derived the Wiener process model by expanding the equations for time development of the distribution function.

The limit results of this section are a slightly more rigorous version of Einstein's method. There are generalisations of these results to less specialised situations and it is a fair statement to make that almost any jump process has some kind of limit which is a diffusion process. However, the precise limits are not always so simple, and there are limits in which certain jump processes become deterministic and are governed by Liouville's equation (Sect.3.5.3) rather than the full Fokker-Planck equation. These results are presented in Sect.7.2.

3.8.3 Poisson Process

We have already noted the Poisson process in Sect.1.4.1. The process in which electrons arrive at an anode or customers arrive at a shop with a probability per unit time d of arriving, is governed by the Master equation for which

$$W(n+1\,|\,n,t) = d\,; \tag{3.8.47}$$

otherwise,

$$W(n\,|\,m,t) = 0\,. \tag{3.8.48}$$

This Master equation becomes

$$\partial_t P(n,t\,|\,n',t') = d[P(n-1,t\,|\,n',t') - P(n,t\,|\,n',t')] \tag{3.8.49}$$

and by comparison with (3.8.26) also represents a "one-sided" random walk, in which the walker steps to the right only with probability per unit time equal to d.

The characteristic function equation is similar to (3.8.30):

$$\partial_t G(s,t) = d[\exp(\mathrm{i}s) - 1]G(s,t) \tag{3.8.50}$$

with the solution

$$G(s,t) = \exp\{td[\exp(\mathrm{i}s) - 1] \tag{3.8.51}$$

for the initial condition that there are initially no customers (or electrons) at time $t = 0$, yielding

$$P(n,t\,|\,0,0) = \exp(-td)(td)^n/n!\,, \tag{3.8.52}$$

a Poisson distribution with mean given by

$$\langle N(t)\rangle = td\,. \tag{3.8.53}$$

In contrast to the random walk, the only limit that exists is $l \to 0$, with

$$dl \equiv v \tag{3.8.54}$$

held fixed and the limiting characteristic function is

$$\lim_{l\to 0}\{\exp[td(e^{ils}-1)]\} = \exp(itvs) \tag{3.8.55}$$

with the solution

$$p(x, t|0, 0) = \delta(x - vt)\,. \tag{3.8.56}$$

We also see that in this limit, we would obtain from the Master equation (3.8.49) Liouville's equation, whose solution would be the deterministic motion we have derived.

We can do a slightly more refined analysis. We expand the characteristic function up to second order in s in the exponent and find

$$\phi(s, t) = G(ls, t) \simeq \exp[t(ivs - s^2D/2)] \tag{3.8.57}$$

where, as in the previous section,

$$D = l^2d\,.$$

This is the characteristic function of a Gaussian with variance Dt and mean vt, so that we now have

$$p(x, t|0, 0) \simeq (2\pi Dt)^{-1/2}\exp[-(x - vt)^2/2Dt]\,. \tag{3.8.58}$$

It is also clear that this solution is the solution of

$$\partial_t p(x, t|0, 0) = -v\,\partial_x p(x, t|0,0) + \tfrac{1}{2}D\,\partial_x^2 p(x, t|0, 0 \tag{3.8.59}$$

which is obtained by expanding the Master equation (3.8.49) to order l^2, by writing

$$\begin{aligned} P(n - 1, t|0, 0) &= d\,p(x - l, t|0, 0) \\ &\simeq d\,p(x, t|0,0) - ld\,\partial_x p(x, t|0,0) + \tfrac{1}{2}l^2d\,\partial_x^2 p(x, t|0,0). \end{aligned} \tag{3.8.60}$$

However, this is an *approximation* or an *expansion* and not a limit. The limit $l \to 0$ gives Liouville's equation with the purely deterministic solution (3.8.56). Effectively, the limit $l \to 0$ with well-defined v corresponds to $D = 0$. The kind of approximation just mentioned is a special case of van Kampen's system size expansion which we treat fully in Sect.7.2.3.

3.8.4 The Ornstein-Uhlenbeck Process

All the examples so far have had no stationary distribution, that is, as $t \to \infty$, the distribution at any finite point approaches zero and we see that, with probability one, the point moves to infinity.

If we add a linear drift term to the Wiener process, we have a Fokker-Planck equation of the form

$$\partial_t p = \partial_x(kxp) + \tfrac{1}{2}D\,\partial_x^2 p\,, \tag{3.8.61}$$

where by p we mean $p(x, t \mid x_0, 0)$. This is the *Ornstein-Uhlenbeck process* [3.5]. The equation for the characteristic function

$$\phi(s) = \int_{-\infty}^{\infty} e^{isx} p(x, t \mid x_0, 0) dx \qquad \text{is} \tag{3.8.62}$$

$$\partial_t \phi + ks\partial_s \phi = -\tfrac{1}{2} Ds^2 \phi \,. \tag{3.8.63}$$

The method of characteristics can be used to solve this equation, namely, if

$$u(s, t, \phi) = a \qquad \text{and} \qquad v(s, t, \phi) = b \tag{3.8.64}$$

are two integrals of the subsidary equation (with a and b arbitrary const)

$$\frac{dt}{1} = \frac{ds}{ks} = -\frac{d\phi}{\frac{1}{2}Ds^2\phi}, \tag{3.8.65}$$

then a general solution of (3.8.63) is given by

$$f(u, v) = 0.$$

The particular integrals are readily found by integrating the equation involving dt and ds and that involving ds and $d\phi$; they are

$$u(s, t, \phi) = s \exp(-kt) \qquad \text{and} \tag{3.8.66}$$

$$v(s, t, \phi) = \phi \exp(Ds^2/4k), \tag{3.8.67}$$

and the general solution can clearly be put in the form $v = g(u)$ with $g(u)$ an arbitrary function of u. Thus, the general solution is

$$\phi(s, t) = \exp(-Ds^2/4k) g[s \exp(-kt)] \tag{3.8.68}$$

The boundary condition

$$p(x, 0 \mid x_0, 0) = \delta(x - x_0) \tag{3.8.69}$$

clearly requires

$$\phi(s, 0) = \exp(ix_0 s) \tag{3.8.70}$$

and gives

$$g(s) = \exp(Ds^2/4k + ix_0 s),$$

and hence,

$$\phi(s, t) = \exp\left[\frac{-Ds^2}{4k}(1 - e^{-2kt}) + isx_0 e^{-kt}\right] \tag{3.8.71}$$

which, from Sect.2.8.1, corresponds to a Gaussian with

$$\langle X(t)\rangle = x_0 \exp(-kt) \tag{3.8.72}$$

$$\text{var}\{X(t)\} = \frac{D}{2k}[1 - \exp(-2kt)]\,. \tag{3.8.73}$$

Clearly, as $t \to \infty$, the mean and variance approach limits 0 and $D/2k$, respectively, which gives a limiting stationary solution. This solution can also be obtained directly by requiring $\partial_t p = 0$, so that p satisfies the stationary Fokker-Planck equation

$$\partial_x\left[kxp + \tfrac{1}{2}D\partial_x p\right] = 0 \tag{3.8.74}$$

and integrating once, we find

$$\left[kxp + \tfrac{1}{2}D\partial_x p\right]_{-\infty}^{x} = 0. \tag{3.8.75}$$

The requirement that p vanish at $-\infty$ together with its derivative, is necessary for normalisation. Hence, we have

$$\frac{1}{p}\partial_x p = -\frac{2kx}{D} \tag{3.8.76}$$

$$\text{so that} \quad p_s(x) = (\pi D/k)^{-1/2} \exp(-kx^2/D) \tag{3.8.77}$$

which is a Gaussian with mean 0 and variance $D/2k$, as predicted from the time-dependent solution.

It is clear that a stationary solution can always be obtained for a one variable system by this integration method if such a stationary solution exists. If a stationary solution does not exist, this method gives an unnormalisable solution.

Time Correlation Functions for the Ornstein-Uhlenbeck Process. The time correlation function analogous to that mentioned in connection with the Wiener process can be calculated and is a measurable piece of data in most stochastic systems. However, we have no easy way of computing it other than by definition

$$\langle X(t)X(s)\,|\,[x_0, t_0]\rangle = \iint dx_1 dx_2\, x_1 x_2\, p(x_1, t; x_2, s\,|\,x_0, t_0) \tag{3.8.78}$$

and using the Markov property

$$= \iint dx_1 dx_2\, x_1 x_2\, p(x_1, t\,|\,x_2, s)p(x_2, s\,|\,x_0, t_0) \tag{3.8.79}$$

on the assumption that

$$t \geqslant s \geqslant t_0\,. \tag{3.8.80}$$

The correlation function with a definite initial condition is not normally of as much interest as the *stationary correlation function*, which is obtained by allowing the system to approach the stationary distribution. It is achieved by putting the initial condition in the remote past, as pointed out in Sect. 3.7.2. Letting $t_0 \to -\infty$, we find

$$\lim_{t_0 \to -\infty} p(x_2, s \mid x_0, t_0) = p_s(x_2) = (\pi D/k)^{-1/2} \exp(-kx_2^2/D) . \tag{3.8.81}$$

and by straightforward substitution and integration and noting that the stationary mean is zero, we get

$$\langle X(t)X(s)\rangle_s = \langle X(t),X(s)\rangle_s = \frac{D}{2k} \exp(-k|t-s|) . \tag{3.8.82}$$

This result demonstrates the general property of stationary processes: that the correlation functions depend only on time differences. It is also a general result [3.6] that the process we have described in this section is the only stationary Gaussian Markov process in one real variable.

The results of this subsection are very easily obtained by the stochastic differential equation methods which will be developed in Chap.4.

The Ornstein-Uhlenbeck process is a simple, explicitly representable process, which has a stationary solution. In its stationary state, it is often used to model a realistic noise signal, in which $X(t)$ and $X(s)$ are only significantly correlated if

$$|t-s| \sim 1/k \equiv \tau . \tag{3.8.83}$$

(More precisely, τ, known as the *correlation time* can be defined for arbitrary processes $X(s)$ by

$$\tau = \int_0^\infty dt \, \langle X(t), X(0)\rangle_s / \mathrm{var}\{X\}_s \tag{3.8.84}$$

which is independent of the precise functional form of the correlation function).

3.8.5 Random Telegraph Process

We consider a signal $X(t)$ which can have either of two values a and b and switches from one to the other with certain probabilities per unit time. Thus, we have a Master equation

$$\left.\begin{aligned} \partial_t P(a, t \mid x, t_0) &= -\lambda P(a, t \mid x, t_0) + \mu P(b, t \mid x, t_0) \\ \partial_t P(b, t \mid x, t_0) &= \lambda P(a, t \mid x, t_0) - \mu P(b, t \mid x, t_0) \end{aligned}\right\} \tag{3.8.85}$$

for which the solution can simply be found by noting that

$$P(a, t \mid x, t_0) + P(b, t \mid x, t_0) = 1$$

and that a simple equation can be derived for $\lambda P(a, t|x, t_0) - \mu P(b, t|x, t_0)$, whose solution is, because of the initial condition

$$P(x', t_0|x, t_0) = \delta_{x,x'}\,, \tag{3.8.86}$$

$$\lambda P(a, t|x, t_0) - \mu P(b, t|x, t_0) = \exp[-(\lambda + \mu)(t - t_0)](\lambda\delta_{a,x} - \mu\delta_{b,x}) \tag{3.8.87}$$

so that

$$\left.\begin{aligned} P(a, t|x, t_0) &= \frac{\mu}{\lambda+\mu} + \exp[-(\lambda+\mu)(t - t_0)]\left(\frac{\lambda}{\lambda+\mu}\delta_{a,x} - \frac{\mu}{\lambda+\mu}\delta_{b,x}\right) \\ P(b, t|x, t_0) &= \frac{\lambda}{\lambda+\mu} - \exp[-(\lambda+\mu)(t - t_0)]\left(\frac{\lambda}{\lambda+\mu}\delta_{a,x} - \frac{\mu}{\lambda+\mu}\delta_{b,x}\right) \end{aligned}\right\} \tag{3.8.88}$$

This process clearly has the stationary solution obtained by letting $t_0 \to -\infty$:

$$\left.\begin{aligned} P_s(a) &= \frac{\mu}{\lambda+\mu} \\ P_s(b) &= \frac{\lambda}{\lambda+\mu} \end{aligned}\right\} \tag{3.8.89}$$

which is, of course, obvious from the Master equation.

The mean of $X(t)$ and its variance are straightforwardly computed:

$$\begin{aligned} \langle X(t)|[x_0, t_0]\rangle &= \sum xP(x, t|x_0, t_0) \\ &= \frac{a\mu + b\lambda}{\mu+\lambda} + \exp[-(\lambda+\mu)(t - t_0)]\left(x_0 - \frac{a\mu + b\lambda}{\mu+\lambda}\right) \end{aligned} \tag{3.8.90}$$

so that

$$\langle X\rangle_s = \frac{a\mu + b\lambda}{\mu+\lambda}\,. \tag{3.8.91}$$

The variance can also be computed but is a very messy expression. The stationary variance is easily computed to be

$$\mathrm{var}\{X\}_s = \frac{(a-b)^2\mu\lambda}{(\lambda+\mu)^2}\,. \tag{3.8.92}$$

To compute the stationary time correlation function, we write (assuming $t \geqslant s$)

$$\langle X(t)X(s)\rangle_s = \sum_{xx'} xx'P(x, t|x', s)P_s(x') \tag{3.8.93}$$

$$= \sum_{x'} x'\langle X(t)|[x', s]\rangle P_s(x')\,. \tag{3.8.94}$$

we now use (3.8.90–92) to obtain

$$\langle X(t)X(s)\rangle_s = \langle X\rangle_s^2 + \exp[-(\lambda+\mu)(t-s)](\langle X^2\rangle_s - \langle X\rangle_s^2) \tag{3.8.95}$$

$$= \left(\frac{a\mu + b\lambda}{\mu+\lambda}\right)^2 + \exp[-(\lambda+\mu)(t-s)]\frac{(a-b)^2\mu\lambda}{(\lambda+\mu)^2}. \tag{3.8.96}$$

Hence,

$$\langle X(t), X(s)\rangle_s = \langle X(t)X(s)\rangle_s - \langle X\rangle_s^2 = \exp[-(\lambda+\mu)|t-s|]\frac{(a-b)^2\mu\lambda}{(\lambda+\mu)^2}. \tag{3.8.97}$$

Notice that this time correlation function is of exactly the same form as that of the Ornstein-Uhlenbeck process. Higher-order correlation functions are not the same of course, but because of this simple correlation function and the simplicity of the two state process, the random telegraph signal also finds wide application in model building.

4. The Ito Calculus and Stochastic Differential Equations

4.1 Motivation

In Sect.1.2.2 we met for the first time the equation which is the prototype of what is now known as a Langevin equation, which can be described heuristically as an ordinary differential equation in which a rapidly and irregularly fluctuating random function of time [the term $X(t)$ in Langevin's original equation] occurs. The simplicity of Langevin's derivation of Einstein's results is in itself sufficient motivation to attempt to put the concept of such an equation on a reasonably precise footing.

The simple-minded Langevin equation that turns up most often can be written in the form

$$\frac{dx}{dt} = a(x, t) + b(x, t)\xi(t) , \tag{4.1.1}$$

where x is the variable of interest, $a(x, t)$ and $b(x, t)$ are certain known functions and $\xi(t)$ is the rapidly fluctuating random term. An idealised mathematical formulation of the concept of a "rapidly varying, highly irregular function" is that for $t \neq t'$, $\xi(t)$ and $\xi(t')$ are statistically independent. We also require $\langle \xi(t) \rangle = 0$, since any non-zero mean can be absorbed into the definition of $a(x, t)$, and thus require that

$$\langle \xi(t)\xi(t') \rangle = \delta(t - t') \tag{4.1.2}$$

which satisfies the requirement of no correlation at different times and furthermore, has the rather pathological result that $\xi(t)$ has infinite variance. From a realistic point of view, we know that no quantity can have such an infinite variance, but the concept of *white noise* as an *idealisation* of a realistic fluctuating signal does have some meaning, and has already been mentioned in Sect.1.4.2 in connection with Johnson noise in electrical circuits. We have already met two sources which might be considered realistic versions of almost uncorrelated noise, namely, the Ornstein-Uhlenbeck process and the random telegraph signal. For both of these the second-order correlation function can, up to a constant factor, be put in the form

$$\langle X(t), X(t') \rangle = \frac{\gamma}{2} e^{-\gamma|t-t'|} . \tag{4.1.3}$$

Now the essential difference between these two is that the sample paths of the random telegraph signal are discontinuous, while those of the Ornstein-Uhlenbeck process are not. If (4.1.1) is to be regarded as a real differential equation, in which $\xi(t)$

is not white noise with a delta function correlation, but rather a noise with a finite correlation time, then the choice of a continuous function for $\xi(t)$ seems essential to make this equation realistic: we do not expect dx/dt to change discontinuously. The limit as $\gamma \to \infty$ of the correlation function (4.1.3) is clearly the Dirac delta function since

$$\int_{-\infty}^{\infty} \frac{\gamma}{2} e^{-\gamma|t-t'|} dt' = 1 \tag{4.1.4}$$

and for $t \neq t'$,

$$\lim_{\gamma \to \infty} \frac{\gamma}{2} e^{-\gamma|t-t'|} = 0 . \tag{4.1.5}$$

This means that a possible model of the $\xi(t)$ could be obtained by taking some kind of limit as $\gamma \to \infty$ of the Ornstein-Uhlenbeck process. This would correspond, in the notation of Sect. 3.8.4, to the limit

$$k \to \infty \tag{4.1.6}$$

with $D = k^2$. This limit simply does not exist. Any such limit must clearly be taken after calculating measurable quantities. Such a procedure is possible but too cumbersome to use as a calculational tool.

An alternative approach is called for. Since we write the differential equation (4.1.1), we must expect it to be integrable and hence must expect that

$$u(t) = \int_0^t dt' \, \xi(t') \tag{4.1.7}$$

exists.

Suppose we now demand the ordinary property of an integral, that $u(t)$ is a continuous function of t. This implies that $u(t)$ is a Markov process since we can write

$$u(t') = \int_0^t ds \, \xi(s) + \int_t^{t'} ds \, \xi(s) \tag{4.1.8}$$

$$= \lim_{\varepsilon \to 0} \left[\int_0^{t-\varepsilon} ds \, \xi(s) \right] + \int_t^{t'} ds \, \xi(s) \tag{4.1.9}$$

and for any $\varepsilon > 0$, the $\xi(s)$ in the first integral are independent of the $\xi(s)$ in the second integral. Hence, by continuity, $u(t)$ and $u(t') - u(t)$ are statistically independent and further, $u(t') - u(t)$ is independent of $u(t'')$ for all $t'' < t$. This means that $u(t')$ is fully determined (probabilistically) from the knowledge of the value of $u(t)$ and not by any past values. Hence, $u(t)$ is a Markov process.

Since the sample functions of $u(t)$ are continuous, we must be able to describe $u(t)$ by a Fokker-Planck equation. We can compute the drift and diffusion coefficients for this process by using the formulae of Sect.3.5.2. We can write

$$\langle u(t+\Delta t) - u_0 | [u_0, t] \rangle = \left\langle \int_t^{t+\Delta t} \xi(s) ds \right\rangle = 0 \tag{4.1.10}$$

and

$$\langle [u(t+\Delta t) - u_0]^2 | [u_0, t] \rangle = \int_t^{t+\Delta t} ds \int_t^{t+\Delta t} ds' \langle \xi(s)\xi(s') \rangle \tag{4.1.11}$$

$$= \int_t^{t+\Delta t} ds \int_t^{t+\Delta t} ds' \delta(s - s') = \Delta t \tag{4.1.12}$$

so that the drift and diffusion coefficients are

$$A(u_0, t) = \lim_{\Delta t \to 0} \frac{\langle u(t+\Delta t) - u_0 | [u_0, t] \rangle}{\Delta t} = 0 \tag{4.1.13}$$

$$B(u_0, t) = \lim_{\Delta t \to 0} \frac{\langle [u(t+\Delta t) - u_0]^2 | [u_0, t] \rangle}{\Delta t} = 1 . \tag{4.1.14}$$

Thus, the Fokker-Planck equation is that of the Wiener process and we can write

$$\int_0^t \xi(t') dt' = u(t) = W(t) . \tag{4.1.15}$$

Thus, we have the paradox that the integral of $\xi(t)$ is $W(t)$, which is itself not differentiable, as shown in Sect.3.8.1. This means that mathematically speaking, the Langevin equation (4.1.1) does not exist. However, the corresponding *integral equation*

$$x(t) - x(0) = \int_0^t a[x(s), s] ds + \int_0^t b[x(s), s] \xi(s) ds \tag{4.1.16}$$

can be interpreted consistently.

We make the replacement, which follows directly from the interpretation of the integral of $\xi(t)$ as the Wiener process $W(t)$, that

$$dW(t) \equiv W(t + dt) - W(t) = \xi(t) dt \tag{4.1.17}$$

and thus write the second integral as

$$\int_0^t b[x(s), s] dW(s) , \tag{4.1.18}$$

which is a kind of stochastic Stieltjes integral with respect to a sample function $W(t)$. Such an integral can be defined and we will carry this out in the next section.

Before doing so, it should be noted that the requirement that $u(t)$ be continuous, while very natural, can be relaxed to yield a way of defining jump processes as stochastic differential equations. This has already been hinted at in the treatment of shot noise in Sect. 1.4.1. However, it does not seem to be nearly so useful and will not be treated in this book. The interested reader is referred to [4.1].

As a final point, we should note that one normally *assumes* that $\xi(t)$ is Gaussian, and satisfies the conditions (4.1.2) as well. The above did not require this: the Gaussian nature follows in fact from the *assumed* continuity of $u(t)$. Which of these assumptions is made is, in a strict sense, a matter of taste. However, the continuity of $u(t)$ seems a much more natural assumption to make than the Gaussian nature of $\xi(t)$, which involves in principle the determination of moments of arbitrarily high order.

4.2 Stochastic Integration

4.2.1 Definition of the Stochastic Integral

Suppose $G(t)$ is an arbitrary function of time and $W(t)$ is the Wiener process. We define the stochastic integral $\int_{t_0}^{t} G(t')dW(t')$ as a kind of Riemann-Stieltjes integral. Namely, we divide the interval $[t_0, t]$ into n subintervals by means of partitioning points (as in Fig. 4.1)

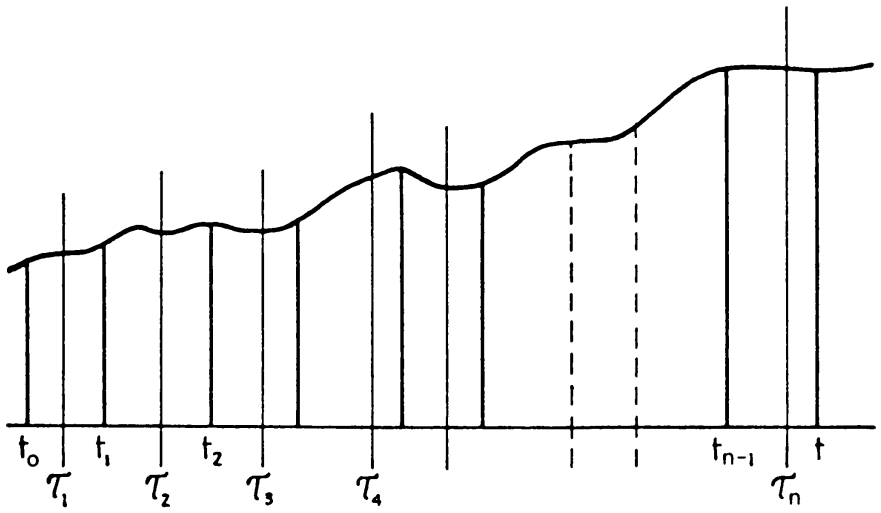


Fig. 4.1. Partitioning of the time interval used in the definition of stochastic integration

$$t_0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_{n-1} \leqslant t\,, \tag{4.2.1}$$

and define intermediate points τ_i such that

$$t_{i-1} \leqslant \tau_i \leqslant t_i\,. \tag{4.2.2}$$

The stochastic integral $\int_{t_0}^{t} G(t')dW(t')$ is defined as a limit of the partial sums.

$$S_n = \sum_{i=1}^{n} G(\tau_i)[W(t_i) - W(t_{i-1})]\,. \tag{4.2.3}$$

It is heuristically quite easy to see that, in general, the integral defined as the limit of

S_n depends on the particular choice of intermediate point τ_i. For, if we take the choice of $G(\tau_i) = W(\tau_i)$,

$$\langle S_n \rangle = \langle \sum_{i=1}^{n} W(\tau_i)[W(t_i) - W(t_{i-1})] \rangle \tag{4.2.4}$$

$$= \sum_{i=1}^{n} [\min(\tau_i, t_i) - \min(\tau_i, t_{i-1})] \tag{4.2.5}$$

$$= \sum_{i=1}^{n} (\tau_i - t_{i-1}). \tag{4.2.6}$$

If, for example, we choose for all i

$$\tau_i = \alpha t_i + (1 - \alpha) t_{i-1} \quad (0 < \alpha < 1), \tag{4.2.7}$$

$$\text{then } \langle S_n \rangle = \sum_{i=1}^{n} (t_i - t_{i-1})\, \alpha = (t - t_0)\, \alpha \tag{4.2.8}$$

So that the mean value of the integral can be anything between zero and $(t - t_0)$, depending on the choice of intermediate points.

We therefore make the choice of intermediate points characterised by $\alpha = 0$, that is, we choose

$$\tau_i = t_{i-1} \tag{4.2.9}$$

and thus define the *Ito stochastic integral* of the function $G(t)$ by

$$\int_{t_0}^{t} G(t')dW(t') = \underset{n\to\infty}{\text{ms-lim}} \left\{ \sum_{i=1}^{n} G(t_{i-1})[W(t_i) - W(t_{i-1})] \right\}. \tag{4.2.10}$$

By ms-lim we mean the *mean square limit*, as defined in Sect.2.9.2.

4.2.2 Example $\int_{t_0}^{t} W(t')dW(t')$

An exact calculation is possible. We write [writing W_i for $W(t_i)$]

$$S_n = \sum_{i=1}^{n} W_{i-1}(W_i - W_{i-1}) \equiv \sum_{i=1}^{n} W_{i-1}\Delta W_i \tag{4.2.12}$$

$$= \tfrac{1}{2} \sum_{i=1}^{n} [(W_{i-1} + \Delta W_i)^2 - (W_{i-1})^2 - (\Delta W_i)^2] \tag{4.2.13}$$

$$= \tfrac{1}{2}[W(t)^2 - W(t_0)^2] - \tfrac{1}{2} \sum_{i=1}^{n} (\Delta W_i)^2.$$

We can calculate the mean square limit of the last term. Notice that

$$\langle \sum (\Delta W_i)^2 \rangle = \sum_i \langle (W_i - W_{i-1})^2 \rangle = \sum_i (t_i - t_{i-1}) = t - t_0. \tag{4.2.14}$$

Because of this,

$$\langle[\sum_i (W_i - W_{i-1})^2 - (t - t_0]^2\rangle$$
$$= \langle \sum_i (W_i - W_{i-1})^4 + 2 \sum_{i>j} (W_i - W_{i-1})^2 (W_j - W_{j-1})^2$$
$$- 2(t - t_0) \sum_i (W_i - W_{i-1})^2 + (t - t_0)^2\rangle. \tag{4.2.15}$$

Notice the $W_i - W_{i-1}$ is a Gaussian variable and is independent of $W_j - W_{j-1}$. Hence, we can factorise. Thus,

$$\langle (W_i - W_{i-1})^2 (W_j - W_{j-1})^2\rangle = (t_i - t_{i-1})(t_j - t_{j-1}) \tag{4.2.16}$$

and also [using (2.8.6)]

$$\langle (W_i - W_{i-1})^4\rangle = 3 \langle (W_i - W_{i-1})^2\rangle^2 = 3(t_i - t_{i-1})^2 \tag{4.2.17}$$

which combined with (4.2.16) gives

$$\langle[\sum_i (W_i - W_{i-1}) - (t - t_0)]^2\rangle = 2 \sum_i (t_i - t_{i-i})^2 \tag{4.2.18}$$
$$+ \sum_{i,j} [(t_i - t_{i-1}) - (t - t_0)][(t_j - t_{j-1}) - (t - t_0)]$$
$$= 2 \sum_i (t_i - t_{i-1})^2$$
$$\to 0 \quad \text{as} \quad n \to \infty.$$

Thus,

$$\underset{n\to\infty}{\text{ms-lim}} \sum_i (W_i - W_{i-1})^2 = t - t_0 \tag{4.2.19}$$

by definition of the mean square limit, so

$$\boxed{\int_{t_0}^{t} W(t')dW(t') = \tfrac{1}{2}[W(t)^2 - W(t_0)^2 - (t - t_0)].} \tag{4.2.20}$$

Comments

i) $$\langle \int_{t_0}^{t} W(t)dW(t)\rangle = \tfrac{1}{2}[\langle W(t)^2\rangle - \langle W(t_0)^2\rangle - (t - t_0)] = 0 \tag{4.2.21}$$

This is also obvious by definition, since in the individual terms we have $\langle W_{i-1}\, \Delta W_i\rangle$ which vanishes because ΔW_i is statistically independent of W_{i-1}, as was demonstrated in Sect.3.8.1.

ii) The result for the integral is no longer the same as for the ordinary Riemann-Stieltjes integral in which the term $(t - t_0)$ would be absent. The reason for this is that $|W(t + \Delta t) - W(t)|$ is almost always of the order $\sqrt{t}$, so that in contrast to ordinary integration, terms of second order in $\Delta W(t)$ do not vanish on taking the limit.

4.2.3 The Stratonovich Integral

An alternative definition was introduced by *Stratonovich* [4.2] as a stochastic integral in which the anomalous term $(t - t_0)$ does not occur. We define this fully in Sect. 4.3.6 – in the cases considered so far, it amounts to evaluating the integrand as a function of $W(t)$ at the value $\frac{1}{2}[W(t_i) + W(t_{i-1})]$. It is straightforward to show that

$$S\int_{t_0}^{t} W(t')\,dW(t')$$

$$= \operatorname*{ms\text{-}lim}_{n\to\infty} \sum_{i=1}^{n} \frac{W(t_i) + W(t_{i-1})}{2}[W(t_i) - W(t_{i-1})] \tag{4.2.22}$$

$$= \tfrac{1}{2}[W(t)^2 - W(t_0)^2]\,. \tag{4.2.23}$$

However, the integral as defined by Stratonovich [which we will always designate by a prefixed S as in (4.2.24)] has no general relationship with that defined by Ito. That is, for *arbitrary* functions $G(t)$, there is no connection between the two integrals. [In the case, however, where we can specify that $G(t)$ is related to some stochastic differential equation, a formula can be given relating one to the other, see Sect.4.3.6].

4.2.4 Nonanticipating Functions

The concept of a nonanticipating function can be easily made quite obscure by complex notation, but is really quite simple. We have in mind a situation in which all functions can be expressed as functions or functionals of a certain Wiener process $W(t)$ through the mechanism of a stochastic differential (or integral) equation of the form

$$x(t) - x(t_0) = \int_{t_0}^{t} a[x(t'),\ t']dt' + \int_{t_0}^{t} b[x(t'), t']dW(t')\,. \tag{4.2.24}$$

A function $G(t)$ is called a *nonanticipating function of* t if for all s and t such that $t < s$, $G(t)$ is statistically independent of $W(s) - W(t)$. This means that $G(t)$ is independent of the behaviour of the Wiener process in the future of t. This is clearly a rather reasonable requirement for a physical function which could be a solution of an equation like (4.2.24) in which it seems heruistically obvious that $x(t)$ involves $W(t')$ only for $t' \leqslant t$.

For example, specific nonanticipating functions of t are:

i) $W(t)$

ii) $\int_{t_0}^{t} dt' F[W(t')]$

iii) $\int_{t_0}^{t} dW(t')\,F[W(t')]$

iv) $\int_{t_0}^{t} dt' G(t')$

v) $\int_{t_0}^{t} dW(t') G(t')$

} when $G(t)$ is itself a nonanticipating function

Results (iii) and (v) depend on the fact that the Ito stochastic integral, as defined in (4.2.10), is a limit of the sequence in which only $G(t')$ for $t' < t$ and $W(t')$ for $t' \leqslant t$ are involved.

The reasons for considering nonanticipating functions specifically are:
i) many results can be derived, which are only true for such functions;
ii) they occur naturally in situations, such as in the study of differential equations involving time, in which some kind of *causality* is expected in the sense that the unknown future cannot affect the present;
iii) the definition of stochastic differential equations requires such functions.

4.2.5 Proof that $dW(t)^2 = dt$ and $dW(t)^{2+N} = 0$

The formulae in the heading are the key to the use of the Ito calculus as an ordinary computational tool. However, as written they are not very precise and what is really meant is

$$\begin{aligned} \int_{t_0}^{t} [dW(t')]^{2+N} G(t') &\equiv \underset{n\to\infty}{\text{ms-lim}} \sum_i G_{i-1} \Delta W_i^{2+N} \\ &= \int_{t_0}^{t} dt' G(t') \text{ for } \quad N = 0 \\ &= 0 \quad \text{for } N > 0 \end{aligned} \tag{4.2.25}$$

for an arbitrary *nonanticipating function* $G(t)$.

The proof is quite straightforward. For $N = 0$, let us define

$$I = \lim_{n\to\infty} \langle [\sum_i G_{t-1}(\Delta W_i^2 - \Delta t_i)]^2 \rangle \tag{4.2.26}$$

$$= \lim_{n\to\infty} \{ \langle \sum_i \underbrace{(G_{t-1})^2}_{\text{Independent}} \underbrace{(\Delta W_i^2 - \Delta t_i)^2}_{\text{Independent}} + \sum_{i>j} \underbrace{2 G_{t-1} G_{j-1} (\Delta W_j^2 - \Delta t_j)}_{\text{Independent}} \underbrace{(\Delta W_i^2 - \Delta t_i)}_{\text{Independent}} \rangle \}. \tag{4.2.27}$$

The horizontal braces indicate factors which are statistically independent of each other because of the properties of the Wiener process, and because the G_t are values of a nonanticipating function which are independent of all ΔW_t for $j > i$.

Using this independence, we can factorise the means, and also using

i) $\langle \Delta W_t^2 \rangle = \Delta t_t$

ii) $\langle (\Delta W_t^2 - \Delta t_t)^2 \rangle = 2\Delta t_t^2$ (from Gaussian nature of ΔW_t),

we find

$$I = 2 \lim_{n\to\infty} [\sum_i \Delta t_i^2 \langle (G_{t-1})^2 \rangle] \, . \tag{4.2.28}$$

Under reasonably mild conditions on $G(t)$ (e.g., boundedness), this means that

$$\underset{n\to\infty}{\text{ms-lim}} \left(\sum_i G_{t-1}\Delta W_i^2 - \sum_i G_{t-1}\Delta t_i\right) = 0 \tag{4.2.29}$$

and since

$$\underset{n\to\infty}{\text{ms-lim}} \sum_i G_{t-1}\Delta t_i = \int_{t_0}^{t} dt' G(t') \,, \tag{4.2.30}$$

we have

$$\boxed{\int_{t_0}^{t} [dW(t')]^2 G(t') = \int_{t_0}^{t} dt' G(t') \,.}$$

Comments

i) The proof $\int_{t_0}^{t} G(t)[dW(t)]^{2+N} = 0$ for $N > 0$ is similar and uses the explicit expressions for the higher moments of a Gaussian (Sect.2.8.1).

ii) $dW(t)$ only occurs in integrals so that when we restrict ourselves to nonanticipating functions, we can simply write

$$dW(t)^2 \equiv dt \tag{4.2.31}$$

$$dW(t)^{2+N} \equiv 0 \; (N > 0) \,. \tag{4.2.32}$$

iii) The results are only valid for the Ito integral, since we have used the fact that ΔW_t is independent of G_{t-1}. In the Stratonovich integral,

$$\Delta W_t = W(t_t) - W(t_{t-1}) \tag{4.2.33}$$

$$G_{t-1} = G[\tfrac{1}{2}(t_t + t_{t-1})] \tag{4.2.34}$$

and although $G(t)$ is nonanticipating, this is not sufficient to guarantee the independence of ΔW_t and G_{t-1} as thus defined.

iv) By similar methods one can prove that

$$\int_{t_0}^{t} G(t')dt'dW(t') \equiv \underset{n\to\infty}{\text{ms-lim}} \sum G_{t-1}\Delta W_t \Delta t_t = 0 \tag{4.2.35}$$

and similarly for higher powers. The simplest way of characterising these results is to say that $dW(t)$ is an infinitesimal order of $\frac{1}{2}$ and that in calculating differentials, infinitesimals of order higher than 1 are discarded.

4.2.6 Properties of the Ito Stochastic Integral

a) Existence

One can show that the Ito stochastic integral $\int_{t_0}^{t} G(t')dW(t')$ exists whenever the function $G(t')$ is *continuous* and *nonanticipating* on the closed interval $[t_0, t]$ [4.3].

b) Integration of Polynomials
We can formally use the result of Sect.4.2.5:

$$d[W(t)]^n = [W(t) + dW(t)]^n - W(t)^n = \sum_{r=1}^{n} \binom{n}{r} W(t)^{n-r} dW(t)^r$$

and using the fact that $dW(t)^r \to 0$ for all $r > 2$,

$$= nW(t)^{n-1}dW(t) + \frac{n(n-1)}{2} W(t)^{n-2}dt \tag{4.2.36}$$

so that

$$\boxed{\int_{t_0}^{t} W(t')^n dW(t') = \frac{1}{n+1}[W(t)^{n+1} - W(t_0)^{n+1}] - \frac{n}{2}\int_{t_0}^{t} W(t)^{n-1}dt.} \tag{4.2.37}$$

c) Two Kinds of Integral
We note that for each $G(t)$ there are two kinds of integrals, namely,

$$\int_{t_0}^{t} G(t')dt' \quad \text{and} \quad \int_{t_0}^{t} G(t')dW(t'),$$

both of which occur in the previous equation. There is, in general, no connection between these two kinds of integral.

d) General Differentiation Rules
In forming differentials [as in(b) above], one must keep all terms up to second order in $dW(t)$. This means that, for example,

$$\begin{aligned} d\{\exp[W(t)]\} &= \exp[W(t) + dW(t)] - \exp[W(t)] \\ &= \exp[W(t)][dW(t) + \tfrac{1}{2}dW(t)^2] \\ &= \exp[W(t)][dW(t) + \tfrac{1}{2}dt] \end{aligned} \tag{4.2.38}$$

or more generally,

$$\begin{aligned} df[W(t), t] = \frac{\partial f}{\partial t}dt + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial f}{\partial W}dW(t) + \frac{1}{2}\frac{\partial^2 f}{\partial W}[dW(t)]^2 \\ + \frac{\partial^2 f}{\partial W \partial t}dtdW(t) + \ldots \end{aligned}$$

and we use $(dt)^2 \to 0$

$dtdW(t) \to 0$ [Sect. 4.2.5, comment (iv)]

$[dW(t)]^2 = dt$

and all higher powers vanish, to arrive at

$$\boxed{df[W(t), t] = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W^2}\right)dt + \frac{\partial f}{\partial W}dW(t)\,.} \tag{4.2.39}$$

e) Mean Value Formula
For nonanticipating $G(t)$,

$$\boxed{\left\langle \int_{t_0}^{t} G(t')dW(t') \right\rangle = 0\,.} \tag{4.2.40}$$

For, since $G(t)$ is nonanticipating, in the definition of the stochastic integral,

$$\left\langle \sum_i G_{i-1}\Delta W_i \right\rangle = \sum_i \langle G_{i-1}\rangle\langle \Delta W_i\rangle = 0 \tag{4.2.41}$$

and we know from Sect. 2.9.5 that operations of ms-lim and $\langle\ \rangle$ may be interchanged. Hence, taking the limit of (4.2.41), we have the result.

This result is *not true* for Stratonovich's integral, since the value of G_{t-1} is chosen in the middle of the interval, and can be correlated with ΔW_t.

f) Correlation Formula
If $G(t)$ and $H(t)$ are arbitrary continuous nonanticipating functions,

$$\boxed{\left\langle \int_{t_0}^{t} G(t')dW(t') \int_{t_0}^{t} H(t')dW(t') \right\rangle = \int_{t_0}^{t} dt\, \langle G(t')H(t')\rangle\,.} \tag{4.2.42}$$

Proof. Notice that

$$\left\langle \sum_i G_{i-1}\Delta W_i \sum_j H_{j-1}\Delta W_j \right\rangle = \left\langle \sum_i G_{i-1}H_{i-1}(\Delta W_i)^2 \right\rangle + \left\langle \sum_{i>j} (G_{i-1}H_{j-1} + G_{j-1}H_{i-1})\Delta W_j \Delta W_i \right\rangle\,. \tag{4.2.43}$$

In the second term, ΔW_i is independent of all other terms since $j < i$, and G and H are nonanticipating. Hence, we may factorise out the term $\langle \Delta W_i\rangle = 0$ so that this term vanishes. Using

$$\langle \Delta W_i^2\rangle = \Delta t_i \tag{4.2.44}$$

and interchanging mean and limit operations, the result follows.

Formally, this is equivalent to the idea that Langevin terms $\xi(t)$ are delta correlated and uncorrelated with $F(t)$ and $G(t)$. For, rewriting

$$dW(t) \to \xi(t)dt\,, \tag{4.2.45}$$

it is clear that if $F(t)$ and $G(t)$ are nonanticipating, $\xi(t)$ is independent of them, and we get

$$\int_{t_0}^{t} dt' \int_{t_0}^{t} ds' \langle G(t')H(s')\xi(t')\xi(s')\rangle = \int_{t_0}^{t}\int_{t_0}^{t} dt'ds' \langle G(t')H(s')\rangle\langle \xi(t')\xi(s')\rangle$$
$$= \int_{t_0}^{t} dt' \langle G(t')H(t')\rangle \tag{4.2.46}$$

which implies

$$\langle \xi(t)\xi(s)\rangle = \delta(t-s)\,.$$

An important point of definition arises here, however. In integrals involving delta functions, it frequently occurs in the study of stochastic differential equations that the argument of the delta function is equal to either the upper or the lower limit of the integral, that is, we find integrals like

$$I_1 = \int_{t_1}^{t_2} dt\, f(t)\delta(t-t_1) \tag{4.2.47}$$

or

$$I_2 = \int_{t_1}^{t_2} dt\, f(t)\delta(t-t_2) \tag{4.2 48}$$

and various conventions can be made concerning the value of such integrals. We will show that in the present context, we must always make the interpretation

$$I_1 = f(t_1) \tag{4.2.49}$$

$$I_2 = 0 \tag{4.2.50}$$

corresponding to counting all the weight of a delta function at the lower limit of an integral, and none of the weight at the upper limit. To demonstrate this, note that

$$\langle \int_{t_0}^{t} G(t')dW(t') \,[\int_{t_0}^{t'} H(s')dW(s')]\rangle = 0 \tag{4.2.51}$$

This follows, since the function defined by the integral inside the square bracket is, by Sect.4.2.4 comment (v), a nonanticipating function and hence the complete integrand, [obtained by multiplying by $G(t')$ which is also nonanticipating] is itself nonanticipating. Hence the average vanishes by the result of Sect. 4.2.6e.

Now using the formulation in terms of the Langevin source $\xi(t)$, we can rewrite (4.2.52) as

$$\int_{t_0}^{t} dt' \int_{t_0}^{t'} ds' \langle G(t')H(s')\rangle \delta(t'-s') = 0 \tag{4.2.52}$$

which corresponds to not counting the weight of the delta function at the upper limit. Consequently, the full weight must be counted at the lower limit.

This property is a direct consequence of the definition of the Ito integral as in (4.2.10), in which the increment points "towards the future". That is, we can interpret

$$dW(t) = W(t + dt) - W(t)\,. \tag{4.2.53}$$

In the case of the Stratonovich integral, we get quite a different formula, which is by no means as simple to prove as in the Ito case, but which amounts to choosing

$$\left.\begin{aligned} I_1 &= \tfrac{1}{2}f(t_1) \\ I_2 &= \tfrac{1}{2}f(t_2). \end{aligned}\right\} \text{(Stratonovich)} \tag{4.2.54}$$

This means that in both cases, the delta function occuring at the limit of an integral has $\frac{1}{2}$ its weight counted. This formula, although intuitively more satisfying than the Ito form, is more complicated to use, especially in the perturbation theory of stochastic differential equations, where the Ito method makes very many terms vanish.

4.3 Stochastic Differential Equations (SDE)

We concluded in Sect.4.1, that the most satisfactory interpretation of the Langevin equation

$$\frac{dx}{dt} = a(x, t) + b(x, t)\xi(t) \tag{4.3.1}$$

is a stochastic integral equation

$$x(t) - x(0) = \int_0^t dt' a[x(t'), t'] + \int_0^t dW(t')b[x(t'), t']\,. \tag{4.3.2}$$

Unfortunately, the kind of stochastic integral to be used is not given by the reasoning of Sect.4.1. The Ito integral is mathematically and technically the most satisfactory, but unfortunately, it is not always the most natural choice physically. The Stratonovich integral is the natural choice for an interpretation which assumes $\xi(t)$ is a real noise (not a white noise) with finite correlation time, which is then allowed to become infinitesimally small after calculating measurable quantities. Furthermore, a Stratonovich interpretation enables us to use ordinary calculus, which is not possible for an Ito interpretation.

From a mathematical point of view, the choice is made clear by the near impossibility of carrying out proofs using the Stratonovich integral. We will therefore define the Ito SDE, develop its equivalence with the Stratonovich SDE, and use either form depending on circumstances. The relationship between white noise stochastic differential equations and the real noise systems is explained in Sect.6.5.

4.3.1 Ito Stochastic Differential Equation: Definition

A stochastic quantity $x(t)$ obeys an Ito SDE written as

$$dx(t) = a[x(t), t]dt + b[x(t), t]dW(t) \tag{4.3.3}$$

if for all t and t_0,

$$x(t) = x(t_0) + \int_{t_0}^{t} dt'\, a[x(t'), t'] + \int_{t_0}^{t} dW(t')\, b[x(t'), t'] . \tag{4.3.4}$$

Before considering what conditions must be satisfied by the coefficients in (4.3.4), it is wise to consider what one means by a solution of such an equation and what uniqueness of solution would mean in this context. For this purpose, we can consider a discretised version of the SDE obtained by taking a mesh of points t_i (as illustrated in Fig. 4.2) such that

$$t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = t \tag{4.3.5}$$

and writing the equation as

$$x_{i+1} = x_i + a(x_i, t_i)\Delta t_i + b(x_i, t_i)\Delta W_i . \tag{4.3.6}$$

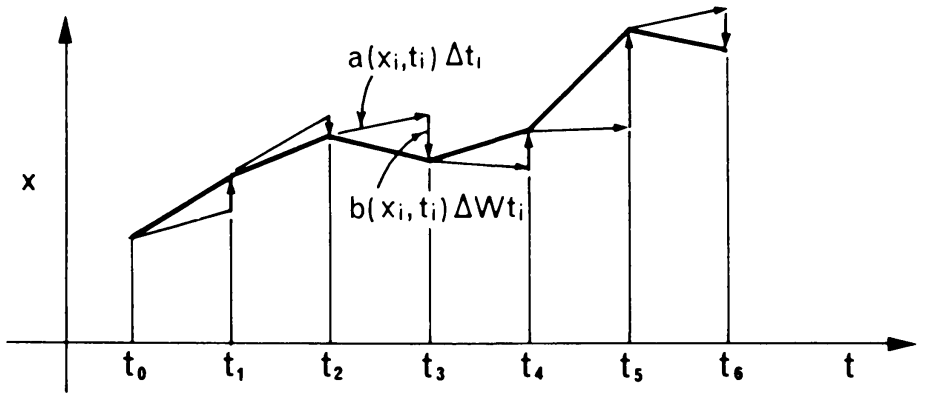


Fig. 4.2. Illustration of the Cauchy-Euler procedure for constructing an approximate solution of the stochastic differential equation $dx(t) = a[x(t), t]dt + b[x(t), t]dW(t)$

Here,

$$\left.\begin{aligned} x_i &= x(t_i) \\ \Delta t_i &= t_{i+1} - t_i \\ \Delta W_i &= W(t_{i+1}) - W(t_i). \end{aligned}\right\} \tag{4.3.7}$$

We see from (4.3.6) that an approximate procedure for solving the equation is to calculate x_{i+1} from the knowledge of x_i by adding a deterministic term

$$a(x_i, t_i)\Delta t_i \tag{4.3.8}$$

and a stochastic term

$$b(x_t, t_i)\Delta W_i \,. \tag{4.3.9}$$

The stochastic term contains an element ΔW_t, which is the increment of the Wiener process, but is statistically independent of x_t if (i) x_0 is itself independent of all $W(t) - W(t_0)$ for $t > t_0$ (thus, the initial conditions if considered random, must be nonanticipating) and (ii) $a(x, t)$ is a nonanticipating function of t for any fixed x.

Constructing an approximate solution iteratively by use of (4.3.6), we see that x_t is always independent of ΔW_j for $j \geqslant i$.

The solution is then formally constructed by letting the mesh size go to zero. To say that the solution is unique means that for a given sample function $\tilde{W}(t)$ of the random Wiener process $W(t)$, the particular solution of the equation which arises is unique. To say that the solution exists means that with probability one, a solution exists for any choice of sample function $\tilde{W}(t)$ of the Wiener process $W(t)$.

The method of constructing a solution outlined above is called the *Cauchy-Euler* method, and can be used to generate simulations.

However, this is not the way uniqueness and existence are usually demonstrated, though it is possible to demonstrate these properties this way. Existence and uniqueness will not be proved here. The interested reader will find proofs in [4.3]. The conditions which are required for existence and uniqueness in a time interval $[t_0, T]$ are:

i) *Lipschitz condition:* a K exists such that

$$|a(x, t) - a(y, t)| + |b(x, t) - b(y, t)| \leqslant K|x - y| \tag{4.3.10}$$

for all x and y, and all t in the range $[t_0, T]$.

ii) *growth condition:* a K exists such that for all t in the range $[t_0, T]$,

$$|a(x, t)|^2 + |b(x, t)|^2 \leqslant K^2(1 + |x|^2) \,. \tag{4.3.11}$$

Under these conditions there will be a unique nonanticipating solution $x(t)$ in the range $[t_0, T]$.

Almost every stochastic differential equation encountered in practice satisfies the Lipschitz condition since it is essentially a smoothness condition. However, the growth condition is often violated. This does not mean that no solution exists; rather, it means the solution may "explode" to infinity, that is, the value of x can become infinite in a finite time; in practice, a finite random time. This phenomenon occurs in ordinary differential equations, for example,

$$\frac{dx}{dt} = \frac{1}{2} a x^3 \tag{4.3.12}$$

has the general solution with an initial condition $x = x_0$ at $t = 0$,

$$x(t) = (-at + 1/x_0^2)^{-1/2} \,. \tag{4.3.13}$$

If a is positive, this becomes infinite when $x_0 = (at)^{-1/2}$ but if a is negative, the solution never explodes. Failing to satisfy the Lipschitz condition does not guarantee the solution will explode. More precise stability results are required for one to be certain of that [4.3].

4.3.2 Markov Property of the Solution of an Ito Stochastic Differential Equation

We now show that $x(t)$, the solution to the stochastic differential equation (4.3.4), is a Markov Process. Heuristically, the result is obvious, since with a given initial condition $x(t_0)$, the future time development is uniquely (stochastically) determined, that is, $x(t)$ for $t > t_0$ is determined only by

i) the particular sample path of $W(t)$ for $t > t_0$;
ii) the value of $x(t_0)$.

Since $x(t)$ is a nonanticipating function of t, $W(t)$ for $t > t_0$ is independent of $x(t)$ for $t < t_0$. Thus, the time development of $x(t)$ for $t > t_0$ is independent of $x(t)$ for $t < t_0$ provided $x(t_0)$ is known. Hence, $x(t)$ is a Markov process. For a precise proof see [4.3].

4.3.3 Change of Variables: Ito's Formula

Consider an arbitrary function of $x(t)$: $f[x(t)]$. What stochastic differential equation does it obey? We use the results of Sect.4.2.5 to expand $df[x(t)]$ to second order in $dW(t)$:

$$\begin{aligned} df[x(t)] &= f[x(t) + dx(t)] - f[x(t)] \\ &= f'[x(t)]dx(t) + \tfrac{1}{2}f''[x(t)]dx(t)^2 + \dots \\ &= f'[x(t)]\{a[x(t), t]dt + b[x(t), t]dW(t)\} \\ &\quad + \tfrac{1}{2}f''[x(t)]b[x(t), t]^2[dW(t)]^2 + \dots , \end{aligned}$$

where all other terms have been discarded since they are of higher order. Now use $[dW(t)]^2 = dt$ to obtain

$$\boxed{\begin{aligned} df[x(t)] = &\{a[x(t), t]f'[x(t)] + \tfrac{1}{2}b[x(t), t]^2 f''[x(t)]\}\, dt \\ &+ b[x(t), t]f'[x(t)]dW(t) . \end{aligned}} \tag{4.3.14}$$

This formula is known as Ito's *formula* and shows that changing variables is not given by ordinary calculus unless $f[x(t)]$ is merely linear in $x(t)$.

Many Variables. In practice, Ito's formula becomes very complicated and the easiest method is to simply use the multivariate form of the rule that $dW(t)$ is an infinitesmial of order $\frac{1}{2}$. By similar methods to those used in Sect.4.2.5, we can show that for an n dimensional Wiener process $\boldsymbol{W}(t)$,

$$dW_i(t)dW_j(t) = \delta_{ij}dt \tag{4.3.15a}$$

$$[dW_i(t)]^{N+2} = 0 \quad (N > 0) \tag{4.3.15b}$$

$$dW_i(t)dt = 0 \tag{4.3.15c}$$

$$dt^{1+N} = 0 \quad (N > 0)\,. \tag{4.3.15d}$$

which imply that $dW_i(t)$ is an infinitesimal of order $\frac{1}{2}$. Note, however, that (4.3.15a) is a consequence of the independence of $dW_i(t)$ and $dW_j(t)$. To develop Ito's formula for functions of an n dimensional vector $\boldsymbol{x}(t)$ satisfying the stochastic differential equation

$$d\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x}, t)dt + \underset{\sim}{\boldsymbol{B}}(\boldsymbol{x}, t)d\boldsymbol{W}(t)\,, \tag{4.3.16}$$

we simply follow this procedure. The result is

$$\boxed{\begin{aligned} df(\boldsymbol{x}) = &\{\sum_i A_i(\boldsymbol{x}, t)\partial_i f(\boldsymbol{x}) + \tfrac{1}{2}\sum_{i,j}[\underset{\sim}{\boldsymbol{B}}(\boldsymbol{x}, t)\underset{\sim}{\boldsymbol{B}}^{\mathrm{T}}(\boldsymbol{x}, t)]_{ij}\partial_i\partial_j f(\boldsymbol{x})\}dt \\ &+ \sum_{i,j} B_{ij}(\boldsymbol{x}, t)\partial_i f(\boldsymbol{x})dW_j(t)\,. \end{aligned}} \tag{4.3.17}$$

4.3.4 Connection Between Fokker-Planck Equation and Stochastic Differential Equation

We now consider the time development of an arbitrary $f(x(t))$. Using Ito's formula

$$\begin{aligned} \langle df[x(t)]\rangle/dt = \left\langle \frac{df[x(t)]}{dt} \right\rangle &= \frac{d}{dt}\langle f[x(t)]\rangle \\ &= \langle a[x(t), t]\partial_x f + \tfrac{1}{2}b[x(t), t]^2\partial_x^2 f\rangle\,. \end{aligned} \tag{4.3.18}$$

However, $x(t)$ has a conditional probability density $p(x, t\,|\,x_0, t_0)$ and

$$\begin{aligned} \frac{d}{dt}\langle f[x(t)]\rangle &= \int dx\, f(x)\partial_t p(x, t\,|\,x_0, t_0) \\ &= \int dx[a(x, t)\partial_x f + \tfrac{1}{2}b(x, t)^2\partial_x^2 f]p(x, t\,|\,x_0, t_0)\,. \end{aligned} \tag{4.3.19}$$

This is now of the same form as (3.4.16) Sect.3.4.1. Under the same conditions as there, we integrate by parts and discard surface terms to obtain

$$\int dx\, f(x)\partial_t p = \int dx f(x)\,\{-\partial_x[a(x, t)p] + \tfrac{1}{2}\partial_x^2[b(x, t)^2 p]\}$$

and hence, since $f(x)$ is arbitrary,

$$\boxed{\partial_t p(x, t\,|\,x_0, t_0) = -\partial_x[a(x, t)p(x, t\,|\,x_0, t_0)] + \tfrac{1}{2}\partial_x^2[b(x, t)^2 p(x, t\,|\,x_0, t_0)]\,.} \tag{4.3.20}$$

We have thus a complete equivalence to a diffusion process defined by a drift coefficient $a(x, t)$ and a diffusion coefficient $b(x, t)^2$.

The results are precisely analogous to those of Sect.3.5.3, in which it was shown that the diffusion process could be locally approximated by an equation resembling an Ito stochastic differential equation.

4.3.5 Multivariable Systems

In general, many variable systems of stochastic differential equations can be defined for n variables by

$$d\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x}, t)dt + \underset{\sim}{\boldsymbol{B}}(\boldsymbol{x}, t)d\boldsymbol{W}(t), \tag{4.3.21}$$

where $d\boldsymbol{W}(t)$ is an n variable Wiener process, as defined in Sect.3.8.1. The many variable version of the reasoning used in Sect. 4.3.4 shows that the Fokker-Planck equation for the conditional probability density $p(\boldsymbol{x}, t \mid \boldsymbol{x}_0, t_0) \equiv p$ is

$$\partial_t p = -\sum_i \partial_i [A_i(\boldsymbol{x}, t)p] + \tfrac{1}{2} \sum_{i,j} \partial_i \partial_j \{[\underset{\sim}{\boldsymbol{B}}(\boldsymbol{x}, t)\underset{\sim}{\boldsymbol{B}}^{\mathrm{T}}(\boldsymbol{x}, t)]_{ij} p\}. \tag{4.3.22}$$

Notice that the same Fokker-Planck equation arises from all matrices $\underset{\sim}{\boldsymbol{B}}$ such that $\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{B}}^{\mathrm{T}}$ is the same. This means that we can obtain the same Fokker-Planck equation by replacing $\underset{\sim}{\boldsymbol{B}}$ by $\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{S}}$ where $\underset{\sim}{\boldsymbol{S}}$ is orthogonal, i.e., $\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{S}}^{\mathrm{T}} = 1$. Notice that $\underset{\sim}{\boldsymbol{S}}$ may depend on $\boldsymbol{x}(t)$. This can be seen more directly. Suppose $\underset{\sim}{\boldsymbol{S}}(t)$ is an orthogonal matrix with an arbitrary *nonanticipating* dependence on t. Then define

$$d\boldsymbol{V}(t) = \underset{\sim}{\boldsymbol{S}}(t)d\boldsymbol{W}(t). \tag{4.3.23}$$

Now the vector $d\boldsymbol{V}(t)$ is a linear combination of Gaussian variables $d\boldsymbol{W}(t)$ with coefficients $\underset{\sim}{\boldsymbol{S}}(t)$ which are independent of $d\boldsymbol{W}(t)$, since $\underset{\sim}{\boldsymbol{S}}(t)$ is nonanticipating. For any fixed value of $\underset{\sim}{\boldsymbol{S}}(t)$, the $d\boldsymbol{V}(t)$ are thus Gaussian and their correlation matrix is

$$\begin{aligned} \langle dV_i(t)dV_j(t)\rangle &= \sum_{l,m} S_{il}(t)S_{jm}(t)\langle dW_l(t)dW_m(t)\rangle \\ &= \sum_l S_{il}(t)S_{jl}(t)dt = \delta_{ij}dt \end{aligned} \tag{4.3.24}$$

since $\underset{\sim}{\boldsymbol{S}}(t)$ is orthogonal. Hence, all the moments are independent of $\underset{\sim}{\boldsymbol{S}}(t)$ and are the same as those of $d\boldsymbol{W}(t)$, so $d\boldsymbol{V}(t)$ is itself Gaussian with the same correlation matrix as $d\boldsymbol{W}(t)$. Finally, averages at different times factorise, for example, if $t > t'$ in

$$\sum_{i,k} \langle [dW_i(t)S_{ij}(t)]^m [dW_k(t')S_{kl}(t')]^n \rangle , \tag{4.3.25}$$

we can factorise out the averages of $dW_i(t)$ to various powers since $dW_i(t)$ is independent of all other terms. Evaluating these we will find that the orthogonal nature of $\underset{\sim}{\boldsymbol{S}}(t)$ gives, after averaging over $dW_i(t)$, simply

$$\sum_k \langle [dW_j(t)]^m \rangle \langle [dW_k(t') S_{kl}(t')]^n \rangle \tag{4.3.26}$$

which similarly gives $\langle [dW_j(t)]^m [dW_l(t')]^n \rangle$. Hence, the $dV(t)$ are also increments of a Wiener process. The orthogonal transformation simply mixes up different sample paths of the process, without changing its stochastic nature.

Hence, instead of (4.3.21) we can write

$$d\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x}, t)dt + \underline{\boldsymbol{B}}(\boldsymbol{x}, t)\underline{\boldsymbol{S}}^{\mathrm{T}}(t)\underline{\boldsymbol{S}}(t)d\boldsymbol{W}(t) \tag{4.3.27}$$

$$= \boldsymbol{A}(\boldsymbol{x}, t)dt + \underline{\boldsymbol{B}}(\boldsymbol{x}, t)\underline{\boldsymbol{S}}^{\mathrm{T}}(t)d\boldsymbol{V}(t)\,, \tag{4.3.28}$$

and since $V(t)$ is itself simply a Wiener process, this equation is equivalent to

$$d\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x}, t)dt + \underline{\boldsymbol{B}}(\boldsymbol{x}, t)\underline{\boldsymbol{S}}^{\mathrm{T}}(t)d\boldsymbol{W}(t) \tag{4.3.29}$$

which has exactly the same Fokker-Planck equation (4.3.22).

We will return to some examples in which this identity is relevant in Sect.4.4.6.

4.3.6 Stratonovich's Stochastic Differential Equation

Stratonovich [4.2] has defined a stochastic integral of an integrand which is a function of $x(t)$ and t by

$$S\int_{t_0}^{t} G[x(t'), t']\, dW(t') = \underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n} G\left\{\frac{x(t_i) + x(t_{i-1})}{2}, t_{i-1}\right\}[W(t_i) - W(t_{i-1})] \tag{4.3.30}$$

It should be noted that only the dependence on $x(t)$ is averaged. If $G(z, t)$ is differentiable in t, the integral is independent of the particular choice of value for t in the range $[t_{i-1}, t_i]$.

It is possible to write a stochastic differential equation (SDE) using Stratonovich's integral

$$x(t) = x(t_0) + \int_{t_0}^{t} dt'\alpha[x(t'), t'] + S\int_{t_0}^{t} dW(t')\beta[x(t'), t']\,, \tag{4.3.31}$$

and we shall show that is equivalent to an appropriate Ito SDE.

Let us assume that $x(t)$ is a solution of

$$dx(t) = a[x(t), t]dt + b[x(t), t]dW(t) \tag{4.3.32}$$

and deduce the corresponding α and β. In both cases, the solution $x(t)$ is the same function. We first compute the connection between $S\int_{t_0}^{t} dW(t')\beta[x(t'), t']$ and $\int_{t_0}^{t} dW(t')b[x(t'), t']$. Then,

$$S\int_{t_0}^{t} dW(t')\,\beta[x(t'), t'] \simeq \sum_i \beta\left[\frac{x(t_i) + x(t_{i-1})}{2}, t_{i-1}\right][W(t_i) - W(t_{i-1})]. \tag{4.3.33}$$

In (4.3.33) we write

$$x(t_i) = x(t_{i-1}) + dx(t_{i-1})$$

and use the Ito SDE (4.3.32) to write

$$dx(t_i) = a[x(t_{i-1}), t_{i-1}](t_i - t_{i-1}) + b[x(t_{i-1}), t_{i-1}][W(t_i) - W(t_{i-1})]. \tag{4.3.34}$$

Then, applying Ito's formula, we can write

$$\begin{aligned}\beta\left[\frac{x(t_i) + x(t_{i-1})}{2}, t_{i-1}\right] &= \beta[x(t_{i-1}) + \tfrac{1}{2}dx(t_{i-1}), t_{i-1}] \\ &= \beta(t_{i-1}) + [a(t_{i-1})\,\partial_x\beta(t_{i-1}) + \tfrac{1}{4}b^2(t_{i-1})][\tfrac{1}{2}(t_i - t_{i-1})] \\ &\quad + \tfrac{1}{2}b(t_{i-1})\,\partial_x^2\beta(t_{i-1})[W(t_i) - W(t_{i-1})]\end{aligned} \tag{4.3.35}$$

(For simplicity, we write $\beta(t_i)$ etc, instead of $\beta[x(t_i), t_i]$ wherever possible). Putting all these back in the original equation (4.3.32) and dropping as usual dt^2, $dt\,dW$, and setting $dW^2 = dt$, we find

$$\begin{aligned}S\int &= \sum_i \beta(t_{i-1})\{W(t_i) - W(t_{i-1})\} \\ &\quad + \tfrac{1}{2}\sum_i b(t_{i-1})\,\partial_x\beta(t_{i-1})(t_i - t_{i-1}).\end{aligned}$$

Hence we derive

$$\boxed{S\int_{t_0}^{t}\beta[x(t'), t']dW(t') = \int_{t_0}^{t}\beta[x(t'), t']dW(t') + \tfrac{1}{2}\int_{t_0}^{t} b[x(t'), t']\partial_x\beta[x(t'), t']dt'.} \tag{4.3.37}$$

This formula gives a connection between the Ito and Stratonovich integrals of functions $\beta[x(t'), t']$, in which $x(t')$ is the solution of the Ito SDE (4.3.31). It does not give a general connection between the Ito and Stratonovich integrals of arbitrary functions.

If we now make the choice

$$\begin{aligned}\alpha(x, t) &= a(x, t) - \tfrac{1}{2}b(x, t)\partial_x b(x, t) \\ \beta(x, t) &= b(x, t)\end{aligned} \tag{4.3.38}$$

We see that the Ito SDE	$dx = a\,dt + b\,dW(t),$	(4.3.39a)
is the same as the Stratonovich SDE	$dx = [a - \tfrac{1}{2}b\partial_x b]dt + b\,dW(t),$	(4.3.39b)

or conversely,

the Stratonovich SDE	$dx = \alpha\,dt + \beta\,dW(t)$	(4.3.40a)
is the same as the Ito SDE	$dx = [\alpha + \tfrac{1}{2}\beta\partial_x\beta]dt + \beta\,dW(t).$	(4.3.40b)

Comments

i) Using Ito's formula (4.3.14) we can show that the rule for a change of variables in Stratonovich SDE is exactly the same as in ordinary calculus. Start with the Stratonovich SDE (4.3.40a), and convert to the Ito SDE (4.3.40b). Change to the new variable $y = f(x)$ with the inverse $x = g(y)$.

Define $\bar{\alpha}(y) = \alpha[g(y)]$

$\bar{\beta}(y) = \beta[g(y)]$.

Use Ito's formula and note that $df/dx = (dg/dy)^{-1}$ to obtain the Ito SDE

$$dy = \left[\bar{\alpha}\left(\frac{dg}{dy}\right)^{-1} + \tfrac{1}{2}\bar{\beta}\partial_y\bar{\beta}\left(\frac{dg}{dy}\right)^{-2} - \tfrac{1}{2}\frac{d^2g}{dy^2}\left(\frac{dg}{dy}\right)^{-3}\bar{\beta}^2\right]dt + \left(\frac{dg}{dy}\right)^{-1}\bar{\beta}dW .$$

Now convert back to a Stratonovich equation using (4.3.39); we obtain

$$dy = (\bar{\alpha}dt + \bar{\beta}dW)\left(\frac{dg}{dy}\right)^{-1}$$

or

$$df[x(t)] = \{\alpha[x(t), t]dt + \beta[x(t), t]dW(t)\}f'[x(t)] \tag{4.3.41}$$

which is the same as in ordinary calculus.

ii) **Many Variables.** If a many variable Ito equation is

$$d\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x}, t)dt + \underline{\boldsymbol{B}}(\boldsymbol{x}, t)d\boldsymbol{W}(t) , \tag{4.3.42}$$

then the corresponding Stratonovich equation can be shown similarly to be given by replacing

$$A_i^s = A_i - \tfrac{1}{2}\sum_{j,k} B_{kj}\partial_k B_{ij}$$

$$B_{ij}^s = B_{ij} . \tag{4.3.43}$$

iii) **Fokker-Planck Equation** corresponding to the Stratonovich SDE,

$$(S)\ d\boldsymbol{x} = \boldsymbol{A}^s(\boldsymbol{x}, t)dt + \underline{\boldsymbol{B}}^s(\boldsymbol{x}, t)d\boldsymbol{W}(t) \tag{4.3.44}$$

can, by use of (4.3.43) and the known correspondence (Sect.4.3.5) between the Ito SDE and Fokker-Planck equation, be put in the form

$$\boxed{\partial_t p = -\sum_i \partial_i\{A_i^s p\} + \tfrac{1}{2}\sum_{i,j,k}\partial_i\{B_{ik}^s\partial_j[B_{jk}^s p]\}} \tag{4.3.45}$$

which is often known as the "Stratonovich form" of the Fokker-Planck equation. In contrast to the two forms of the SDEs, the two forms of Fokker-Planck equation have a different appearance but are (of course) interpreted with the same rules —

those of ordinary calculus. We will find later that the Stratonovich form of the Fokker-Planck equation does arise very naturally in certain contexts (Sect.6.6).

iv) **Comparison of the Ito and Stratonovich Integrals.** The Stratonovich integral as defined in (4.3.30) is quite a specialised concept, for it can only be defined in terms of a function $G(z, t)$ of *two* variables. The more "obvious" definition in terms of $G(x[\frac{1}{2}(t_i + t_{i-1})], \frac{1}{2}(t_i + t_{i-1}))$ was not used by Stratonovich in his original definition, although the view that this provides the definition of the Stratonovich integral is widespread in the literature (including the first edition of this book). Apparently, the more obvious definition cannot be proved to converge – see [4.6]. In practise, the precise definition of the Stratonovich integral from first principles is of no great interest, whereas the property that the rule for change of variables is given by ordinary calculus is of great significance, and this is ensured not so much by the definition as by the relations (4.3.37, 43) between the two kinds of integral. One could simply choose to *define* the Stratonovich integral as being given by (4.3.37) when the function obeys the SDE (4.3.31), and this would be mathematically completely satisfactory, and much less confusing.

4.3.7 Dependence on Initial Conditions and Parameters

In exactly the same way as in the case of deterministic differential equations, if the functions which occur in a stochastic differential equation depend continuously on parameters, then the solution normally depends continuously on that parameter. Similarly, the solution depends continuously on the initial conditions. Let us formulate this more precisely. Consider a one-variable equation

$$\left.\begin{aligned} &dx = a(\lambda, x, t)dt + b(\lambda, x, t)dW(t) \\ &\text{with initial condition} \\ &x(t_0) = c(\lambda) \end{aligned}\right\} \tag{4.3.46}$$

where λ is a parameter. Let the solution of (4.3.49) be $x(\lambda, t)$. Suppose

i) $\underset{\lambda \to \lambda_0}{\text{st-lim}}\, c(\lambda) = c(\lambda_0)$;

ii) $\lim_{\lambda \to \lambda_0} \{\sup_{t \in [t_0, T]} [|a(\lambda, x, t) - a(\lambda_0, x, t)| + |b(\lambda, x, t) - b(\lambda_0, x, t)|]\} = 0$;

iii) a K independent of λ exists such that

$$|a(\lambda, x, t)|^2 + |b(\lambda, x, t)|^2 \leqslant K^2(1 + |x^2|) .$$

Then,

$$\underset{\lambda \to \lambda_0}{\text{st-lim}} \{ \sup_{t \in [t_0, T]} |x(\lambda, t) - x(\lambda_0, t)| \} = 0 . \tag{4.3.47}$$

For a proof see [4.1].

Comments

i) Recalling the definition of stochastic limit, the interpretation of the limit (4.3.50) is that as $\lambda \to \lambda_0$, the probability that the maximum deviation over any finite interval $[t_0, T]$ between $x(\lambda, t)$ and $x(\lambda_0, t)$ is nonzero, goes to zero.

ii) Dependence on the initial condition is achieved by letting a and b be independent of λ.

iii) The result will be very useful in justifying perturbation expansions.

iv) Condition (ii) is written in the most natural form for the case that the functions $a(x, t)$ and $b(x, t)$ are not themselves stochastic. It often arises that $a(x, t)$ and $b(x, t)$ are themselves stochastic (nonanticipating) functions. In this case, condition (ii) must be replaced by a probabilistic statement. It is, in fact, sufficient to replace $\lim_{\lambda \to \lambda_0}$ by $\text{st-lim}_{\lambda \to \lambda_0}$.

4.4 Some Examples and Solutions

4.4.1 Coefficients Without x Dependence

The simple equation

$$dx = a(t)dt + b(t)dW(t) \tag{4.4.1}$$

with $a(t)$ and $b(t)$ nonrandom functions of time, is solved simply by integrating

$$x(t) = x_0 + \int_{t_0}^{t} a(t)dt + \int_{t_0}^{t} b(t)dW(t)\,. \tag{4.4.2}$$

Here, x_0 can be either a nonrandom initial condition or may be random, but must be independent of $W(t) - W(t_0)$ for $t > t_0$; otherwise, $x(t)$ is not nonanticipating.

As constructed, $x(t)$ is Gaussian, provided x_0 is either nonrandom or itself Gaussian, since

$$\int_{t_0}^{t} b(t)dW(t)$$

is simply a linear combination of infinitesimal Gaussian variables.Further,

$$\langle x(t)\rangle = \langle x_0\rangle + \int_{t_0}^{t} a(t)dt$$

(since the mean of the Ito integral vanishes)
and

$$\langle [x(t) - \langle x(t)\rangle][x(s) - \langle x(s)\rangle]\rangle \equiv \langle x(t), x(s)\rangle$$

$$= \langle \int_{t_0}^{t} b(t')dW(t') \int_{t_0}^{s} b(s')dW(s')\rangle = \int_{t_0}^{\min(t,s)} [b(t')]^2 dt' ,$$

where we have used the result (4.2.42) with, however,

$$\left.\begin{aligned} G(t') &= b(t') \\ &= 0 \end{aligned}\right\} \begin{aligned} &t' < t \\ &t' \geqslant t \end{aligned}$$

$$\left.\begin{aligned} H(t') &= b(t') \\ &= 0 \end{aligned}\right\} \begin{aligned} &t' < s \\ &t' \geqslant s . \end{aligned}$$

The process is thus completely determined.

4.4.2 Multiplicative Linear White Noise Process

The equation

$$dx = cx\, dW(t) \tag{4.4.3}$$

is known as a multiplicative white noise process because it is linear in x, but the "noise term" $dW(t)$ multiplies x. We can solve this exactly by using Ito's formula Let us define a new variable by

$$y = \log x , \tag{4.4.4}$$

so that

$$dy = \frac{1}{x} dx - \frac{1}{2x^2}(dx)^2$$

$$= c\, dW(t) - \tfrac{1}{2}c^2 dt . \tag{4.4.5}$$

This equation can now be directly integrated, so we obtain

$$y(t) = y(t_0) + c[W(t) - W(t_0)] - \tfrac{1}{2}c^2(t - t_0) \tag{4.4.6}$$

and hence,

$$x(t) = x(t_0) \exp\{c[W(t) - W(t_0)] - \tfrac{1}{2}c^2(t - t_0)\} . \tag{4.4.7}$$

We can calculate the mean by using the formula for any Gaussian variable z with zero mean

$$\langle \exp z\rangle = \exp(\langle z^2\rangle/2) \tag{4.4.8}$$

so that

$$\langle x(t)\rangle = \langle x(t_0)\rangle \exp\left[\tfrac{1}{2}c^2(t-t_0) - \tfrac{1}{2}c^2(t-t_0)\right]$$
$$= \langle x(t_0)\rangle \,. \tag{4.4.9}$$

This result is also obvious from definition, since

$$\langle dx\rangle = \langle cx\, dW(t)\rangle = 0 \quad \text{so that}$$

$$\frac{d\langle x\rangle}{dt} = 0 \,.$$

We can also calculate the autocorrelation function

$$\langle x(t)x(s)\rangle = \langle x(t_0)^2\rangle\langle \exp\{c[W(t) + W(s) - 2W(t_0)] - \tfrac{1}{2}c^2(t+s-2t_0)\}\rangle$$
$$= \langle x(t_0)^2\rangle \exp\{\tfrac{1}{2}c^2[\langle [W(t) + W(s) - 2W(t_0)]^2\rangle - (t+s-2t_0)]\}$$
$$= \langle x(t_0)^2\rangle \exp\{\tfrac{1}{2}c^2[t+s-2t_0 + 2\min(t,s) - (t+s-2t_0)]\}$$
$$= \langle x(t_0)^2\rangle \exp\left[c^2 \min(t-t_0, s-t_0)\right] \,. \tag{4.4.10}$$

Stratonovich Interpretation. The solution of this equation interpreted as a Stratonovich equation can also be obtained, but ordinary calculus would then be valid. Thus, instead of (4.4.5) we would obtain

$$dy = c\, dW(t)$$

and hence,

$$x(t) = x(t_0) \exp\{c[W(t) - W(t_0)]\} \,. \tag{4.4.11}$$

In this case,

$$\langle x(t)\rangle = \langle x(t_0)\rangle \exp\left[\tfrac{1}{2}c^2(t-t_0)\right] \tag{4.4.12}$$

and

$$\langle x(t)x(s)\rangle = \langle x(t_0)^2\rangle \exp\{\tfrac{1}{2}c^2[t+s-2t_0 + 2\min(t-t_0, s-t_0)]\} \,. \tag{4.4.13}$$

One sees that there is a clear difference between these two answers.

4.4.3 Complex Oscillator with Noisy Frequency

This is a simplification of a model due to *Kubo* [4.4] and is a slight generalisation of the previous example for complex variables. We consider

$$\frac{dz}{dt} = \mathrm{i}[\omega + \sqrt{2\gamma}\,\xi(t)]z \tag{4.4.14}$$

which formally represents a simple model of an oscillator with a mean frequency ω perturbed by a noise term $\xi(t)$.

Physically, this is best modelled by writing a Stratonovich equation

$$(S) \quad dz = \mathrm{i}[\omega dt + \sqrt{2\gamma}\, dW(t)]z \tag{4.4.15}$$

which is equivalent to the Ito equation (from Sect. 4.3.6)

$$dz = [(\mathrm{i}\omega - \gamma)dt + \mathrm{i}\sqrt{2\gamma}\, dW(t)]z\,. \tag{4.4.16}$$

Taking the mean value, we see immediately that

$$\frac{d\langle z\rangle}{dt} = (\mathrm{i}\omega - \gamma)\langle z\rangle \tag{4.4.17}$$

with the *damped* oscillatory solution

$$\langle z(t)\rangle = \exp[(\mathrm{i}\omega - \gamma)t]\langle z(0)\rangle\,. \tag{4.4.18}$$

We shall show fully in Sect. 6.6, why the Stratonovich model is more appropriate. The most obvious way to see this is to note that $\xi(t)$ would, in practice, be somewhat smoother than a white noise and ordinary calculus would apply, as is the case in the Stratonovich interpretation.

Now in this case, the correlation function obtained from solving the original Stratonovich equation is

$$\langle z(t)z(s)\rangle = \langle z(0)^2\rangle \exp[(\mathrm{i}\omega - \gamma)(t+s) - 2\gamma\min(t,s)]\,. \tag{4.4.19}$$

In the limit $t, s \to \infty$, with $t + \tau = s$,

$$\lim_{t\to\infty}\langle z(t+\tau)z(t)\rangle = 0\,. \tag{4.4.20}$$

However, the correlation function of physical interest is the complex correlation

$$\begin{aligned}\langle z(t)z^*(s)\rangle &= \langle |z(0)|^2\rangle\langle \exp\{\mathrm{i}\omega(t-s) + \mathrm{i}\sqrt{2\gamma}[W(t) - W(s)]\}\rangle \\ &= \langle |z(0)|^2\rangle \exp\{\mathrm{i}\omega(t-s) - \gamma[t+s-2\min(t,s)]\} \\ &= \langle |z(0)|^2\rangle \exp[\mathrm{i}\omega(t-s) - \gamma|t-s|]\,.\end{aligned} \tag{4.4.21}$$

Thus, the complex correlation function has a damping term which arises purely from the noise. It may be thought of as a noise induced dephasing effect, whereby for large time differences, $z(t)$ and $z^*(s)$ become independent.

A realistic oscillator cannot be described by this model of a complex oscillator, as discussed by *van Kampen* [4.5]. However the qualitative behaviour is very simi-

lar, and this model may be regarded as a prototype model of oscillators with noisy frequency.

4.4.4 Ornstein-Uhlenbeck Process

Taking the Fokker-Planck equation given for the Ornstein-Uhlenbeck process (Sect.3.8.4), we can immediately write down the SDE using the result of Sect.4.3.4:

$$dx = -kx\,dt + \sqrt{D}\,dW(t), \tag{4.4.22}$$

and solve this directly. Putting

$$y = x\,e^{kt}, \tag{4.4.23}$$

then

$$\begin{aligned} dy &= (dx)d(e^{kt}) + (dx)e^{kt} + xd(e^{kt}) \\ &= [-kx\,dt + \sqrt{D}\,dW(t)]k\,e^{kt}dt \\ &\quad + [-kx\,dt + \sqrt{D}\,dW(t)]e^{kt} + kx\,e^{kt}dt\,. \end{aligned} \tag{4.4.24}$$

We note that the first product vanishes, involving only dt^2, and $dW(t)dt$ (in fact, it can be seen that this will always happen if we simply multiply x by a deterministic function of time). We get

$$dy = \sqrt{D}\,e^{kt}dW(t) \tag{4.4.25}$$

so that integrating and resubstituting for y, we get

$$x(t) = x(0)e^{-kt} + \sqrt{D}\int_0^t e^{-k(t-t')}dW(t')\,. \tag{4.4.26}$$

If the initial condition is deterministic or Gaussian distributed, then $x(t)$ is clearly Gaussian with mean and variance

$$\langle x(t)\rangle = \langle x(0)\rangle e^{-kt} \tag{4.4.27}$$

$$\mathrm{var}\{x(t)\} = \langle\{[x(0) - \langle x(0)\rangle]e^{-kt} + \sqrt{D}\int_0^t e^{-k(t-t')}dW(t')\}^2\rangle\,. \tag{4.4.28}$$

Taking the initial condition to be nonanticipating, that is, independent of $dW(t)$ for $t > 0$, we can write using the result of Sect.4.2.6f

$$\mathrm{var}\{x(t)\} = \mathrm{var}\{x(0)\}\,e^{-2kt} + D\int_0^t e^{-2k(t-t')}dt'$$

$$= \{[\text{var}\{x(0)\} - D/2k\}\, e^{-2kt} + D/2k\,. \tag{4.4.29}$$

These equations are the same as those obtained directly by solving the Fokker-Planck equation in Sect.3.8.4, with the added generalisation of a nonanticipating random initial condition. Added to the fact that the solution is a Gaussian variable, we also have the correct conditional probability.

The time correlation function can also be calculated directly and is,

$$\begin{aligned}\langle x(t), x(s)\rangle &= \text{var}\{x(0)\}\, e^{-k(t+s)} + D\Big\langle \int_0^t e^{-k(t-t')}\, dW(t') \int_0^s e^{-k(s-s')}\, dW(s')\Big\rangle \\ &= \text{var}\{x(0)\}\, e^{-k(t+s)} + D\int_0^{\min(t,s)} e^{-k(t+s-2t')}\, dt' \\ &= \left[\text{var}\{x(0)\} - \frac{D}{2k}\right] e^{-k(t+s)} + \frac{D}{2k}\, e^{-k|t-s|}\,. \end{aligned} \tag{4.4.30}$$

Notice that if $k > 0$, as $t, s \to \infty$ with finite $|t - s|$, the correlation function becomes stationary and of the form deduced in Sect.3.8.4.

In fact, if we set the initial time at $-\infty$ rather than 0, the solution (4.4.26) becomes

$$x(t) = \sqrt{D} \int_{-\infty}^{t} e^{-k(t-t)}\, dW(t')\,. \tag{4.4.31}$$

in which the correlation function and the mean obviously assume their stationary values. Since the process is Gaussian, this makes it stationary.

4.4.5 Conversion from Cartesian to Polar Coordinates

A model often used to describe an optical field is given by a pair of Ornstein-Uhlenbeck processes describing the real and imaginary components of the electric field, i.e.,

$$\begin{aligned} dE_1(t) &= -\gamma E_1(t)\, dt + \varepsilon\, dW_1(t) \\ dE_2(t) &= -\gamma E_2(t)\, dt + \varepsilon\, dW_2(t)\,. \end{aligned} \tag{4.4.32}$$

It is of interest to convert to polar coordinates. We set

$$\begin{aligned} E_1(t) &= a(t)\cos\phi(t) \\ E_2(t) &= a(t)\sin\phi(t) \end{aligned} \tag{4.4.33}$$

and for simplicity, also define

$$\mu(t) = \log a(t) \tag{4.4.34}$$

so that

$$\mu(t) + i\phi(t) = \log\,[E_1(t) + iE_2(t)]\,. \tag{4.4.35}$$

We then use the Ito calculus to derive

$$\begin{aligned} d(\mu + i\phi) &= \frac{d(E_1 + iE_2)}{E_1 + iE_2} - \frac{1}{2}\frac{[d(E_1 + iE_2)]^2}{(E_1 + iE_2)^2} \\ &= -\frac{\gamma(E_1 + iE_2)}{E_1 + iE_2}\,dt + \frac{\varepsilon[dW_1(t) + i dW_2(t)]}{(E_1 + iE_2)} \\ &\qquad - \frac{1}{2}\frac{\varepsilon^2[dW_1(t) + i\,dW_2(t)]^2}{(E_1 + iE_2)^2} \end{aligned} \tag{4.4.36}$$

and noting $dW_1(t)dW_2(t) = 0$, $dW_1(t)^2 = dW_2(t)^2 = dt$, it can be seen that the last term vanishes, so we find

$$d[\mu(t) + i\phi(t)] = -\gamma dt + \varepsilon\exp\,[-\mu(t) - i\phi(t)]\,\{dW_1(t) + i\,dW_2(t)\}\,. \tag{4.4.37}$$

We now take the real part, set $a(t) = \exp\,[\mu(t)]$ and using the Ito calculus find

$$da(t) = \{-\gamma a(t) + \tfrac{1}{2}\varepsilon^2/a(t)\}\,dt + \varepsilon\,\{dW_1(t)\cos\phi(t) + dW_2(t)\sin\phi(t)]\}\,. \tag{4.4.38}$$

The imaginary part yields

$$d\phi(t) = [\varepsilon/a(t)]\,[-dW_1(t)\sin\phi(t) + dW_2\cos\phi(t)]\,. \tag{4.4.39}$$

We now define

$$\left.\begin{aligned} dW_a(t) &= dW_1(t)\cos\phi(t) + dW_2(t)\sin\phi(t) \\ dW_\phi(t) &= -\,dW_1(t)\sin\phi(t) + dW_2(t)\cos\phi(t)\,. \end{aligned}\right\} \tag{4.4.40}$$

We note that this is an orthogonal transformation of the kind mentioned in Sect. 4.3.5, so that we may take $dW_a(t)$ and dW_ϕ(t) as increments of independent Wiener processes $W_a(t)$ and $W_\phi(t)$.

Hence, the stochastic differential equations for phase and amplitude are

$$d\phi(t) = [\varepsilon/a(t)]dW_\phi(t) \tag{4.4.41a}$$

$$da(t) = [-\gamma a(t) + \tfrac{1}{2}\varepsilon^2/a(t)]dt + \varepsilon dW_a(t)\,. \tag{4.4.41b}$$

Comment. Using the rules given in Sect. 4.3.6 (ii), it is possible to convert both the Cartesian equation (4.4.32) and the polar equations (4.4.41) to the Stratonovich form, and to find that both are exactly the same as the Ito form. Nevertheless, a direct conversion using ordinary calculus is not possible. Doing so we would get the same result until (4.4.38) where the term $[\frac{1}{2}\varepsilon^2/a(t)]dt$ would not be found. This must be compensated by an extra term which arises from the fact that the Stratonovich increments $dW_i(t)$ are *correlated* with $\phi(t)$ and thus, $dW_a(t)$ and

$dW_\phi(t)$ cannot simply be defined by (4.4.40). We see the advantage of the Ito method which retains the statistical independence of $dW(t)$ and variables evaluated at time t.

Unfortunately, the equations in Polar form are not soluble, as the corresponding Cartesian equations are. There is an advantage, however, in dealing with polar equations in the laser, whose equations are similar, but have an added term proportional to $a(t)^2 dt$ in (4.4.41b).

4.4.6 Multivariate Ornstein-Uhlenbeck Process

we define the process by the SDE

$$d\boldsymbol{x}(t) = -A\boldsymbol{x}(t)dt + B\, d\boldsymbol{W}(t), \tag{4.4.42}$$

(A and B are constant matrices) for which the solution is easily obtained (as in Sect. 4.4.4):

$$\boldsymbol{x}(t) = \exp(-At)\boldsymbol{x}(0) + \int_0^t \exp[-A(t-t')]B\, d\boldsymbol{W}(t') . \tag{4.4.43}$$

The mean is

$$\langle \boldsymbol{x}(t) \rangle = \exp(-At)\langle \boldsymbol{x}(0) \rangle . \tag{4.4.44}$$

The correlation function follows similarly

$$\begin{aligned} \langle \boldsymbol{x}(t), \boldsymbol{x}^{\mathrm{T}}(s) \rangle &\equiv \langle [\boldsymbol{x}(t) - \langle \boldsymbol{x}(t) \rangle][\boldsymbol{x}(s) - \langle \boldsymbol{x}(s) \rangle]^{\mathrm{T}} \rangle \\ &= \exp(-At)\langle \boldsymbol{x}(0), \boldsymbol{x}^{\mathrm{T}}(0) \rangle \exp(-As) \\ &\quad + \int_0^{\min(t,s)} \exp[-A(t-t')]BB^{\mathrm{T}} \exp[-A^{\mathrm{T}}(s-t')]dt' . \end{aligned} \tag{4.4.45}$$

The integral can be explicitly evaluated in certain special cases, and for particular low-dimensional problems, it is possible to simply multiply everything out term by term. In the remainder we set $\langle \boldsymbol{x}(0), \boldsymbol{x}^{\mathrm{T}}(0) \rangle = 0$, corresponding to a deterministic initial condition, and evaluate a few special cases.

a) Suppose $AA^{\mathrm{T}} = A^{\mathrm{T}}A$

Then we can find a unitary matrix S such that

$$\begin{aligned} &SS^{+} = 1 \\ &SAS^{+} = SA^{\mathrm{T}}S^{+} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots \lambda_n) . \end{aligned} \tag{4.4.46}$$

For simplicity, assume $t \geqslant s$. Then

$$\langle \boldsymbol{x}(t), \boldsymbol{x}^{\mathrm{T}}(s) \rangle = S^{+}G(t, s)S ,$$

where

$$[G(t,s)]_{ij} = \frac{(BB^{\mathrm{T}})_{ij}}{\lambda_i + \lambda_j}[\exp(-\lambda_i|t-s|) - \exp(-\lambda_i t - \lambda_j s)] . \tag{4.4.47}$$

b) Variance in Stationary Solution

If A has only eigenvalues with positive real part, a stationary solution exists of the form

$$\boldsymbol{x}_s(t) = \int_{-\infty}^{t} \exp[-A(t-t')]B\, d\boldsymbol{W}(t) . \tag{4.4.48}$$

We have of course

$$\left.\begin{aligned} &\langle \boldsymbol{x}_s(t)\rangle = 0 \\ &\text{and} \\ &\langle \boldsymbol{x}_s(t), \boldsymbol{x}_s^{\mathrm{T}}(s)\rangle = \int_{-\infty}^{\min(t,s)} \exp[-A(t-t')]BB^{\mathrm{T}}\exp[-A^{\mathrm{T}}(s-t')]dt' . \end{aligned}\right\} \tag{4.4.49}$$

Let us define the stationary covariance matrix σ by

$$\sigma = \langle \boldsymbol{x}_s(t), \boldsymbol{x}_s^{\mathrm{T}}(t)\rangle . \tag{4.4.50}$$

Then the evaluation of this quantity can be achieved algebraically for

$$\begin{aligned} A\sigma + \sigma A^{\mathrm{T}} &= \int_{-\infty}^{t} A\exp[-A(t-t')]BB^{\mathrm{T}}\exp[-A^{\mathrm{T}}(t-t')]dt' \\ &\quad + \int_{-\infty}^{t} \exp[-A(t-t')]BB^{\mathrm{T}}\exp[-A^{\mathrm{T}}(t-t')]A^{\mathrm{T}}dt' \\ &= \int_{-\infty}^{t} \frac{d}{dt'}\{\exp[-A(t-t')]BB^{\mathrm{T}}\exp[-A(t-t')]\}\,dt' . \end{aligned}$$

Carrying out the integral, we find that the lower limit vanishes by the assumed positivity of the eigenvalues of A and hence only the upper limit remains, giving

$$\boxed{A\sigma + \sigma A^{\mathrm{T}} = BB^{\mathrm{T}}} \tag{4.4.51}$$

as an algebraic equation for the stationary covariance matrix.

c) Stationary Variance for Two Dimensions

We note that if A is a 2×2 matrix, it satisfies the characteristic equation

$$A^2 - (\mathrm{Tr}\ A)A + (\mathrm{Det}\ A) = 0 \tag{4.4.52}$$

and from (4.4.49) and the fact that (4.4.52) implies $\exp(-At)$ is a polynomial of degree 1 in A, we must be able to write

$$\sigma = \alpha BB^{\mathrm{T}} + \beta(ABB^{\mathrm{T}} + BB^{\mathrm{T}}A^{\mathrm{T}}) + \gamma ABB^{\mathrm{T}}A^{\mathrm{T}} .$$

Using (4.4.52), we find (4.4.51) is satisfied if

$$\alpha + (\mathrm{Tr}\ A)\beta - (\mathrm{Det}\ A)\gamma = 0$$
$$2\beta(\mathrm{Det}\ A) + 1 = 0$$
$$\beta + (\mathrm{Tr}\ A)\gamma = 0\,.$$

From which,

$$\boxed{\sigma = \frac{(\mathrm{Det}\ A)BB^{\mathrm{T}} + [A - (\mathrm{Tr}\ A)1]BB^{\mathrm{T}}[A - (\mathrm{Tr}\ A)1]^{\mathrm{T}}}{2\ (\mathrm{Tr}\ A)\ (\mathrm{Det}\ A)}\,.} \tag{4.4.53}$$

d) Time Correlation Matrix in the Stationary Sate

From the solution of (4.4.49), we see that if $t > s$,

$$\langle \boldsymbol{x}_s(t), \boldsymbol{x}_s^{\mathrm{T}}(s)\rangle = \exp[-A(t-s)] \int_{-\infty}^{s} \exp[-A(s-t')]BB^{\mathrm{T}}\exp[-A^{\mathrm{T}}(s-t')]dt'$$
$$= \exp[-A(t-s)]\sigma \qquad t > s \tag{4.4.54a}$$

and similarly,

$$= \sigma \exp[-A^{\mathrm{T}}(s-t)] \qquad t < s \tag{4.4.54b}$$

which gives the dependence on $s - t$ as expected of a stationary solution. Defining then

$$G_s(t-s) = \langle \boldsymbol{x}_s(t), \boldsymbol{x}_s^{\mathrm{T}}(s)\rangle\,, \tag{4.4.55}$$

we see (remembering $\sigma = \sigma^{\mathrm{T}}$) that

$$G_s(t-s) = [G_s(s-t)]^{\mathrm{T}}\,. \tag{4.4.56}$$

e) Spectrum Matrix in Stationary State

The spectrum matrix turns out to be rather simple. We define similarly to Sect. 1.4.2:

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}\omega\tau}G_s(\tau)d\tau \tag{4.4.57}$$
$$= \frac{1}{2\pi}\left\{\int_0^{\infty} \exp[-(\mathrm{i}\omega + A)\tau]\sigma\, d\tau + \int_{-\infty}^{0} \sigma \exp[(-\mathrm{i}\omega + A^{\mathrm{T}})\tau]\,d\tau\right\}$$
$$= \frac{1}{2\pi}[(A + \mathrm{i}\omega)^{-1}\sigma + \sigma(A^{\mathrm{T}} - \mathrm{i}\omega)^{-1}]\,.$$

Hence, $(A + \mathrm{i}\omega)S(\omega)(A^{\mathrm{T}} - \mathrm{i}\omega) = \frac{1}{2\pi}(\sigma A^{\mathrm{T}} + A\sigma)$ and using (4.4.51), we get

$$\boxed{S(\omega) = \frac{1}{2\pi}(A + \mathrm{i}\omega)^{-1}BB^{\mathrm{T}}(A^{\mathrm{T}} - \mathrm{i}\omega)^{-1}\,.} \tag{4.4.58}$$

f) Regression Theorem

The result (4.4.54a) is also known as a regression theorem in that it states that the time development $G_s(\tau)$ is for $\tau > 0$ governed by the same law of time development of the mean [e.g.,(4.4.44)]. It is a consequence of the Markovian linear nature of the problem. For

$$\frac{d}{d\tau}[G_s(\tau)]d\tau = \frac{d}{d\tau}\langle \boldsymbol{x}_s(\tau), \boldsymbol{x}_s^{\mathrm{T}}(0)\rangle d\tau$$
$$= \langle [-A\boldsymbol{x}_s(\tau)d\tau + B\,d\boldsymbol{W}(\tau)], \boldsymbol{x}_s^{\mathrm{T}}(0)\rangle \tag{4.4.59}$$

and since $\tau > 0$, $d\boldsymbol{W}(\tau)$ is uncorrelated with $\boldsymbol{x}_s^{\mathrm{T}}(0)$, so

$$\boxed{\frac{d}{d\tau}[G_s(\tau)] = -A\,G_s(\tau)\,.} \tag{4.4.59}$$

Thus, computation of $G_s(\tau)$ requires the knowledge of $G_s(0) = \sigma$ and the time development equation of the mean. This result is similar to those of Sect.3.7.4.

4.4.7 The General Single Variable Linear Equation

a) Homogeneous Case

We consider firstly the homogeneous case

$$dx = [b(t)dt + g(t)dW(t)]x \tag{4.4.60}$$

and using the usual Ito rules, write

$$y = \log x \tag{4.4.61}$$

so that

$$dy = dx/x - \tfrac{1}{2}(dx)^2/x^2$$
$$= [b(t)dt + g(t)dW(t)] - \tfrac{1}{2}g(t)^2dt, \tag{4.4.62}$$

and integrating and inverting (4.4.61), we get

$$x(t) = x(0)\exp\left\{\int_0^t [b(t') - \tfrac{1}{2}g(t')^2]dt' + \int_0^t g(t')dW(t')\right\} \tag{4.4.63}$$
$$\equiv x(0)\phi(t) \tag{4.4.64}$$

which serves to define $\phi(t)$.

We note that [using (4.4.8)]

$$\langle [x(t)]^n\rangle = \langle [x(0)]^n\rangle \left\langle \exp\left\{n\int_0^t [b(t') - \tfrac{1}{2}g(t')^2]dt' + n\int_0^t g(t')dW(t')\right\}\right\rangle$$
$$= \langle [x(0)]^n\rangle \exp\left\{n\int b(t')dt' + \tfrac{1}{2}n(n-1)\int g(t')^2dt'\right\}. \tag{4.4.65}$$

b) Inhomogeneous Case

Now consider

$$dx = [a(t) + b(t)x]dt + [f(t) + g(t)x]dW(t) \tag{4.4.66}$$

and write

$$z(t) = x(t)[\phi(t)]^{-1} \tag{4.4.67}$$

with $\phi(t)$ as defined in (4.4.64) and a solution of the homogeneous equation (4.4.60). Then we write

$$dz = dx[\phi(t)]^{-1} + x\, d[\phi(t)^{-1}] + dx\, d[\phi(t)^{-1}] .$$

Noting that $d[\phi(t)]^{-1} = -d\phi(t)[\phi(t)]^{-2} + [d\phi(t)]^2[\phi(t)]^{-3}$ and using Ito rules, we find

$$dz = \{[a(t) - f(t)g(t)]dt + f(t)dW(t)\}\,\phi(t)^{-1} \tag{4.4.68}$$

which is directly integrable. Hence, the solution is

$$x(t) = \phi(t)\left\{x(0) + \int_0^t \phi(t')^{-1}\{[a(t') - f(t')g(t')]\, dt' + f(t')dW(t')\}\right\} . \tag{4.4.69}$$

c) Moments and Autocorrelation

It is better to derive equations for the moments from (4.4.66) rather than calculate moments and autocorrelation directly from the solution (4.4.69).

For we have

$$\begin{aligned} d[x(t)^n] &= nx(t)^{n-1}dx(t) + \frac{n(n-1)}{2}x(t)^{n-2}[dx(t)]^2 \\ &= nx(t)^{n-1}dx(t) + \frac{n(n-1)}{2}x(t)^{n-2}[f(t) + g(t)x(t)]^2dt . \end{aligned} \tag{4.4.70}$$

Hence,

$$\boxed{\begin{aligned} \frac{d}{dt}\langle x(t)^n\rangle = \langle x(t)^n\rangle &\left[nb(t) + \frac{n(n-1)}{2}g(t)^2\right] \\ &+ \langle x(t)^{n-1}\rangle[na(t) + n(n-1)f(t)g(t)] \\ &+ \langle x(t)^{n-2}\rangle\left[\frac{n(n-1)}{2}f(t)^2\right] . \end{aligned}} \tag{4.4.71}$$

These equations from a hierarchy in which the nth equation involves the solutions of the previous two, and can be integrated successively.

4.4.8 Multivariable Linear Equations

a) Homogeneous Case

The equation is

$$d\boldsymbol{x}(t) = [B(t)dt + \sum_i G_i(t)dW_i(t)]\boldsymbol{x}(t)\,, \tag{4.4.72}$$

where $B(t)$, $G_i(t)$ are matrices. The equation is not, in general, soluble in closed form unless all the matrices $B(t)$, $G_i(t')$ commute at all times with each other, i.e.

$$\begin{aligned} &G_i(t)G_j(t') = G_j(t')G_i(t) \\ &B(t)G_i(t') = G_i(t')B(t) \\ &B(t)B(t') = B(t')B(t) \end{aligned} \tag{4.4.73}$$

In this case, the solution is completely analogous to the one variable case and we have

$$\boldsymbol{x}(t) = \Phi(t)\boldsymbol{x}(0)$$

with

$$\Phi(t) = \exp\left\{\int_0^t [B(t) - \tfrac{1}{2}\sum_i G_i(t)^2]dt + \int_0^t \sum_i G_i(t)dW_i(t)\right\}. \tag{4.4.74}$$

b) Inhomogeneous Case

We can reduce the inhomogeneous case to the homogeneous case in exactly the same way as in one dimension. Thus, we consider

$$d\boldsymbol{x}(t) = [\boldsymbol{A}(t) + B(t)\boldsymbol{x}]dt + \sum_i [\boldsymbol{F}_i(t) + G_i(t)\boldsymbol{x}]dW(t) \tag{4.4.75}$$

and write

$$\boldsymbol{y}(t) = \psi(t)^{-1}\boldsymbol{x}(t)\,, \tag{4.4.76}$$

where $\psi(t)$ is a matrix solution of the homogeneous equation (4.4.72). We first have to evaluate $d[\psi^{-1}]$. For any matrix M we have $MM^{-1} = 1$, so, expanding to second order, $Md[M^{-1}] + dM\,M^{-1} + dM\,d[M^{-1}] = 0$.
Hence, $d[M^{-1}] = -[M + dM]^{-1}\,dM\,M^{-1}$ and again to second order

$$d[M^{-1}] = -M^{-1}dM\,M^{-1} + M^{-1}dM\,M^{-1}dM\,M^{-1} \tag{4.4.77}$$

and thus, since $\psi(t)$ satisfies the homogeneous equation,

$$d[\psi(t)^{-1}] = \psi(t)^{-1}\{[-B(t) + \sum_i G_i(t)^2]dt - \sum_i G_i(t)dW_i(t)\}\,.$$

and, again taking differentials

$$d\boldsymbol{y}(t) = \psi(t)^{-1}\{[\boldsymbol{A}(t) - \sum_i \boldsymbol{G}_i(t)\boldsymbol{F}_i(t)]dt + \sum_i \boldsymbol{F}_i(t)dW_i(t)\}.$$

Hence,

$$\boldsymbol{x}(t) = \psi(t)\{\boldsymbol{x}(0) + \int_0^t \psi(t')^{-1}\{[\boldsymbol{A}(t') - \sum_i \boldsymbol{G}_i(t')\boldsymbol{F}_i(t')]dt' + \sum \boldsymbol{F}_i(t')dW_i(t')\}\}. \tag{4.4.78}$$

This solution is not very useful for practical purposes, even when the solution for the homogeneous equation is known, because of the difficulty in evaluating means and correlation functions.

4.4.9 Time-Dependent Ornstein-Uhlenbeck Process

This is a particular case of the previous general linear equation which is soluble. It is a generalisation of the multivariate Ornstein-Uhlenbeck process (Sect.4.4.6) to include time-dependent parameters, namely,

$$d\boldsymbol{x}(t) = -\boldsymbol{A}(t)\boldsymbol{x}(t)dt + \boldsymbol{B}(t)d\boldsymbol{W}(t). \tag{4.4.79}$$

This is clearly of the same form as (4.4.75) with the replacements

$$\begin{aligned} &\boldsymbol{A}(t) \to 0 \\ &\boldsymbol{B}(t) \to -\boldsymbol{A}(t) \\ &\sum_i \boldsymbol{F}_i(t)dW_i(t) \to \boldsymbol{B}(t)d\boldsymbol{W}(t) \\ &\boldsymbol{G}_i(t) \to 0. \end{aligned} \tag{4.4.80}$$

The corresponding homogeneous equation is simply the deterministic equation

$$d\boldsymbol{x}(t) = -\boldsymbol{A}(t)\boldsymbol{x}(t)\,dt \tag{4.4.81}$$

which is soluble provided $\boldsymbol{A}(t)\boldsymbol{A}(t') = \boldsymbol{A}(t')\boldsymbol{A}(t)$ and has the solution

$$\boldsymbol{x}(t) = \psi(t)\boldsymbol{x}(0)$$

with

$$\psi(t) = \exp[-\int_0^t \boldsymbol{A}(t')dt']. \tag{4.4.82}$$

Thus, applying (4.4.78),

$$\boldsymbol{x}(t) = \exp[-\int_0^t \boldsymbol{A}(t')dt']\boldsymbol{x}(0) + \int_0^t \{\exp[-\int_{t'}^t \boldsymbol{A}(s)ds]\}\,\boldsymbol{B}(t')d\boldsymbol{W}(t'). \tag{4.4.83}$$

This is very similar to the solution of the time-independent Ornstein-Uhlenbeck process, as derived in Sect. 4.4.6 (4.4.43).

From this we have

$$\langle \boldsymbol{x}(t) \rangle = \exp\left[-\int_0^t A(t')dt'\right]\langle \boldsymbol{x}(0) \rangle \tag{4.4.84}$$

$$\langle \boldsymbol{x}(t), \boldsymbol{x}^{\mathrm{T}}(t) \rangle = \exp\left[-\int_0^t A(t')dt'\right]\langle \boldsymbol{x}(0), \boldsymbol{x}(0)^{\mathrm{T}} \rangle \exp\left[-\int_0^t A^{\mathrm{T}}(t')dt'\right]$$
$$+ \int_0^t dt' \exp\left[-\int_{t'}^t A(s)ds\right] B(t')B^{\mathrm{T}}(t') \exp\left[-\int_{t'}^t A^{\mathrm{T}}(s)ds\right]. \tag{4.4.85}$$

The time-dependent Ornstein-Uhlenbeck process will arise very naturally in connection with the development of asymptotic methods in low-noise systems.

5. The Fokker-Planck Equation

In this rather long chapter, the theory of continuous Markov processes is developed from the point of view of the corresponding Fokker-Planck equation, which gives the time evolution of the probability density function for the system. The chapter divides into two main subjects—single variable and multivariable processes. There are a large number of exact results for single variable systems, which makes the separate treatment of such systems appropriate. Thus Sect. 5.2 and its subsections treat all aspects of one variable systems, whereas Sect. 5.3 deals with multivariable systems. However, the construction of appropriate boundary conditions is of fundamental importance in both cases, and is carried out in general in Sect. 5.2.1. A corresponding treatment for the boundary conditions on the backward Fokker-Planck equation is given in Sect. 5.2.4. The remaining subsections of Sect. 5.2 are devoted to a range of exact results, on stationary distribution functions, properties of eigenfunctions, and exit problems, most of which can be explicitly solved in the one variable case.

Section 5.3 and its subsections explore exact results for many variable systems. These results are not as explicit as for the one variable case. An extra feature which is included is the concept of *detailed balance* in multivariable systems, which is almost trivial in one variable systems, but leads to very interesting conclusions in multivariable systems.

The chapter concludes with a treatment of exact results in exit problems for multivariable Fokker-Planck equations.

5.1 Background

We have already met the Fokker-Planck equation in several contexts, starting from Einstein's original derivation and use of the diffusion equation (Sect.1.2), again as a particular case of the differential Chapman-Kolmogorov equation (Sect.3.5.2), and finally, in connection with stochastic differential equations (Sect.4.3.4). There are many techniques associated with the use of Fokker-Planck equations which lead to results more directly than by direct use of the corresponding stochastic differential equation; the reverse is also true. To obtain a full picture of the nature of diffusion processes, one must study both points of view.

The origin of the name "Fokker-Planck Equation" is from the work of *Fokker* (1914) [5.1] and *Planck* (1917) [5.2] where the former investigated Brownian motion in a radiation field and the latter attempted to build a complete theory of fluctuations based on it. Mathematically oriented works tend to use the term "Kolmogorov's Equation" because of Kolmogorov's work in developing its rigorous basis

[5.3]. Yet others use the term "Smoluchowski Equation" because of Smoluchowski's original use of this equation. Without in any way assessing the merits of this terminology, I shall use the term "Fokker-Planck equation" as that most commonly used by the audience to which this book is addressed.

5.2 Fokker-Planck Equation in One Dimension

In one dimension, the Fokker-Planck equation (FPE) takes the simple form

$$\frac{\partial f(x,t)}{\partial t} = -\frac{\partial}{\partial x}[A(x,t)f(x,t)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[B(x,t)f(x,t)] . \tag{5.2.1}$$

In Sects.3.4,5, the FPE was shown to be valid for the conditional probability, that is, the choice

$$f(x,t) = p(x,t|x_0,t_0) \tag{5.2.2}$$

for any initial x_0, t_0, and with the initial condition

$$p(x,t_0|x_0,t_0) = \delta(x-x_0) . \tag{5.2.3}$$

However, using the definition for the one time probability

$$p(x,t) = \int dx_0\, p(x,t;x_0,t_0) = \int dx_0\, p(x,t|x_0,t_0)p(x_0,t_0) , \tag{5.2.4}$$

we see that it is also valid for $p(x,t)$ with the initial condition

$$p(x,t)|_{t=t_0} = p(x,t_0) , \tag{5.2.5}$$

which is generally less singular than (5.2.3).

From the result of Sect.4.3.4, we know that the stochastic process described by a conditional probability satisfying the FPE is equivalent to the Ito stochastic differential equation (SDE)

$$dx(t) = A[x(t),t]dt + \sqrt{B[x(t),t]}dW(t)$$

and that the two descriptions are to be regarded as complementary to each other. We will see that perturbation theories based on the FPE are very different from those based on the SDE and both have their uses.

5.2.1 Boundary Conditions

The FPE is a second-order parabolic partial differential equation, and for solutions we need an initial condition such as (5.2.5) and boundary conditions at the end of the interval inside which x is constrained. These take on a variety of forms.

It is simpler to derive the boundary conditions in general, than to restrict consideration to the one variable situation. We consider the forward equation

$$\partial_t p(\boldsymbol{z}, t) = -\sum_i \frac{\partial}{\partial z_i} A_i(\boldsymbol{z}, t) p(\boldsymbol{z}, t) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial z_i \partial z_j} B_{ij}(\boldsymbol{z}, t) p(\boldsymbol{z}, t) \tag{5.2.6}$$

We note that this can also be written

$$\frac{\partial p(\boldsymbol{z}, t)}{\partial t} + \sum_i \frac{\partial}{\partial z_i} J_i(\boldsymbol{z}, t) = 0 \tag{5.2.7}$$

where we define the *probability current*

$$J_i(\boldsymbol{z}, t) = A_i(\boldsymbol{z}, t) p(\boldsymbol{z}, t) - \frac{1}{2} \sum_j \frac{\partial}{\partial z_j} B_{ij}(\boldsymbol{z}, t) p(\boldsymbol{z}, t) \tag{5.2.8}$$

Equation (5.2.7) has the form of a local conservation equation, and can be written in an integral form as follows. Consider some region R with a boundary $\boldsymbol{S}$ and define

$$P(R, t) = \int_R d\boldsymbol{z}\, p(\boldsymbol{z}, t)$$

then (5.2.7) is equivalent to

$$\frac{\partial P(R, t)}{\partial t} = -\int_S dS\, \boldsymbol{n} \cdot \boldsymbol{J}(\boldsymbol{z}, t) \tag{5.2.9}$$

where $\boldsymbol{n}$ is the outward pointing normal to $\boldsymbol{S}$. Thus (5.2.9) indicates that the total loss of probability is given by the surface integral of $\boldsymbol{J}$ over the boundary of R. We can show as well that the current $\boldsymbol{J}$ does have the somewhat stronger property, that a surface integral over any surface $\boldsymbol{S}$ gives the net flow of probability across that surface.

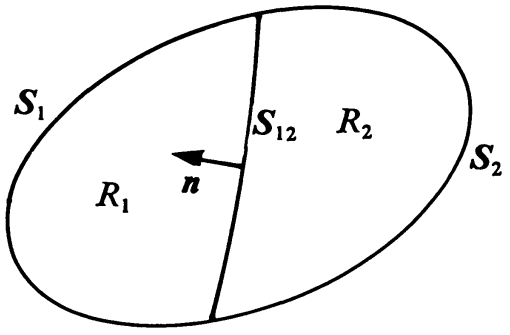


Fig. 5.1. Regions used to demonstrate that the probability current is the flow of probability

For consider two adjacent regions R_1 and R_2, separated by a surface $\boldsymbol{S}_{12}$. Let $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ be the surfaces which, together with $\boldsymbol{S}_{12}$, enclose respectively R_1 and R_2 (see Fig. 5.1).

Then the net flow of probability can be computed by noting that we are dealing here with a process with continuous sample paths, so that, in a sufficiently short

time Δt, the probability of crossing S_{12} from R_2 to R_1 is the joint probability of being in R_2 at time t and R_1 at time $t + \Delta t$,

$$= \int_{R_1} dx \int_{R_2} dy\, p(x, t + \Delta t; y, t).$$

The *net flow* of probability from R_2 to R_1 is obtained by subtracting from this the probability of crossing in the reverse direction, and dividing by Δt; i.e.

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{R_1} dx \int_{R_2} dy\, [p(x, t + \Delta t; y, t) - p(y, t + \Delta t; x, t)]. \tag{5.2.10}$$

Note that

$$\int_{R_1} dx \int_{R_2} dy\, p(x, t; y, t) = 0$$

since this is the probability of being in R_1 and R_2 simultaneously. Thus, we can write

$$(5.2.10) = \int_{R_1} dx \int_{R_2} dy\, [\partial_{t'} p(x, t'; y, t) - \partial_{t'} p(y, t'; x, t)]_{t'=t}$$

and using the Fokker-Planck equation in the form (5.2.7)

$$= -\int_{R_1} dx \sum_i \frac{\partial}{\partial x_i} J_i(x, t; R_2, t) + \int_{R_2} dy \sum_i \frac{\partial}{\partial y_i} J_i(y, t; R_1, t) \tag{5.2.11}$$

where $J_i(x, t; R_2, t)$ is formed from

$$p(x, t; R_2, t) = \int_{R_2} dy\, p(x, t; y, t)$$

in the same way as $J(z, t)$ is formed from $p(z, t)$ in (5.2.8) and $J_i(y, t; R_1, t)$ is defined similarly. We now convert the integrals to surface integrals. The integral over S_2 vanishes, since it will involve $p(x, t; R_2, t)$, with x not in R_2 or on its boundary (except for a set of measure zero.) Similarly the integral over S_1 vanishes, but those over S_{12} do not, since here the integration is simply over part of the boundaries of R_1 and R_2.

Thus we find, the net flow from R_2 to R_1 is

$$\int_{S_{12}} dS\, n \cdot \{J(x, t; R_1, t) + J(x, t; R_2, t)\}$$

and we finally conclude, since x belongs the union of R_1 and R_2, that the net flow of probability per unit time from R_2 to R_1

$$\equiv \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{R_1} dx \int_{R_2} dy\, [p(x, t + \Delta t; y, t) - p(y, t + \Delta t; x, t)] = \int_{S_{12}} dS\, n \cdot J(x, t)$$

where n points from R_2 to R_1 (5.2.12)

We can now consider the various boundary conditions separately.

a) Reflecting Barrier

We can consider the situation where the particle cannot leave a region R, hence there is zero net flow of probability across S, the boundary of R. Thus we require

$$\boldsymbol{n}\cdot\boldsymbol{J}(\boldsymbol{z}, t) = 0 \quad \text{for } \boldsymbol{z} \in \boldsymbol{S}, \quad \boldsymbol{n} = \text{normal to } \boldsymbol{S} \tag{5.2.13}$$

where $\boldsymbol{J}(\boldsymbol{z}, t)$ is given by (5.2.8).

Since the particle cannot cross S, it must be reflected there, and hence the name *reflecting barrjer* for this condition.

b) Absorbing Barrier

Here, one assumes that the moment the particle reaches S, it is removed from the system, thus the barrier absorbs. Consequently, the probability of being on the boundary is zero, i.e.

$$p(\boldsymbol{z}, t) = 0 \qquad \text{for } \boldsymbol{z} \in \boldsymbol{S} \tag{5.2.14}$$

c) Boundary Conditions at a Discontinuity

It is possible for both the A_i and B_{ij} coefficients to be discontinuous at a surface S, but for there to be free motion across S. Consequently, the probability and the normal component of the current must both be continuous across S,

$$\boldsymbol{n}\cdot\boldsymbol{J}(\boldsymbol{z})|_{S_+} = \boldsymbol{n}\cdot\boldsymbol{J}(\boldsymbol{z})|_{S_-} \tag{5.2.15}$$

$$p(\boldsymbol{z})|_{S_+} = p(\boldsymbol{z})|_{S_-} \tag{5.2.16}$$

where S_+, S_-, as subscripts, mean the limits of the quantities from the left and right hand sides of the surface.

The definition (5.2.8) of the current, indicates that the derivatives of $p(\boldsymbol{z})$ are not necessarily continuous at S.

d) Periodic Boundary Condition

We assume that the process takes place on an interval $[a, b]$ in which the two end points are identified with each other. (this occurs, for example, if the diffusion is on a circle). Then we impose boundary conditions derived from those for a discontinuity, i.e.,

$$\text{I: } \lim_{x\to b-} p(x, t) = \lim_{x\to a+} p(x, t) \tag{5.2.17}$$

$$\text{II: } \lim_{x\to b-} J(x, t) = \lim_{x\to a+} J(x, t). \tag{5.2.18}$$

Most frequently, periodic boundary conditions are imposed when the functions $A(x, t)$ and $B(x, t)$ are periodic on the same interval so that we have

$$\begin{aligned} A(b, t) &= A(a, t) \\ B(b, t) &= B(a, t) \end{aligned} \tag{5.2.19}$$

and this means that I and II simply reduce to an equality of $p(x, t)$ and its derivatives at the points a and b.

e) Prescribed Boundaries

If the diffusion coefficient vanishes at a boundary, we have a situation in which the kind of boundary may be automatically prescribed. Suppose the motion occurs only for $x > a$. If a Lipschitz condition is obeyed by $A(x, t)$ and $\sqrt{B(x, t)}$ at $x = a$ (Sect. 4.3.1) and $B(x, t)$ is differentiable at $x = a$ then

$$\partial_x B(a, t) = 0. \tag{5.2.20}$$

The SDE then has solutions, and we may write

$$dx(t) = A(x, t)\, dt + \sqrt{B(x, t)}\, dW(t) \tag{5.2.21}$$

In this rather special case, the situation is determined by the sign of $A(x, t)$. Three cases then occur, as follows.

i) **Exit boundary.** In this case, we suppose

$$A(a, t) < 0 \tag{5.2.22}$$

so that if the particle reaches the point a, it will certainly proceed out of region to $x < a$. Hence the name "exit boundary"

ii) **Entrance boundary.** Suppose

$$A(a, t) > 0. \tag{5.2.23}$$

In this case, if the particle reaches the point a, the sign of $A(a, t)$ is such as to return it to $x > a$; thus a particle placed to the right of a can never leave the region. However, a particle introduced at $x = a$ will certainly enter the region. Hence the name, "entrance boundary".

iii) **Natural boundary.** Finally consider

$$A(a, t) = 0. \tag{5.2.24}$$

The particle, once it reaches $x = a$, will remain there. However it can be demonstrated that it cannot ever reach this point. This is a boundary from which we can neither absorb nor at which we can introduce any particles.

Feller [5.4] has shown that in general the boundaries can be assigned to one of the four types; regular, entrance, exit and natural. His general criteria for the classification of these boundaries are as follows. Define

$$f(x) = \exp\left[-2\int_{x_0}^{x} ds\, A(s)/B(s)\right] \tag{5.2.25}$$

$$g(x) = 2/[B(x)f(x)] \tag{5.2.26}$$

$$h_1(x) = f(x) \int_{x_0}^{x} g(s)\, ds \tag{5.2.27}$$

$$h_2(x) = g(x) \int_{x_0}^{x} f(s)\, ds\,. \tag{5.2.28}$$

Here $x_0 \in (a, b)$, and is fixed. Denote by

$$\mathcal{L}(x_1, x_2) \tag{5.2.29}$$

the space of all functions integrable on the interval (x_1, x_2); then the boundary at a can be classified as

---

I: Regular: if $f(x) \in \mathcal{L}(a, x_0)$, and $g(x) \in \mathcal{L}(a, x_0)$

II: Exit: if $g(x) \notin \mathcal{L}(a, x_0)$, and $h_1(x) \in \mathcal{L}(a, x_0)$

III: Entrance: if $g(x) \in \mathcal{L}(a, x_0)$, and $h_2(x) \in \mathcal{L}(a, x_0)$

IV: Natural : all other cases.

It can be seen from the results of Sect. 5.2.2 that for an exit boundary there is no normalisable stationary solution of the FPE, and that the mean time to reach the boundary, (5.2.161), is finite. Similarly, if the boundary is exit, a stationary solution can exist, but the mean time to reach the boundary is infinite. In the case of a regular boundary, the mean time to reach the boundary is finite, but a stationary solution with a reflecting boundary at a does exist. The case of natural boundaries is harder to analyse. The reader is referred to [5.5] for a more complete description.

f) Boundaries at Infinity

All of the above kinds of boundary can occur at infinity, provided we can simultaneously guarantee the normalisation of the probability which, if $p(x)$ is reasonably well behaved, requires

$$\lim_{x \to \infty} p(x, t) = 0\,. \tag{5.2.30}$$

If $\partial_x p(x)$ is reasonably well behaved (i.e., does not oscillate infinitely rapidly as $x \to \infty$),

$$\lim_{x \to \infty} \partial_x p(x, t) = 0 \tag{5.2.31}$$

so that a nonzero current at infinity will usually require either $A(x, t)$ or $B(x, t)$ to become infinite there. Treatment of such cases is usually best carried out by changing to another variable which is finite at $x = \infty$.

Where there are boundaries at $x = \pm\infty$ and nonzero currents at infinity are permitted, we have two possibilities which do not allow for loss of probability:

i) $J(\pm\infty, t) = 0$ (5.2.32)

ii) $J(+\infty, t) = J(-\infty, t)$. (5.2.33)

These are the limits of reflecting and periodic boundary conditions, respectively.

5.2.2 Stationary Solutions for Homogeneous Fokker-Planck Equations

We recall (Sect.3.7.2) that in a homogeneous process, the drift and diffusion coefficients are time independent. In such a case, the equation satisfied by the stationary distribution is

$$\frac{d}{dx}[A(x)p_s(x)] - \frac{1}{2}\frac{d^2}{dx^2}[B(x)p_s(x)] = 0 \tag{5.2.34}$$

which can also be written simply in terms of the current (as defined in Sect.5.2.1)

$$\frac{dJ(x)}{dx} = 0 \tag{5.2.35}$$

which clearly has the solution

$$J(x) = \text{constant}. \tag{5.2.36}$$

Suppose the process takes place on an interval (a, b). Then we must have

$$J(a) = J(x) = J(b) \equiv J \tag{5.2.37}$$

and if one of the boundary conditions is reflecting, this means that both are reflecting, and $J = 0$.

If the boundaries are not reflecting, (5.2.37) requires them to be periodic. We then use the boundary conditions given by (5.2.17,18).

a) Zero Current—Potential Solution

Setting $J = 0$, we rewrite (5.2.37) as

$$A(x)p_s(x) = \frac{1}{2}\frac{d}{dx}[B(x)p_s(x)] = 0 \tag{5.2.38}$$

for which the solution is

$$p_s(x) = \frac{\mathcal{N}}{B(x)} \exp[2\int_a^x dx' A(x')/B(x')], \tag{5.2.39}$$

where $\mathcal{N}$ is a normalisation constant such that

$$\int_a^b dx\, p_s(x) = 1. \tag{5.2.40}$$

Such a solution is known as a *potential solution*, for various historical reasons, but chiefly because the stationary solution is obtained by a single integration (the full significance of this term will be treated in Sect.5.3.3).

b) Periodic Boundary Condition

Here we have nonzero current J and we rewrite (5.2.36) as

$$A(x)p_s(x) - \frac{1}{2}\frac{d}{dx}[B(x)p_s(x)] = J\,. \tag{5.2.41}$$

However, J is not arbitrary, but is determined by normalisation and the periodic boundary condition

$$p_s(a) = p_s(b) \tag{5.2.42}$$

$$J(a) = J(b). \tag{5.2.43}$$

For convenience, define

$$\psi(x) = \exp\left[2\int_a^x dx' A(x')/B(x')\right]. \tag{5.2.44}$$

Then we can easily integrate (5.2.41) to get

$$p_s(x)B(x)/\psi(x) = p_s(a)B(a)/\psi(a) - 2J\int_a^x dx'/\psi(x')\,. \tag{5.2.45}$$

By imposing the boundary condition (5.2.42) we find that

$$J = [B(b)/\psi(b) - B(a)/\psi(a)]p_s(a)/\left[\int_a^b dx'/\psi(x')\right] \tag{5.2.46}$$

so that

$$p_s(x) = p_s(a)\left[\frac{\int_a^x \frac{dx'}{\psi(x')}\frac{B(b)}{\psi(b)} + \int_x^b \frac{dx'}{\psi(x')}\frac{B(a)}{\psi(a)}}{\frac{B(x)}{\psi(x)}\int_a^b \frac{dx'}{\psi(x')}}\right]. \tag{5.2.47}$$

c) Infinite Range and Singular Boundaries

In either of these cases, one or the other of the above possibilities may turn out to be forbidden because of divergences, etc. A full enumeration of the possibilities is, in general, very complicated. We shall demonstrate these by means of the examples given in the next section.

5.2.3 Examples of Stationary Solutions

a) Diffusion in a Gravitational Field

A strongly damped Brownian particle moving in a constant gravitational field is often described by the SDE (Sect.6.4)

$$dx = -g\,dt + \sqrt{D}\,dW(t) \tag{5.2.48}$$

for which the Fokker-Planck equation is

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}(gp) + \frac{1}{2}D\frac{\partial^2 p}{\partial x^2}. \tag{5.2.49}$$

On the interval (a, b) with reflecting boundary conditions, the stationary solution is given by (5.2.39), i.e.

$$p_s(x) = \mathcal{N} \exp[-2gx/D], \tag{5.2.50}$$

where we have absorbed constant factors into the definition of $\mathcal{N}$.

Clearly this solution is normalisable on (a, b) only if a a is finite, though b may be infinite. The result is no more profound than to say that particles diffusing in a beaker of fluid will fall down, and if the beaker is infinitely deep, they will never stop falling! Diffusion upwards against gravity is possible for any distance but with exponentially small probability.

Now assume periodic boundary conditions on (a, b). Substitution into (5.2.47) yields

$$p_s(x) = p_s(a); \tag{5.2.51}$$

a constant distribution.

The interpretation is that the particles pass freely from a to b *and* back.

b) Ornstein Uhlenbeck Process

We use the notation of Sect.3.8.4 where the Fokker-Planck equation was

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}(kxp) + \frac{1}{2}D\frac{\partial^2 p}{\partial x^2}, \tag{5.2.52}$$

whose stationary solution on the interval (a, b) with reflecting barriers is

$$p_s(x) = \mathcal{N} \exp(-kx^2/D). \tag{5.2.53}$$

Provided $k > 0$, this is normalisable on $(-\infty, \infty)$. If $k < 0$, one can only make sense of it on a finite interval.

Suppose

$$a = -b < 0. \tag{5.2.54}$$

Then in this case,

$$\psi(x) = \exp\left[-\frac{k}{D}(x^2 - a^2)\right] \tag{5.2.55}$$

and if we consider the periodic boundary condition on this interval, by noting

$$\psi(a) = \psi(-a), \tag{5.2.56}$$

we find that

$$\begin{aligned} p_s(x) &= p_s(a)\psi(x)/\psi(a) \\ &= p_s(a)\exp\left[-\frac{k}{D}(x^2 - a^2)\right] \end{aligned}$$

so that the symmetry yields the same solution as in the case of reflecting barriers.

Letting $a \to \infty$, we see that we still have the same solution. The result is also true if $a \to \infty$ independently of $b \to -\infty$, provided $k > 0$.

c) A Chemical Reaction Model

Although chemical reactions are normally best modelled by a birth-death master equation formalism (as in Chap. 7), approximate treatments are often given by means of a FPE. The reaction

$$X + A \rightleftharpoons 2X \tag{5.2.57}$$

is of interest since it possesses an exit boundary at $x = 0$ (where x is the number of molecules of X). Clearly if there is no X, a collision between X and A cannot occur so no more X is produced.

The FPE is derived in Sect.7.6.1 and is

$$\partial_t p(x, t) = -\partial_x[(ax - x^2)p(x, t)] + \tfrac{1}{2}\partial_x^2[(ax + x^2)p(x, t)]\,. \tag{5.2.58}$$

We introduce reflecting boundaries at $x = \alpha$ and $x = \beta$. In this case, the stationary solution is

$$p_s(x) = e^{-2x}(a + x)^{4a-1}x^{-1} \tag{5.2.59}$$

which is not normalisable if $\alpha = 0$. The pole at $x = 0$ is a result of the absorption there. In fact, comparing with (5.2.28), we see that

$$\begin{aligned} B(0, t) &\equiv (ax + x^2)_{x=0} = 0 \\ A(0, t) &\equiv (ax - x^2)_{x=0} = 0 \\ \partial_x B(0, t) &\equiv (a + 2x)_{x=0} > 0 \end{aligned} \tag{5.2.60}$$

so we indeed have an exit boundary. The stationary solution has relevance only

if $\alpha > 0$ since it is otherwise not normalisable. The physical meaning of a reflecting barrier is quite simple: whenever a molecule of X disappears, we simply add another one immediately. A plot of $p_s(x)$ is given in Fig. 5.2. The time for all x to disappear is in practice extraordinarily long, and the stationary solution (5.2.59) is, in practice, a good representation of the distribution except near $x = 0$.

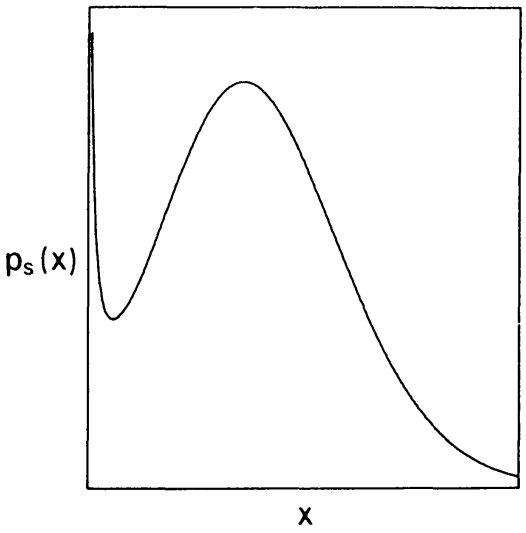


Fig. 5.2. Non-normalisable "stationary" $p_s(x)$ for the reaction $X + A \rightleftharpoons 2X$

5.2.4 Boundary Conditions for the Backward Fokker-Planck Equation

We suppose that $p(\boldsymbol{x}, t \mid \boldsymbol{x}', t')$ obeys the forward Fokker-Planck equation for a set of $\boldsymbol{x}, t$ and $\boldsymbol{x}', t'$, and that the process is confined to a region R with boundary $\boldsymbol{S}$. Then, if s is a time between t and t',

$$0 = \frac{\partial}{\partial s} p(\boldsymbol{x}, t \mid \boldsymbol{x}', t') = \frac{\partial}{\partial s} \int d\boldsymbol{y} p(\boldsymbol{x}, t \mid \boldsymbol{y}, s) p(\boldsymbol{y}, s \mid \boldsymbol{x}', t') , \tag{5.2.61}$$

where we have used the Chapman-Kolmogorov equation. We take the derivative $\partial/\partial s$ inside the integral, use the forward Fokker-Planck equation for the second factor and the backward equation for the first factor. For brevity, let us write

$$\begin{aligned} p(\boldsymbol{y}, s) &= p(\boldsymbol{y}, s \mid \boldsymbol{x}', t') \\ \bar{p}(\boldsymbol{y}, s) &= p(\boldsymbol{x}, t \mid \boldsymbol{y}, s) . \end{aligned} \tag{5.2.62}$$

Then,

$$\begin{aligned} 0 = \int_R d\boldsymbol{y} \Big[- \sum_i \frac{\partial}{\partial y_i} (A_i p) + \sum_{i,j} \frac{\partial^2}{\partial y_i \partial y_j} (B_{ij} p) \Big] \bar{p} \\ + \int_R d\boldsymbol{y} \Big[- \sum_i A_i \frac{\partial \bar{p}}{\partial y_i} - \sum_{i,j} B_{ij} \frac{\partial^2 \bar{p}}{\partial y_i \partial y_j} \Big] p \end{aligned} \tag{5.2.63}$$

and after some manipulation

$$= \int_R d\boldsymbol{y} \sum_i \frac{\partial}{\partial y_i} \Big\{ - A_i p \bar{p} + \frac{1}{2} \sum_j \Big[\bar{p} \frac{\partial}{\partial y_j} (B_{ij} p) - p B_{ij} \frac{\partial \bar{p}}{\partial y_j} \Big] \Big\} \tag{5.2.64}$$

$$= \int_S \sum_i dS_i \left\{ \bar{p} \left[-A_i p + \frac{1}{2} \sum_j \frac{\partial}{\partial y_i} (B_{ij} p) \right] \right\}$$

$$- \frac{1}{2} \int_S \sum_i dS_i p \left(\sum_j B_{ij} \frac{\partial \bar{p}}{\partial y_j} \right). \tag{5.2.65}$$

We now treat the various cases individually.

a) Absorbing Boundaries

This requires $p = 0$ on the boundary. That it also requires $\bar{p}(\boldsymbol{y}, t) = 0$ on the boundary is easily seen to be consistent with (5.2.65) since on substituting $p = 0$ in that equation, we get

$$0 = \int_S \bar{p} \sum_{i,j} dS_i B_{ij} \frac{\partial p}{\partial y_j}. \tag{5.2.66}$$

However, if the boundary is absorbing, clearly

$$p(\boldsymbol{x}, t | \boldsymbol{y}, s) = 0, \qquad \text{for } \boldsymbol{y} \in \text{boundary} \tag{5.2.67}$$

since this merely states that the probability of X re-entering R from the boundary is zero.

b) Reflecting Boundaries

Here the condition on the forward equation makes the first integral vanish in (5.2.65). The final factor vanishes for arbitrary p only if

$$\sum_{i,j} n_i B_{ij}(\boldsymbol{y}) \frac{\partial}{\partial y_j} [p(\boldsymbol{x}, t | \boldsymbol{y}, s)] = 0 . \tag{5.2.68}$$

In one dimension this reduces to

$$\frac{\partial}{\partial y} p(\boldsymbol{x}, t | \boldsymbol{y}, s) = 0 \tag{5.2.69}$$

unless B vanishes.

c) Other Boundaries

We shall not consider these in this section. For further details see [5.4].

5.2.5 Eigenfunction Methods (Homogeneous Processes)

We shall now show how, in the case of homogeneous processes, solutions can most naturally be expressed in terms of eigenfunctions. We consider reflecting and absorbing boundaries.

a) Eigenfunctions for Reflecting Boundaries

We consider a Fokker-Planck equation for a process on an interval (a, b) with reflecting boundaries. We suppose the FPE to have a stationary solution $p_s(x)$ and the form

$$\partial_t p(x, t) = -\partial_x[A(x)p(x, t)] + \tfrac{1}{2}\,\partial_x^2\,[B(x)p(x, t)]\,. \tag{5.2.70}$$

We define a function $q(x, t)$ by

$$p(x, t) = p_s(x)q(x, t) \tag{5.2.71}$$

and, by direct substitution, find that $q(x, t)$ satisfies the *backward equation*

$$\partial_t q(x, t) = A(x)\partial_x q(x, t) + \tfrac{1}{2}\,B(x)\partial_x^2 q(x, t)\,. \tag{5.2.72}$$

We now wish to consider solutions of the form

$$p(x, t) = P_\lambda(x)\mathrm{e}^{-\lambda t} \tag{5.2.73}$$

$$q(x, t) = Q_\lambda(x)\mathrm{e}^{-\lambda t} \tag{5.2.74}$$

which obey the eigenfunction equations

$$-\partial_x[A(x)P_\lambda(x)] + \tfrac{1}{2}\,\partial_x^2[B(x)P_\lambda(x)] = -\lambda P_\lambda(x) \tag{5.2.75}$$

$$A(x)\partial_x Q_{\lambda'}(x) + \tfrac{1}{2}\,B(x)\partial_x^2 Q_{\lambda'}(x) = -\lambda' Q_{\lambda'}(x)\,. \tag{5.2.76}$$

Then we can straightforwardly show by partial integration that

$$\begin{aligned}(\lambda' - \lambda)\int_a^b dx P_\lambda(x)Q_{\lambda'}(x) = [Q_{\lambda'}(x)\{-A(x)P_\lambda(x) + \tfrac{1}{2}\,\partial_x[B(x)P_\lambda(x)]\} \\ - \tfrac{1}{2}\,B(x)P_\lambda(x)\partial_x Q_{\lambda'}(x)]_a^b,\end{aligned} \tag{5.2.77}$$

and using the reflecting boundary condition on the coefficient of $Q_{\lambda'}(x)$, we see that it vanishes. Further, using the definition of $q(x, t)$ in terms of the stationary solution (5.2.71), it is simple to show that

$$\tfrac{1}{2}\,B(x)\partial_x Q_{\lambda'}(x) = -A(x)P_{\lambda'}(x) + \tfrac{1}{2}\,\partial_x[B(x)P_{\lambda'}(x)] \tag{5.2.78}$$

so that term vanishes also. Hence, the $Q_\lambda(x)$ and $P_\lambda(x)$ form a bi-orthogonal system

$$\int_a^b dx\; P_\lambda(x)Q_{\lambda'}(x) = \delta_{\lambda\lambda'} \tag{5.2.79}$$

or, there are two alternative orthogonality systems,

$$\int_a^b dx\; p_s(x)Q_\lambda(x)Q_{\lambda'}(x) = \delta_{\lambda\lambda'} \tag{5.2.80}$$

$$\int_a^b dx[p_s(x)]^{-1}P_\lambda(x)P_{\lambda'}(x) = \delta_{\lambda\lambda'}\,. \tag{5.2.81}$$

It should be noted that setting $\lambda = \lambda' = 0$ gives the normalisation of the stationary solution $p_s(x)$ since

$$P_0(x) = p_s(x) \tag{5.2.82}$$

$$Q_0(x) = 1 \,. \tag{5.2.83}$$

Using this orthogonality we can write any solution in terms of eigenfunctions. For if

$$p(x, t) = \sum_\lambda A_\lambda P_\lambda(x) e^{-\lambda t} \,, \tag{5.2.84}$$

then

$$\int_a^b dx\, Q_\lambda(x) p(x, 0) = A_\lambda \,. \tag{5.2.85}$$

For example, the conditional probability $p(x, t \,|\, x_0, 0)$ is given by the initial condition

$$p(x, 0 \,|\, x_0, 0) = \delta(x - x_0) \tag{5.2.86}$$

so that

$$A_\lambda = \int_a^b dx\, Q_\lambda(x) \delta(x - x_0) = Q_\lambda(x_0) \tag{5.2.87}$$

and hence,

$$p(x, t \,|\, x_0, 0) = \sum_\lambda P_\lambda(x) Q_\lambda(x_0) e^{-\lambda t} \,. \tag{5.2.88}$$

We can write the autocorrelation function quite elegantly as

$$\langle x(t)x(0)\rangle = \int dx \int dx_0\, x x_0 p(x, t \,|\, x_0, 0) p_s(x) \tag{5.2.89}$$

$$= \sum_\lambda \left[\int dx\, x P_\lambda(x)\right]^2 e^{-\lambda t} \,, \tag{5.2.90}$$

where we have used the definition of $Q_\lambda(x)$ by (5.2.74).

b) Eigenfunctions for Absorbing Boundaries

This is very similar. We define P_λ and Q_λ as above, except that $p_s(x)$ is still the stationary solution of the Fokker-Planck equation with *reflecting boundary conditions*. With this definition, we find that we must have

$$P_\lambda(a) = Q_\lambda(a) = P_\lambda(b) = Q_\lambda(b) = 0 \tag{5.2.91}$$

and the orthogonality proof still follows through. Eigenfunctions are then computed using this condition and the eigenfunction equations (5.2.75, 76) and all other results look the same. However, the range of λ does not include $\lambda = 0$, and hence $p(x, t \mid x_0, 0) \to 0$ as $t \to \infty$.

5.2.6 Examples

a) A Wiener Process with Absorbing Boundaries
The Fokker-Planck equation

$$\partial_t p = \tfrac{1}{2}\, \partial_x^2 p \tag{5.2.92}$$

it treated on the interval (0, 1). The absorbing boundary condition requires

$$p(0, t) = p(1, t) = 0 \tag{5.2.93}$$

and the appropriate eigenfunctions are $\sin(n\pi x)$ so we expand in a Fourier sine series

$$p(x, t) = \sum_{n=1}^{\infty} b_n(t) \sin(n\pi x) \tag{5.2.94}$$

which automatically satisfies (5.2.93). The initial condition is chosen so that

$$p(x, 0) = \delta(x - x_0) \tag{5.2.95}$$

for which the Fourier coefficients are

$$b_n(0) = 2 \int_0^1 dx\, \delta(x - x_0) \sin(n\pi x) = 2 \sin(n\pi x_0)\,. \tag{5.2.96}$$

Substituting the Fourier expansion (5.2.94) into (5.2.92) gives

$$\dot{b}_n(t) = -\lambda_n b_n(t) \tag{5.2.97}$$

with

$$\lambda_n = n^2\pi^2/2 \tag{5.2.98}$$

and the solution

$$b_n(t) = b_n(0)\exp(-\lambda_n t)\,. \tag{5.2.99}$$

So we have the solution [which by the initial condition (5.2.95) is for the conditional probability $p(x, t \mid x_0, 0)$]

$$p(x, t \mid x_0, 0) = 2 \sum_{n=1}^{\infty} \exp(-\lambda_n t) \sin(n\pi x_0) \sin(n\pi x)\,. \tag{5.2.100}$$

b) Wiener Process with Reflecting Boundaries

Here the boundary condition reduces to [on the interval (0, 1)]

$$\partial_x p(0, t) = \partial_x p(1, t) = 0 \tag{5.2.101}$$

and the eigenfunctions are now $\cos(n\pi x)$, so we make a Fourier cosine expansion

$$p(x, t) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} a_n(t) \cos(n\pi x) \tag{5.2.102}$$

with the same initial condition

$$p(x, 0) = \delta(x - x_0) \tag{5.2.103}$$

so that

$$a_n(0) = 2 \int_0^1 dx \cos(n\pi x)\delta(x - x_0) = 2\cos(n\pi x_0) . \tag{5.2.104}$$

In the same way as before, we find

$$a_n(t) = a_n(0) \exp(-\lambda_n t) \tag{5.2.105}$$

with

$$\lambda_n = n^2\pi^2/2 \tag{5.2.106}$$

so that

$$p(x, t|x_0, 0) = 1 + 2 \sum_{n=1}^{\infty} \cos(n\pi x_0) \cos(n\pi x) \exp(-\lambda_n t) . \tag{5.2.107}$$

As $t \to \infty$, the process becomes stationary, with stationary distribution

$$p_s(x) = \lim_{t\to\infty} p(x, t|x_0, 0) = 1 . \tag{5.2.108}$$

We can compute the stationary autocorrelation function by

$$\langle x(t)x(0)\rangle_s = \int_0^1 \int_0^1 dx\, dx_0\, x x_0 p(x, t|x_0, 0) p_s(x) \tag{5.2.109}$$

and carrying out the integrals explicitly,

$$\langle x(t)x(0)\rangle_s = \frac{1}{4} + \frac{8}{\pi^4} \sum_{n=0}^{\infty} \exp(-\lambda_{2n+1} t) (2n + 1)^{-4} . \tag{5.2.110}$$

We see that as $t \to \infty$, all the exponentials vanish and

$$\langle x(t)x(0)\rangle_s \to \tfrac{1}{4} = [\langle x\rangle_s]^2 \tag{5.2.111}$$

and as $t \to 0$,

$$\langle x(t)x(0)\rangle_s \to \frac{1}{4} + \frac{8}{\pi^4}\sum_{n=0}^{\infty}(2n+1)^{-4} = \frac{1}{3} = \langle x_s^2\rangle \tag{5.2.113}$$

when one takes account of the identity (from the theory of the Riemann zeta-function)

$$\sum_{n=0}^{\infty}(2n+1)^{-4} = \frac{\pi^4}{96}. \tag{5.2.114}$$

c) Ornstein-Uhlenbeck Process

As in Sect.3.8.4, the Fokker-Planck equation is

$$\partial_t p(x,t) = \partial_x[kxp(x,t)] + \tfrac{1}{2}D\partial_x^2 p(x,t)\,. \tag{5.2.115}$$

The eigenfunction equation for Q_λ is

$$d_x^2 Q_\lambda - \frac{2kx}{d}d_x Q_\lambda + \frac{2\lambda}{D}Q_\lambda = 0 \tag{5.2.116}$$

and this becomes the differential equation for *Hermite polynomials* [5.6] on making the replacement

$$y = x\sqrt{k/D}$$
$$d_y^2 Q_\lambda - 2yd_y Q_\lambda + (2\lambda/k)Q_\lambda = 0. \tag{5.2.117}$$

We can write

$$Q_\lambda = (2^n n!)^{-1/2}\mathrm{H}_n(x\sqrt{k/D})\,, \tag{5.2.118}$$

where

$$\lambda = nk\,, \tag{5.2.119}$$

and these solutions are normalised as in (5.2.79–81).

The stationary solution is, as previously found,

$$p_s(x) = (k/\pi D)^{1/2}\exp(-kx^2/D) \tag{5.2.120}$$

and a general solution can be written as

$$p(x,t) = \sum_n \sqrt{[k/(2^n n!\pi D)]}\exp(-kx^2/D)\mathrm{H}_n(x\sqrt{k/D})\mathrm{e}^{-nkt}A_n \tag{5.2.121}$$

with

$$A_n = \int_{-\infty}^{\infty} dx\; p(x,0)\mathrm{H}_n(x\sqrt{k/D})(2^n n!)^{-1/2}\,. \tag{5.2.122}$$

The result can also be obtained directly from the explicit solution for the conditional probability given in Sect.3.8.4 by using generating functions for Hermite polynomials. One sees that the time scale of relaxation to the stationary state is given by the eigenvalues

$$\lambda_n = nk. \tag{5.2.123}$$

Here, k is the rate constant for deterministic relaxation, and it thus determines the slowest time in the relaxation. One can also compute the autocorrelation function directly using (5.2.90). We use the result [5.6] that

$$\mathrm{H}_1(y) = 2y \tag{5.2.124}$$

so that the orthogonality property means that only the eigenfunction corresponding to $n = 1$ has a nonzero coefficient. We must compute

$$\begin{aligned} \int x P_{\lambda_1}(x)dx &= \int_{-\infty}^{\infty} \sqrt{k/(2\pi D)} \exp(-kx^2/D)(2x\sqrt{k/D})x \\ &= \sqrt{D/2k} \end{aligned} \tag{5.2.125}$$

so that

$$\langle x(t)x(0)\rangle_s = \frac{D}{2k} e^{-kt}, \tag{5.2.126}$$

as found previously in Sects.3.8.4, 4.4.4.

d) Rayleigh Process

We take the model of amplitude fluctuations developed in Sect.4.4.5. The Fokker-Planck equation is

$$\partial_t p(x, t) = \partial_x[(\gamma x - \mu/x)p(x, t)] + \mu\partial_x^2 p(x, t), \tag{5.2.127}$$

where

$$\mu = \varepsilon^2/2. \tag{5.2.128}$$

The range here is $(0, \infty)$ and the stationary solution (normalised)

$$p_s(x) = (\gamma x/\mu) \exp(-\gamma x^2/2\mu). \tag{5.2.129}$$

The eigenfunction equation for the $Q_\lambda(x)$ is

$$d_x^2 Q_\lambda + (1/x - \gamma x/\mu)d_x Q_\lambda + (\lambda/\mu)Q_\lambda = 0. \tag{5.2.130}$$

By setting

$$z = x^2\gamma/2\mu, \tag{5.2.131}$$

we obtain

$$z d_z^2 Q_\lambda + (1 - z) d_z Q_\lambda + (\lambda/2\gamma) Q_\lambda = 0 . \tag{5.2.132}$$

This is the differential equation for the *Laguerre polynomials* [5.6] provided

$$\lambda = 2n\gamma . \tag{5.2.133}$$

We can write

$$Q_\lambda(x) = \mathrm{L}_n(\gamma x^2/2\mu) \tag{5.2.134}$$

which is normalised. Hence, the conditional probability is

$$p(x, t \mid x_0, 0) = \sum_{n=0}^{\infty} \frac{\gamma x}{\mu} \exp\left(-\frac{\gamma x^2}{2\mu}\right) \mathrm{L}_n\left(\frac{\gamma x_0^2}{2\mu}\right) \mathrm{L}_n\left(\frac{\gamma x^2}{2\mu}\right) \mathrm{e}^{-2n\gamma t} . \tag{5.2.135}$$

We can compute the autocorrelation function by the method of (5.2.90):

$$\langle x(t)x(0)\rangle = \sum_{n=0}^{\infty} \left[\int_0^{\infty} x\, dx \frac{\gamma x}{\mu} \exp\left(\frac{-\gamma x^2}{2\mu}\right) \mathrm{L}_n\left(\frac{\gamma x^2}{2\mu}\right)\right]^2 \exp(-2n\gamma t) \tag{5.2.136}$$

and using

$$\int_0^{\infty} dz\, z^\alpha \mathrm{e}^{-z} \mathrm{L}_n(z) = (-1)^n \Gamma(\alpha + 1) \binom{\alpha}{n}, \tag{5.2.137}$$

we find for the autocorrelation function

$$\langle x(t)x(0)\rangle = \frac{2\mu}{\gamma} \sum_{n=0}^{\infty} \frac{\pi}{4} \binom{\frac{1}{2}}{n}^2 \exp(-2n\gamma t). \tag{5.2.138}$$

5.2.7 First Passage Times for Homogeneous Processes

It is often of interest to know how long a particle whose position is described by a Fokker-Planck equation remains in a certain region of x. The solution of this problem can be achieved by use of the *backward Fokker-Planck equation*, as described in Sect. 3.6.

a) Two Absorbing Barriers

Let the particle be initially at x at time $t = 0$ and let us ask how long it remains in the interval (a, b) which is assumed to contain x:

$$a \leqslant x \leqslant b \tag{5.2.139}$$

We erect absorbing barriers at a and b so that the particle is removed from the system when it reaches a or b. Hence, if it is still in the interval (a, b), it has never left that interval.

Under these conditions, the probability that at time t the particle is still in (a, b) is

$$\int_a^b dx' p(x', t \mid x, 0) \equiv G(x, t) . \tag{5.2.140}$$

Let the time that the particle leaves (a, b) be T. Then we can rewrite (5.2.140) as

$$\text{Prob}(T \geqslant t) = \int_a^b dx' p(x', t \mid x, 0) \tag{5.2.141}$$

which means that $G(x, t)$ is the same as $\text{Prob}(T \geqslant t)$. Since the system is time homogeneous, we can write

$$p(x', t \mid x, 0) = p(x', 0 \mid x, -t) \tag{5.2.142}$$

and the backward Fokker-Planck equation can be written

$$\partial_t p(x', t \mid x, 0) = A(x)\partial_x p(x', t \mid x, 0) + \tfrac{1}{2} B(x)\partial_x^2 p(x', t \mid x, 0) \tag{5.2.143}$$

and hence, $G(x, t)$ obeys the equation

$$\partial_t G(x, t) = A(x)\partial_x G(x, t) + \tfrac{1}{2} B(x)\partial_x^2 G(x, t) . \tag{5.2.144}$$

The boundary conditions are clearly that

$$p(x', 0 \mid x, 0) = \delta(x - x')$$

and hence,

$$\begin{aligned} G(x, 0) &= 1 \qquad a \leqslant x \leqslant b \\ &= 0 \qquad \text{elsewhere} \end{aligned} \tag{5.2.145}$$

and if $x = a$ or b, the particle is absorbed immediately, so

$$\begin{aligned} &\text{Prob}(T \geqslant t) = 0 \qquad \text{when } x = a \text{ or } b , \qquad \text{i.e.,} \\ &G(a, t) = G(b, t) = 0. \end{aligned} \tag{5.2.146}$$

Since $G(x, t)$ is the probability that $T \geqslant t$, the mean of any function of T is

$$\langle f(T) \rangle = -\int_0^\infty f(t) dG(x, t) . \tag{5.2.147}$$

Thus, the *mean first passage time*

$$T(x) = \langle T \rangle \tag{5.2.148}$$

is given by

$$T(x) = -\int_0^\infty t\, \partial_t G(x, t) dt \tag{5.2.149}$$

$$= \int_0^\infty G(x, t)dt \tag{5.2.150}$$

after integrating by parts.

Similarly, defining

$$T_n(x) = \langle T^n \rangle , \tag{5.2.151}$$

we find

$$T_n(x) = \int_0^\infty t^{n-1} G(x, t)dt . \tag{5.2.152}$$

We can derive a simple ordinary differential equation for $T(x)$ by using (5.2.150) and integrating (5.2.144) over $(0, \infty)$. Noting that

$$\int_0^\infty \partial_t G(x, t)dt = G(x, \infty) - G(x, 0) = -1 , \tag{5.2.153}$$

we derive

$$A(x)\partial_x T(x) + \tfrac{1}{2} B(x)\partial_x^2 T(x) = -1 \tag{5.2.154}$$

with the boundary condition

$$T(a) = T(b) = 0 . \tag{5.2.155}$$

Similarly, we see that

$$-nT_{n-1}(x) = A(x)\partial_x T_n(x) + \tfrac{1}{2} B(x)\partial_x^2 T_n(x) \tag{5.2.156}$$

which means that all the moments of the first passage time can be found by repeated integration.

Solutions of the Equations. Equation (5.2.154) can be solved directly by integration. The solution, after some manipulation, can be written in terms of

$$\boxed{\psi(x) = \exp\left\{\int_a^x dx'[2A(x')/B(x')]\right\} .} \tag{5.2.157}$$

We find

$$T(x) = \frac{2\left[\left(\int_a^x \frac{dy}{\psi(y)}\right)\int_x^b \frac{dy'}{\psi(y')}\int_a^{y'} \frac{dz\psi(z)}{B(z)} - \left(\int_x^b \frac{dy}{\psi(y)}\right)\int_a^x \frac{dy'}{\psi(y')}\int_a^{y'} \frac{dz\psi(z)}{B(z)}\right]}{\int_a^b \frac{dy}{\psi(y)}} . \tag{5.2.158}$$

b) One Absorbing Barrier

We consider motion still in the interval (a, b) but suppose the barrier at a to be reflecting. The boundary conditions then become

$$\partial_x G(a, t) = 0 \tag{5.2.159a}$$

$$G(b, t) = 0 \tag{5.2.159b}$$

which follow from the conditions on the backward Fokker-Planck equation derived in Sect.5.2.4. We solve (5.2.154) with the corresponding boundary condition and obtain

$$\boxed{T(x) = 2\int_x^b \frac{dy}{\psi(y)} \int_a^y \frac{\psi(z)}{B(z)}\, dz \qquad \begin{array}{l} a \text{ reflecting} \\ b \text{ absorbing} \\ a < b \end{array}} \tag{5.2.160}$$

Similarly, one finds

$$\boxed{T(x) = 2\int_a^x \frac{dy}{\psi(y)} \int_y^b \frac{\psi(z)}{B(z)}\, dz \qquad \begin{array}{l} b \text{ reflecting} \\ a \text{ absorbing} \\ a < b \end{array}} \tag{5.2.161}$$

c) Application—Escape Over a Potential Barrier

We suppose that a point moves according to the Fokker-Planck equation

$$\partial_t p(x, t) = \partial_x [U'(x) p(x, t)] + D\partial_x^2 p(x, t) \,. \tag{5.2.162}$$

The potential has maxima and minima, as shown in Fig. 5.3. We suppose that motion is on an infinite range, which means the stationary solution is

$$p_s(x) = \mathcal{N} \exp\left[-U(x)/D\right] \tag{5.2.163}$$

which is bimodal (as shown in Fig. 5.3) so that there is a relatively high probability of being on the left or the right of b, but not near b. What is the mean escape time from the left hand well? By this we mean, what is the mean first passage time from a to x , where x is in the vicinity of b? We use (5.2.160) with the substitutions

$$\begin{aligned} b &\to x_0 \\ a &\to -\infty \\ x &\to a \end{aligned} \tag{5.2.164}$$

so that

$$T(a \to x_0) = \frac{1}{D} \int_a^{x_0} dy \exp\left[U(y)/D\right] \int_{-\infty}^{y} \exp\left[-U(z)/D\right] dz \,. \tag{5.2.165}$$

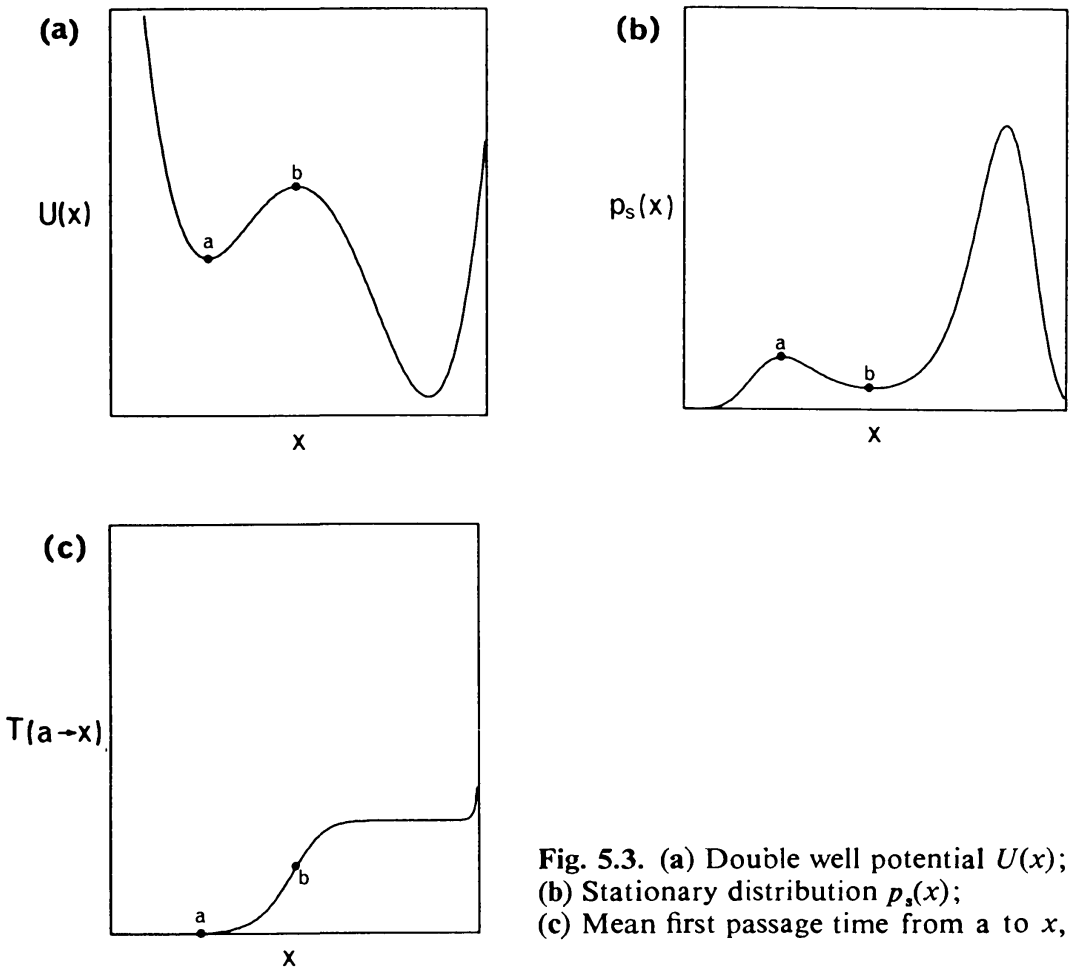


Fig. 5.3. (a) Double well potential $U(x)$;
(b) Stationary distribution $p_s(x)$;
(c) Mean first passage time from a to x, $T(a \to x_0)$

If the central maximum of $U(x)$ is large and D is small, then $\exp[U(y)/D]$ is sharply peaked at $x = b$, while $\exp[-U(z)/D]$ is very small near $z = b$. Therefore, $\int_{-\infty}^{y} \exp[-U(z)/D]dz$ is a very slowly varying function of y near $y = b$. This means that the value of the integral $\int_{-\infty}^{y} \exp[-U(z)/D]dz$ will be approximately constant for those values of y which yield a value of $\exp[U(y)/D]$ which is significantly different from zero. Hence, in the inner integral, we can set $y = b$ and remove the resulting constant factor from inside the integral with respect to y. Hence, we can approximate (5.2.165) by

$$T(a \to x_0) \simeq \left\{\frac{1}{D}\int_{-\infty}^{b} dy \exp[-U(z)/D]\right\}\int_{a}^{x_0} dy \exp[U(y)/D]\,. \tag{5.2.166}$$

Notice that by the definition of $p_s(x)$ in (5.2.163), we can say that

$$\int_{-\infty}^{b} dy \exp[-U(z)/D] = n_a/\mathcal{N} \tag{5.2.167}$$

which means that n_a is the probability that the particle is to the left of b when the system is stationary.

A plot of $T(a \to x_0)$ against x_0 is shown in Fig. 5.3 and shows that the mean first passage time to x_0 is quite small for x_0 in the left well and quite large for x_0 in the

right well. This means that the particle, in going over the barrier to the right well, takes most of the time in actually surmounting the barrier. It is quite meaningful to talk of the *escape time* as that time for the particle, initially at a, to reach a point near c, since this time is quite insensitive to the exact location of the initial and final points. We can evaluate this by further assuming that near b we can write

$$U(x) \simeq U(b) - \frac{1}{2}\left(\frac{x-b}{\delta}\right)^2 \tag{5.2.168}$$

and near a

$$U(x) \simeq U(a) + \frac{1}{2}\left(\frac{x-a}{\alpha}\right)^2 . \tag{5.2.169}$$

The constant factor in (5.2.166) is evaluated as

$$\int_{-\infty}^{b} dz \exp[-U(z)/D] \sim \int_{-\infty}^{\infty} dz \exp\left[-\frac{U(a)}{D} - \frac{(z-a)^2}{2D\alpha^2}\right] \tag{5.2.170}$$

$$\sim \alpha\sqrt{2\pi D} \exp[-U(a)/D] \tag{5.2.171}$$

and the inner factor becomes, on assuming x_0 is well to the right of the central point b,

$$\int_{a}^{x_0} dy \exp U(y)/D \sim \int_{-\infty}^{\infty} dy \exp\left[\frac{U(b)}{D} - \frac{(y-b)^2}{2D\delta^2}\right] \tag{5.2.172}$$

$$= \delta\sqrt{2\pi D} \exp[U(b)/D] . \tag{5.2.173}$$

Putting both of these in (5.2.166), we get

$$T(a \to x_0) \simeq 2\alpha\delta\pi \exp\{[U(b) - U(a)]/D\} . \tag{5.2.174}$$

This is the classical *Arrhenius formula* of chemical reaction theory. In a chemical reaction, we can model the reaction by introducing a coordinate such that $x = a$ is species A and $x = c$ is species C. The reaction is modelled by the above diffusion process and the two distinct chemical species are separated by the potential barrier at b. In the chemical reaction, statistical mechanics gives the value

$$D = kT , \tag{5.2.175}$$

where k is Boltzmann's constant and T is the absolute temperature. We see that the most important dependence on temperature comes from the exponential factor which is often written

$$\exp(\Delta E/kT) \tag{5.2.176}$$

and predicts a very characteristic dependence on temperature. Intuitively, the answer is obvious. The exponential factor represents the probability that the energy

will exceed that of the barrier when the system is in thermal equilibrium. Those molecules that reach this energy then react, with a certain finite probability.

We will come back to problems like this in great detail in Chap.9.

5.2.8 Probability of Exit Through a Particular End of the Interval

What is the probability that the particle, initially at x in (a, b), exits through a, and what is the mean exit time?

The total probability that the particle exits through a after time t is given by the time integral of the probability current at a. We thus define this probability by

$$g_a(x, t) = -\int_t^\infty dt' \, J(a, t' | x, 0) \tag{5.2.177}$$

$$= \int_t^\infty dt' \{-A(a)p(a, t' | x, 0) + \tfrac{1}{2}\partial_a[B(a)p(a, t' | x, 0)]\} \tag{5.2.178}$$

(the negative sign is chosen since we need the current pointing to the left) and

$$g_b(x, t) = \int_t^\infty dt' \{A(b)p(b, t' | x, 0) - \tfrac{1}{2}\partial_b[B(b)p(b, t' | x, 0)]\} . \tag{5.2.179}$$

These quantities give the probabilities that the particle exits through a or b after time t, respectively. The probability that (given that it exits through a) it exits after time t is

$$\text{Prob}(T_a > t) = g_a(x, t)/g_a(x, 0) . \tag{5.2.180}$$

We now find an equation for $g_a(x, t)$. We use the fact that $p(a, t | x, 0)$ satisfies a backward Fokker-Planck equation. Thus,

$$\begin{aligned} A(x)\partial_x g_a(x, t) + \tfrac{1}{2}B(x)\partial_x^2 g_a(x, t) &= -\int_t^\infty dt' \partial_{t'} J(a, t' | x, 0) \\ &= J(a, t | x, 0) \\ &= \partial_t g_a(x, t) . \end{aligned} \tag{5.2.181}$$

The mean exit time, given that exit is through a, is

$$T(a, x) = -\int_0^\infty t\partial_t \text{Prob}\ (T_a > t)\, dt = \int_0^\infty g_a(x, t)dt/g_a(x, \infty) . \tag{5.2.182}$$

Simply integrating (5.2.181) with respect to t, we get

$$\boxed{A(x)\partial_x[\pi_a(x)T(a, x)] + \tfrac{1}{2}B(x)\partial_x^2[\pi_a(x)T(a, x)] = -\pi_a(x) ,} \tag{5.2.183}$$

where we define

$$\pi_a(x) = (\text{probability of exit through } a) = g_a(x, 0) . \tag{5.2.184}$$

The boundary conditions on (5.2.183) are quite straightforward since they follow from those for the backward Fokker-Planck equation, namely,

$$\pi_a(a)T(a, a) = \pi_a(b)T(a, b) = 0 . \tag{5.2.185}$$

In the first of these clearly $T(a, a)$ is zero (the time to reach a from a is zero) and in the second, $\pi_a(b)$ is zero (the probability of exiting through a, starting from b, is zero).

By letting $t \to 0$ in (5.2.181), we see that $J(a, 0|x, 0)$ must vanish if $a \neq x$, since $p(a, 0|x, 0) = \delta(x - a)$. Hence, the right-hand side tends to zero and we get

$$\boxed{A(x)\partial_x \pi_a(x) + \tfrac{1}{2}B(x)\partial_x^2 \pi_a(x) = 0 ,} \tag{5.2.186}$$

the boundary condition this time being

$$\boxed{\begin{aligned} \pi_a(a) &= 1 \\ \pi_a(b) &= 0 . \end{aligned}} \tag{5.2.187}$$

The solution of (5.2.186) subject to this boundary condition and the condition

$$\pi_a(x) + \pi_b(x) = 1 \tag{5.2.188}$$

is

$$\boxed{\pi_a(x) = \left[\int_x^b dy\, \psi(y)\right] \Big/ \int_a^b dy\, \psi(y)} \tag{5.2.189}$$

$$\boxed{\pi_b(x) = \left[\int_a^x dy\, \psi(y)\right] \Big/ \int_a^b dy\, \psi(y)} \tag{5.2.190}$$

with $\psi(x)$ as defined in (5.2.157).

These formulae find application in the problem of relaxation of a distribution initially concentrated at an unstable stationary point (Sect.9.1.4).

5.3 Fokker-Planck Equations in Several Dimensions

In many variable situations, Fokker-Planck equations take on an essentially more complex range of behaviour than is possible in the case of one variable. Boundaries are no longer simple end points of a line but rather curves or surfaces, and the nature of the boundary can change from place to place. Stationary solutions even with reflecting boundaries can correspond to nonzero probability currents and eigenfunction methods are no longer so simple.

Nevertheless, the analogies between one and many dimensions are useful, and this section will follow the same general outline as that on one-variable situations.

5.3.1 Change of Variables

Suppose we have a Fokker-Planck equation in variable x_i,

$$\partial_t p(\boldsymbol{x}, t) = -\sum_i \partial_i [A_i(\boldsymbol{x}) p(\boldsymbol{x}, t)] + \tfrac{1}{2} \sum_{i,j} \partial_i \partial_j [B_{ij}(\boldsymbol{x}) p(\boldsymbol{x}, t)] \tag{5.3.1}$$

and we want to know the corresponding equation for the variables

$$y_i = f_i(\boldsymbol{x}) , \tag{5.3.2}$$

where f_i are certain differentiable independent functions. Let us denote by $\tilde{p}(\boldsymbol{y}, t)$ the probability density for the new variable, which is given by

$$\tilde{p}(\boldsymbol{y}, t) = p(\boldsymbol{x}, t) \left| \frac{\partial(x_1, x_2 \ldots)}{\partial(y_1, y_2 \ldots)} \right| . \tag{5.3.3}$$

The simplest way to effect the change of variables is to use Ito's formula on the corresponding SDE

$$d\boldsymbol{x}(t) = A(\boldsymbol{x}) dt + \sqrt{\underset{\sim}{B}(\boldsymbol{x})}\, d\boldsymbol{W}(t) \tag{5.3.4}$$

and then recompute the corresponding FPE for $\tilde{p}(\boldsymbol{y}, t)$ from the resulting SDE as derived in Sect. 4.3.4.

The result is rather complicated. In specific situations, direct implementation (5.3.3) may be preferable. There is no way of avoiding a rather messy calculation unless full use of symmetries and simplifications is made.

Example: Cartesian to Polar Coordinates. As an example, one can consider the transformation to polar coordinates of the Rayleigh process, previously done by the SDE method in Sect.4.4.5. Thus, the Fokker-Planck equation is

$$\partial_t p(E_1, E_2, t) = \gamma \frac{\partial}{\partial E_1} E_1 p + \gamma \frac{\partial}{\partial E_2} E_2 p + \frac{1}{2} \varepsilon^2 \left(\frac{\partial^2 p}{\partial E_1^2} + \frac{\partial^2 p}{\partial E_2^2} \right) \tag{5.3.5}$$

and we want to find the FPE for a and ϕ defined by

$$\begin{aligned} E_1 &= a \cos\phi \\ E_2 &= a \sin\phi . \end{aligned} \tag{5.3.6}$$

The Jacobian is

$$\begin{aligned} |J| &= \frac{\partial(E_1, E_2)}{\partial(a, \phi)} = \begin{vmatrix} \cos\phi & -a \sin\phi \\ \sin\phi & a \cos\phi \end{vmatrix} \\ &= a . \end{aligned} \tag{5.3.7}$$

We use the polar form of the Laplacian to write

$$\frac{\partial^2}{\partial E_1^2} + \frac{\partial^2}{\partial E_2^2} = \frac{1}{a^2}\frac{\partial^2}{\partial \phi^2} + \frac{1}{a}\frac{\partial}{\partial a}\left(a\frac{\partial}{\partial a}\right) \tag{5.3.8}$$

and inverting (5.3.6)

$$\left.\begin{aligned} a &= \sqrt{E_1^2 + E_2^2} \\ \phi &= \tan^{-1}(E_2/E_1)\,, \end{aligned}\right\} \tag{5.3.9}$$

we note

$$\left.\begin{aligned} &\frac{\partial a}{\partial E_1} = \frac{E_1}{\sqrt{E_1^2 + E_2^2}} = \cos\phi\,. \\ &\text{Similarly,} \\ &\frac{\partial a}{\partial E_2} = \sin\phi \end{aligned}\right\} \tag{5.3.10}$$

and

$$\left.\begin{aligned} &\frac{\partial \phi}{\partial E_2} = \frac{E_1}{E_1^2 + E_2^2} = \cos\phi/a\,. \\ &\text{Similarly,} \\ &\frac{\partial \phi}{\partial E_1} = -\sin\phi/a\,. \end{aligned}\right\} \tag{5.3.11}$$

Hence,

$$\begin{aligned} &\frac{\partial}{\partial E_1}E_1 p + \frac{\partial}{\partial E_2}E_2 p \\ &\quad = 2p + E_1\left(\frac{\partial p}{\partial a}\frac{\partial a}{\partial E_1} + \frac{\partial p}{\partial \phi}\frac{\partial \phi}{\partial E_1}\right) + E_2\left(\frac{\partial p}{\partial a}\frac{\partial a}{\partial E_2} + \frac{\partial p}{\partial \phi}\frac{\partial \phi}{\partial E_2}\right) \\ &\quad = 2p + a\frac{\partial p}{\partial a} = \frac{1}{a}\frac{\partial}{\partial a}(a^2 p)\,. \end{aligned} \tag{5.3.12}$$

Let us use the symbol $\tilde{p}(a, \phi)$ for the density function in terms of a and ϕ. The Jacobian formula (5.5.3) tells us that

$$\tilde{p}(a, \phi) = \left|\frac{\partial(E_1, E_2)}{\partial(a, \phi)}\right| p(E_1, E_2) = ap(E_1, E_2)\,. \tag{5.3.13}$$

Putting together (5.3.5, 8, 12, 13), we get

$$\frac{\partial \tilde{p}}{\partial t} = -\frac{\partial}{\partial a}\left[\left(-\gamma a + \frac{\varepsilon^2}{2a}\right)\tilde{p}\right] + \frac{\varepsilon^2}{2}\left(\frac{1}{a^2}\frac{\partial^2 \tilde{p}}{\partial \phi^2} + \frac{\partial^2 \tilde{p}}{\partial a^2}\right) \tag{5.3.14}$$

which (of course) is the FPE, corresponding to the two SDE's in Sect.4.4.5, which were derived by changing variables according to Ito's formula.

5.3.2 Boundary Conditions

We have already touched on boundary conditions in general in Sect.5.2.1 where they were considered in terms of probability current. The full range of boundary conditions for an arbitrary multidimensional Fokker-Planck equation does not seem to have been specified yet. In this book we shall therefore consider mostly *reflecting barrier* boundary conditions at a surface S, namely,

$$\boldsymbol{n}\cdot\boldsymbol{J} = 0 \quad \text{on } \boldsymbol{S}, \tag{5.3.15}$$

where $\boldsymbol{n}$ is the normal to the surface and

$$J_i(\boldsymbol{x}, t) = A_i(\boldsymbol{x}, t)p(\boldsymbol{x}, t) - \frac{1}{2}\sum_j \frac{\partial}{\partial x_j} B_{ij}(\boldsymbol{x}, t)p(\boldsymbol{x}, t) \tag{5.3.16}$$

and *absorbing barrier* boundary conditions

$$p(\boldsymbol{x}, t) = 0 \qquad \text{for } \boldsymbol{x} \text{ on } \boldsymbol{S}\,. \tag{5.3.17}$$

In practice, some part of the surface may be reflecting and another absorbing. At a surface S on which A_i or B_{ij} are discontinuous, we enforce

$$\begin{aligned} &\boldsymbol{n}\cdot\boldsymbol{J}_1 = \boldsymbol{n}\cdot\boldsymbol{J}_2 \qquad \text{on } \boldsymbol{S} \\ &p_1(\boldsymbol{x}) = p_2(\boldsymbol{x}) \qquad \boldsymbol{x} \text{ on } \boldsymbol{S}. \end{aligned} \tag{5.3.18}$$

The tangential current component is permitted to be discontinuous.

The boundary conditions on the backward equation have already been derived in Sect.5.2.4. For completeness, they are

$$\textit{Absorbing Boundary} \quad p(\boldsymbol{x}, t\,|\,\boldsymbol{y}, t') = 0 \qquad \boldsymbol{y} \in \boldsymbol{S} \tag{5.3.19}$$

$$\textit{Reflecting Boundary} \quad \sum_{i,j} n_i B_{ij}(\boldsymbol{y}) \frac{\partial}{\partial y_j} p(\boldsymbol{x}, t\,|\,\boldsymbol{y}, t') = 0 \qquad \boldsymbol{y} \in \boldsymbol{S}\,. \tag{5.3.20}$$

5.3.3 Stationary Solutions: Potential Conditions

A large class of interesting systems is described by Fokker-Planck equations which permit a stationary distribution for which the probability current vanishes for all $\boldsymbol{x}$ in R. Assuming this to be the case, by rearranging the definition of $\boldsymbol{J}$ (5.3.16), we obtain a completely equivalent equation

$$\frac{1}{2}\sum_j B_{ij}(\boldsymbol{x})\frac{\partial p_s(\boldsymbol{x})}{\partial x_j} = p_s(\boldsymbol{x})\left[A_i(\boldsymbol{x}) - \frac{1}{2}\sum_j \frac{\partial}{\partial x_j} B_{ij}(\boldsymbol{x})\right]. \tag{5.3.21}$$

If the matrix $B_{ij}(\boldsymbol{x})$ has an inverse for all $\boldsymbol{x}$, we can rewrite (5.3.21)

$$\frac{\partial}{\partial x_i}\log\left[p_s(\boldsymbol{x})\right] = \sum_k B^{-1}_{ik}(\boldsymbol{x})\left[2A_k(\boldsymbol{x}) - \sum_j \frac{\partial}{\partial x_j} B_{kj}(\boldsymbol{x})\right] \tag{5.3.22}$$

$$\equiv Z_i[A, B, \boldsymbol{x}]\,. \tag{5.3.23}$$

This equation cannot be satisfied for arbitrary $B_{ij}(\boldsymbol{x})$ and $A_i(\boldsymbol{x})$ since the left-hand side is explicitly a gradient. Hence, Z_i must also be a gradient, and a necessary and sufficient condition for that is the vanishing of the curl, i.e.,

$$\frac{\partial Z_i}{\partial x_j} = \frac{\partial Z_j}{\partial x_i}\,. \tag{5.3.24}$$

If this condition is satisfied, the stationary solution can be obtained by simple integration of (5.3.22):

$$p_s(\boldsymbol{x}) = \exp\{\int^{\boldsymbol{x}} d\boldsymbol{x}'\cdot \boldsymbol{Z}[A, B, \boldsymbol{x}']\}\,. \tag{5.3.25}$$

The conditions (5.3.24) are known as *potential conditions* since we derive the quantities Z_i from derivatives of $\log\left[p_s(\boldsymbol{x})\right]$, which, therefore, is often thought of as a potential $-\phi(\boldsymbol{x})$ so that more precisely,

$$p_s(\boldsymbol{x}) = \exp\left[-\phi(\boldsymbol{x})\right] \tag{5.3.26}$$

and

$$\phi(\boldsymbol{x}) = -\int^{\boldsymbol{x}} d\boldsymbol{x}'\cdot \boldsymbol{Z}[A, B, \boldsymbol{x}']\,. \tag{5.3.27}$$

Example: Rayleigh Process in Polar Coordinates. From (5.3.14) we find

$$\boldsymbol{A} = \begin{bmatrix} -\gamma a + \varepsilon^2/2a \\ 0 \end{bmatrix} \tag{5.3.28}$$

$$\underset{\sim}{\boldsymbol{B}} = \begin{bmatrix} \varepsilon^2 & 0 \\ 0 & \varepsilon^2/a^2 \end{bmatrix} \tag{5.3.29}$$

from which

$$\sum_j \frac{\partial}{\partial x_j} B_{a,j} = \sum_j \frac{\partial}{\partial x_j} B_{\phi,j} = 0 \tag{5.3.30}$$

so that

$$\boldsymbol{Z} = 2\,\underset{\sim}{\boldsymbol{B}}^{-1}\boldsymbol{A} = \begin{bmatrix} -2\gamma a/\varepsilon^2 + 1/a \\ 0 \end{bmatrix} \tag{5.3.31}$$

and clearly

$$\frac{\partial Z_a}{\partial \phi} = \frac{\partial Z_\phi}{\partial a} = 0 . \tag{5.3.32}$$

The stationary solution is then

$$p_s(a, \phi) = \exp\left[\int^{(a,\phi)} (d\phi\, Z_\phi + da\, Z_a)\right] \tag{5.3.33}$$

$$= \mathcal{N} \exp\left(-\frac{\gamma a^2}{\varepsilon^2} + \log a\right) \tag{5.3.34}$$

$$= \mathcal{N} a \exp\left(-\frac{\gamma a^2}{\varepsilon^2}\right). \tag{5.3.35}$$

5.3.4 Detailed Balance

a) Definition of Detailed Balance

The fact that the stationary solution of certain Fokker-Planck equations corresponds to a vanishing probability current is a particular version of the physical phenomenon of *detailed balance*. A Markov process satisfies detailed balance if, roughly speaking, in the stationary situation each possible transition balances with the reversed transition. The concept of detailed balance comes from physics, so let us explain more precisely with a physical example. We consider a gas of particles with positions $\boldsymbol{r}$ and velocities $\boldsymbol{v}$. Then a transition corresponds to a particle at some time t with position velocity $(\boldsymbol{r}, \boldsymbol{v})$ having acquired by a later time $t + \tau$ position and velocity $(\boldsymbol{r}', \boldsymbol{v}')$. The probability density of this transition is the *joint probability density* $p(\boldsymbol{r}', \boldsymbol{v}', t + \tau; \boldsymbol{r}, \boldsymbol{v}, t)$.

We may symbolically write this transition as

$$(\boldsymbol{r}, \boldsymbol{v}, t) \rightarrow (\boldsymbol{r}', \boldsymbol{v}', t + \tau) . \tag{5.3.36}$$

The reversed transition is not given simply by interchanging primed and unprimed quantities Rather, it is

$$(\boldsymbol{r}', -\boldsymbol{v}', t) \rightarrow (\boldsymbol{r}, -\boldsymbol{v}, t + \tau) . \tag{5.3.37}$$

It corresponds to the *time reversed transition* and requires the velocities to be reversed because the motion from $\boldsymbol{r}'$ to $\boldsymbol{r}$ is in the opposite direction from that from $\boldsymbol{r}$ to $\boldsymbol{r}'$.

The probability density for the reversed transition is thus the joint probability density

$$p(\boldsymbol{r}, -\boldsymbol{v}, t + \tau; \boldsymbol{r}', -\boldsymbol{v}', t) . \tag{5.3.38}$$

The *principle of detailed balance* requires the equality of these two joint probabilities when the system is in a stationary state. Thus, we may write

$$p_s(\boldsymbol{r}', \boldsymbol{v}', \tau; \boldsymbol{r}, \boldsymbol{v}, 0) = p_s(\boldsymbol{r}, -\boldsymbol{v}, \tau; \boldsymbol{r}', -\boldsymbol{v}', 0) \tag{5.3.39}$$

(The principle can be derived under certain conditions from the laws of physics, see [5.7] and Sect.5.3.6b.)

More explicitly, for a *Markov* process we can rewrite (5.3.39)

$$p(\boldsymbol{r}', \boldsymbol{v}', \tau | \boldsymbol{r}, \boldsymbol{v}, 0) p_s(\boldsymbol{r}, \boldsymbol{v}) = p(\boldsymbol{r}, -\boldsymbol{v}, \tau | \boldsymbol{r}', -\boldsymbol{v}', 0) p_s(\boldsymbol{r}', -\boldsymbol{v}') , \tag{5.3.40}$$

where the conditional probabilities now apply to the corresponding homogeneous Makov process (if the process was not Markov, the conditional probabilities would be for the stationary system only).

In its general form, detailed balance is formulated in terms of arbitrary variables x_i, which under time reyersal, transform to the reversed variables according to the rule

$$x_i \to \varepsilon_i x_i \tag{5.3.41}$$

$$\varepsilon_i = \pm 1 \tag{5.3.42}$$

depending on whether the variable is odd or even under time reversal. In the above, $\boldsymbol{r}$ is even, $\boldsymbol{v}$ is odd.

Then by detailed balance we require

$$p_s(\boldsymbol{x}, t + \tau; \boldsymbol{x}', t) = p_s(\boldsymbol{\varepsilon x}', t + \tau; \boldsymbol{\varepsilon x}, t) . \tag{5.3.43}$$

By $\boldsymbol{\varepsilon x}$, we mean $(\varepsilon_1 x_1, \varepsilon_2 x_2, \ldots)$.
Notice that setting $\tau = 0$ in (5.3.43) we obtain

$$\delta(\boldsymbol{x} - \boldsymbol{x}') p_s(\boldsymbol{x}') = \delta(\boldsymbol{\varepsilon x} - \boldsymbol{\varepsilon x}') p_s(\boldsymbol{\varepsilon x}) . \tag{5.3.44}$$

The two delta functions are equal since only sign changes are involved. Hence,

$$p_s(\boldsymbol{x}) = p_s(\boldsymbol{\varepsilon x}) \tag{5.3.45}$$

is a consequence of the formulation of detailed balance by (5.3.43). Rewriting now in terms of conditional probabilities, we have

$$\boxed{p(\boldsymbol{x}, \tau | \boldsymbol{x}', 0) p_s(\boldsymbol{x}') = p(\boldsymbol{\varepsilon x}', \tau | \boldsymbol{\varepsilon x}, 0) p_s(\boldsymbol{x}) .} \tag{5.3.46}$$

b) General Consequences of Detailed Balance

An important consequence of (5.3.45) is that

$$\langle \boldsymbol{x} \rangle_s = \boldsymbol{\varepsilon} \langle \boldsymbol{x} \rangle_s \tag{5.3.47}$$

(hence all odd variables have zero stationary mean), and for the autocorrelation function

$$G(\tau) \equiv \langle \boldsymbol{x}(\tau) \boldsymbol{x}^{\mathrm{T}}(0) \rangle_s$$

we have

$$G(\tau) = \varepsilon \langle x(0) x^{\mathrm{T}}(\tau) \rangle_s \varepsilon ,$$

hence,

$$\boxed{G(\tau) = \varepsilon G^{\mathrm{T}}(\tau) \varepsilon} \tag{5.3.48}$$

and setting $\tau = 0$ and noting that the covariance matrix σ satisfies $\sigma = \sigma^{\mathrm{T}}$,

$$\sigma \varepsilon = \varepsilon \sigma . \tag{5.3.49}$$

For the spectrum matrix

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}\omega\tau} G(\tau) ,$$

we find from (5.3.48) that

$$S(\omega) = \varepsilon S^{T}(\omega) \varepsilon . \tag{5.3.50}$$

c) Situations in Which Detailed Balance must be Generalised

It is possible that there exist several stationary solutions to a Markov process, and in this situation, a weaker form of detailed balance may hold, namely, instead of (5.3.43), we have

$$p_s^1(x, t + \tau; x', t) = p_s^2(\varepsilon x', t + \tau; \varepsilon x, t) \tag{5.3.51}$$

where the superscripts 1 and 2 refer to two different stationary solutions. Such a situation can exist if one of the variables is odd under time reversal, but does not change with time; for example, in a centrifuge the total angular momentum has this property. A constant magnetic field acts the same way.

Mostly, one writes the detailed balance conditions in such situations as

$$p_s^{\lambda}(x, t + \tau; x^1, t) = p_s^{\varepsilon\lambda}(\varepsilon x', + \tau; \varepsilon x, t) \tag{5.3.52}$$

where λ is a vector of such constant quantities, which change to $\varepsilon\lambda$ under time reversal. According to one point of view, such a situation does not represent detailed balance; since in a given stationary situation, the transitions do *not* balance in detail. It is perhaps better to call the property (5.3.52) *time reversal invariance.*

In the remainder of our considerations, we shall mean by detailed balance the situation (5.3.45), since no strong consequences arise from the form (5.3.52).

5.3.5 Consequences of Detailed Balance

The formulation of detailed balance for the Fokker-Planck equation was done by *van Kampen* [5.7] and independently by *Uhlhorn* [5.8], and *Graham* and *Haken*

[5.9]. We will formulate the conditions in a slightly more direct and more general way. We want necessary and sufficient conditions on the drift and diffusion coefficients and the jump probabilities for a homogeneous Markov process to have stationary solutions which satisfy detailed balance. We shall show that necessary and sufficient conditions are given by

$$\begin{aligned} &\text{(i)} \quad W(\boldsymbol{x}\,|\,\boldsymbol{x}')p_s(\boldsymbol{x}') = W(\varepsilon\boldsymbol{x}'\,|\,\varepsilon\boldsymbol{x})p_s(\boldsymbol{x}) \\ &\text{(ii)} \quad \varepsilon_i A_i(\varepsilon\boldsymbol{x})p_s(\boldsymbol{x}) = -A_i(\boldsymbol{x})p_s(\boldsymbol{x}) + \sum_j \frac{\partial}{\partial x_j}[B_{ij}(\boldsymbol{x})p_s(\boldsymbol{x})] \\ &\text{(iii)} \quad \varepsilon_i\varepsilon_j B_{ij}(\varepsilon\boldsymbol{x}) = B_{ij}(\boldsymbol{x})\,. \end{aligned} \tag{5.3.53}$$

The specialisation to a FPE is simply done by setting the jump probabilities $W(\boldsymbol{x}\,|\,\boldsymbol{x}')$ equal to zero.

Necessary Conditions. It is simpler to formulate conditions for the differential Chapman-Kolmogorov equation than to restrict ourselves to the Fokker-Planck equation. According to Sect. 3.4 which defines the quantities $W(\boldsymbol{x}\,|\,\boldsymbol{x}')$, $A_i(\boldsymbol{x})$ and $B_{ij}(\boldsymbol{x})$ (all of course being time independent, since we are considering homogeneous process), we have the trivial result that detailed balance requires, from (5.3.46)

$$W(\boldsymbol{x}\,|\,\boldsymbol{x}')p_s(\boldsymbol{x}') = W(\varepsilon\boldsymbol{x}'\,|\,\varepsilon\boldsymbol{x})p_s(\boldsymbol{x})\,. \tag{5.3.54}$$

Consider now the drift coefficient. For simplicity write

$$\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{\delta}\,. \tag{5.3.55}$$

Then from (5.3.46) we have

$$\begin{aligned} &\int_{|\delta|<K} d\boldsymbol{\delta}\,\delta_i p(\varepsilon\boldsymbol{x} + \varepsilon\boldsymbol{\delta}, \Delta t\,|\,\varepsilon\boldsymbol{x}, 0)p_s(\boldsymbol{x}) \\ &\quad = \int_{|\delta|<K} d\boldsymbol{\delta}\,\delta_i p(\boldsymbol{x}, \Delta t\,|\,\boldsymbol{x} + \boldsymbol{\delta}, 0)p_s(\boldsymbol{x} + \boldsymbol{\delta}) \end{aligned} \tag{5.3.56}$$

(we use K instead of ε in the range of integration to avoid confusion with ε_i); divide by Δt and take the limit $\Delta t \to 0$, and the left-hand side yields

$$\varepsilon_i A_i(\varepsilon\boldsymbol{x})p_s(\boldsymbol{x}) + O(K)\,. \tag{5.3.57}$$

On the right-hand side we write

$$\begin{aligned} p(\boldsymbol{x} + \boldsymbol{\delta} - \boldsymbol{\delta}, \Delta t\,|\,\boldsymbol{x} + \boldsymbol{\delta}, 0)p_s(\boldsymbol{x} + \boldsymbol{\delta}) &= p(\boldsymbol{x} - \boldsymbol{\delta}, \Delta t\,|\,\boldsymbol{x}, 0)p_s(\boldsymbol{x}) \\ &+ \sum_j \delta_j \frac{\partial}{\partial x_j}[p(\boldsymbol{x} - \boldsymbol{\delta}, \Delta t\,|\,\boldsymbol{x}, 0)p_s(\boldsymbol{x})] + O(\delta^2) \end{aligned} \tag{5.3.58}$$

so that the right-hand side is

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int d\boldsymbol{\delta}\left\{\delta_i p(\boldsymbol{x}-\boldsymbol{\delta},\Delta t\,|\,\boldsymbol{x},0)p_s(\boldsymbol{x})+\sum_j \delta_i\delta_j\frac{\partial}{\partial x_j}[p(\boldsymbol{x}-\boldsymbol{\delta},\Delta t\,|\,\boldsymbol{x},0)p_s(\boldsymbol{x})]\right\}+O(K)$$

$$=-A_i(\boldsymbol{x})p_s(\boldsymbol{x})+\sum_j\frac{\partial}{\partial x_j}[B_{ij}(\boldsymbol{x})p_s(\boldsymbol{x})]+O(K)\,, \tag{5.3.59}$$

where we have used the fact demonstrated in Sect. 3.4 that terms involving higher powers of δ than δ^2 are of order K. Letting $K\to 0$, we find

$$\varepsilon_i A_i(\varepsilon\boldsymbol{x})p_s(\boldsymbol{x})=-A_i(\boldsymbol{x})p_s(\boldsymbol{x})+\sum_j\frac{\partial}{\partial x_j}[B_{ij}(\boldsymbol{x})p_s(\boldsymbol{x})]\,. \tag{5.3.60}$$

The condition on $B_{ij}(\boldsymbol{x})$ is obtained similarly, but in this case no term like the second on the right of (5.3.60) arises, since the principal term is $O(\delta^2)$. We find

$$\varepsilon_i\varepsilon_j B_{ij}(\varepsilon\boldsymbol{x})=B_{ij}(\boldsymbol{x})\,. \tag{5.3.61}$$

A third condition is, of course, that $p_s(\boldsymbol{x})$ be a stationary solution of the differential Chapman-Kolmogorov equation. This is not a trivial condition, and is, in general, independent of the others.

Sufficient Conditions. We now show that (5.3.53) are sufficient. Assume that these conditions are satisfied, that $p_s(\boldsymbol{x})$ is a stationary solution of the differential Chapman-Kolmogorov equation, and that $p(\boldsymbol{x},t\,|\,\boldsymbol{x}',0)$ is a solution of the differential Chapman-Kolmogorov equation. We now consider a quantity

$$\hat{p}(\boldsymbol{x},t\,|\,\boldsymbol{x}',0)\equiv p(\varepsilon\boldsymbol{x}',t\,|\,\varepsilon\boldsymbol{x},0)p_s(\boldsymbol{x})/p_s(\boldsymbol{x}')\,. \tag{5.3.62}$$

Clearly

$$\hat{p}(\boldsymbol{x},0\,|\,\boldsymbol{x}',0)=\delta(\boldsymbol{x}-\boldsymbol{x}')=p(\boldsymbol{x},0\,|\,\boldsymbol{x}',0)\,. \tag{5.3.63}$$

We substitute $\hat{p}$ into the differential Chapman-Kolmogorov equation and show that because $p(\boldsymbol{x}',t\,|\,\boldsymbol{x},0)$ obeys the *backward differential Chapman-Kolmogorov equation* in the variable $\boldsymbol{x}$, the quantity $\hat{p}$ is a solution of the *forward differential Chapman-Kolmogorov equation.*

We do this explicitly. The notation is abbreviated for clarity, so that we write

$$\left.\begin{array}{lll}\hat{p} & \text{for} & \hat{p}(\boldsymbol{x},t\,|\,\boldsymbol{x}',0)\\ p_s & \text{for} & p_s(\boldsymbol{x})\\ p_s' & \text{for} & p_s(\boldsymbol{x}')\\ p(\boldsymbol{x}) & \text{for} & p(\boldsymbol{x}',t\,|\,\boldsymbol{x},0)\,.\end{array}\right\} \tag{5.3.64}$$

We proceed term by term.

i) **Drift Term:**

$$-\sum_i\frac{\partial}{\partial x_i}(A_i\hat{p})=-\sum_i\frac{\partial}{\partial x_i}(A_i p(\varepsilon\boldsymbol{x})p_s/p_s') \tag{5.3.65}$$

$$= -\sum_i \left\{ \frac{\partial}{\partial x_i} [A_i p_s] p(\varepsilon x) + A_i p_s \frac{\partial}{\partial x_i} p(\varepsilon x) \right\} / p_s' \,.$$

ii) **Diffusion Term**:

$$\frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (B_{ij} \hat{p}) = \frac{1}{2} \sum_{i,j} \left[\frac{\partial^2}{\partial x_i \partial x_j} (B_{ij} p_s) p(\varepsilon x) \right.$$
$$\left. + 2 \frac{\partial}{\partial x_i} (B_{ij} p_s) \frac{\partial}{\partial x_j} p(\varepsilon x) + B_{ij} p_s \frac{\partial^2}{\partial x_i \partial x_j} p(\varepsilon x) \right] / p_s' \,. \tag{5.3.66}$$

iii) **Jump Term**:

$$\int dz [W(x \mid z) \hat{p}(z, t \mid x', 0) - W(z \mid x) \hat{p}(x, t \mid x', 0)]$$
$$= \int dz [W(x \mid z) p_s(z) p(\varepsilon x', t \mid \varepsilon z, 0) - W(z \mid x) p_s(x) p(\varepsilon x', t \mid \varepsilon x, 0)] / p_s \,. \tag{5.3.67}$$

We now use the fact that $p_s(x)$ is a solution of the stationary differential Chapman-Kolmogorov equation to write

$$-\sum_i \left[\frac{\partial}{\partial x_i} (A_i p_s) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_i} (B_{ij} p_s) \right] - \int dz \; W(z \mid x) p_s(x)$$
$$= -\int dz \; W(x \mid z) p_s(z) \tag{5.3.68}$$

and using the detailed balance condition (5.3.53(i)) for W

$$= -\int dz \; W(\varepsilon z \mid \varepsilon x) p_s(x) \,. \tag{5.3.69}$$

Now substitute

$$y = \varepsilon x \tag{5.3.70}$$
$$y' = \varepsilon x'$$

and all up all three contributions, taking care of (5.3.68, 69):

$$= \left\{ -\sum_i \varepsilon_i A_i(\varepsilon y) p_s(y) \left[\frac{\partial}{\partial y_i} p(y) \right] + \sum_{i,j} \varepsilon_i \varepsilon_j \frac{\partial}{\partial y_j} [B_{ij}(\varepsilon y) p_s(y)] \left[\frac{\partial}{\partial y_i} p(y) \right] \right.$$
$$+ \frac{1}{2} \sum_{i,j} \varepsilon_i \varepsilon_j B_{ij}(\varepsilon y) p_s(y) \left[\frac{\partial^2}{\partial y_i \partial y_j} p(y) \right] \tag{5.3.71}$$
$$\left. + \int dz [W(\varepsilon y \mid z) p_s(z) p(y', t \mid \varepsilon z, 0) - W(\varepsilon y \mid z) p_s(z) p(y', t \mid y, 0)] \right\} / p_s(y') .$$

We now substitute the detailed balance conditions (5.3.53).

$$= \left\{ \sum_i A_i(y) \frac{\partial}{\partial y_i} p(y', t \mid y, 0) + \frac{1}{2} \sum_{i,j} B_{ij}(y) \frac{\partial^2}{\partial y_i \partial y_j} p(y', t \mid y, 0) \right.$$
$$\left. + \int dz [W(z \mid y) p(y', t \mid z, 0) - W(z \mid y) p(y', t \mid y, 0)] \right\} p_s(y) / p_s(y') \,. \tag{5.3.72}$$

The term in the large curly brackets is now recognisable as the *backward differential Chapman-Kolmogorov operator* [Sect.3.6, (3.6.4)]. Note that the process is homogeneous, so that

$$p(\boldsymbol{y}', t\,|\,\boldsymbol{y}, 0) = p(\boldsymbol{y}, 0\,|\,\boldsymbol{y}, -t)\,.$$

We see that

$$(5.3.72) = \frac{\partial}{\partial t}[p(\boldsymbol{y}', t\,|\,, 0)p_s(\boldsymbol{y})/p_s(\boldsymbol{y}')] = \frac{\partial}{\partial t}\hat{p}(\boldsymbol{x}, t\,|\,\boldsymbol{x}', 0) \tag{5.3.73}$$

which means that $\hat{p}(\boldsymbol{x}, t\,|\,\boldsymbol{x}', 0)$, defined in (5.3.62), satisfies the forward differential Chapman-Kolmogorov equation. Since the initial condition of $p(\boldsymbol{x}, t\,|\,\boldsymbol{x}', 0)$ and $\hat{p}(\boldsymbol{x}, t\,|\,\boldsymbol{x}', 0)$ at $t = 0$ are the same (5.3.63) and the solutions are unique, we have shown that provided the detailed balance conditions (5.3.53) are satisfied, detailed balance is satisfied. Hence, sufficiency is shown.

Comments

i) *Even variables only*: the conditions are considerably simpler if all ε_i are $+1$. In this case, the conditions reduce to

$$W(\boldsymbol{x}\,|\,\boldsymbol{x}')p_s(\boldsymbol{x}') = W(\boldsymbol{x}'\,|\,\boldsymbol{x})p_s(\boldsymbol{x}) \tag{5.3.74}$$

$$A_i(\boldsymbol{x})p_s(\boldsymbol{x}) = \frac{1}{2}\sum_j \frac{\partial}{\partial x_j}[B_{ij}(\boldsymbol{x})p_s(\boldsymbol{x})] \tag{5.3.75}$$

$$B_{ij}(\boldsymbol{x}) = B_{ij}(\boldsymbol{x})\,, \tag{5.3.76}$$

the last of which is trivial. The condition (5.3.75) is exactly the same as the potential condition (5.3.21) which expresses the vanishing of $\boldsymbol{J}$, the probability current in the stationary state.

The conditions (5.3.74, 75) taken together imply that $p_s(\boldsymbol{x})$ satisfies the stationary differential Chapman-Kolmogorov equation, which is not the case for the general conditions (5.3.53).

ii) *Fokker-Planck equations*: *van Kampen*, [5.7], and *Graham* and *Haken* [5.9] introduced the concept of reversible and irreversible drift parts. The irreversible drift is

$$D_i(\boldsymbol{x}) = \tfrac{1}{2}[A_i(\boldsymbol{x}) + \varepsilon_i A_i(\varepsilon\boldsymbol{x})] \tag{5.3.77}$$

and the reversible drift

$$I_i(\boldsymbol{x}) = \tfrac{1}{2}[A_i(\boldsymbol{x}) - \varepsilon_i A_i(\varepsilon\boldsymbol{x}]\,. \tag{5.3.78}$$

Using again the potential defined by

$$p_s(\boldsymbol{x}) = \exp[-\phi(\boldsymbol{x})]\,, \tag{5.3.79}$$

we see that in the case of a Fokker-Planck equation, we can write the conditions for detailed balance as

$$\varepsilon_i \varepsilon_j B_{ij}(\varepsilon \boldsymbol{x}) = B_{ij}(\boldsymbol{x}) \tag{5.3.80}$$

$$D_i(\boldsymbol{x}) - \frac{1}{2}\sum_j \frac{\partial}{\partial x_j}[B_{ij}(\boldsymbol{x})] = -\frac{1}{2}\sum_i B_{ij}(\boldsymbol{x})\frac{\partial \phi(\boldsymbol{x})}{\partial x_j} \tag{5.3.81}$$

$$\sum_i \left[\frac{\partial}{\partial x_i} I_i(\boldsymbol{x}) - I_i(\boldsymbol{x})\frac{\partial \phi(\boldsymbol{x})}{\partial x_i}\right] = 0 \tag{5.3.82}$$

where the last equation is simply the stationary FPE for $p_s(\boldsymbol{x})$, after substituting (3.3.53(i)). As was the case for the potential conditions, it can be seen that (5.3.81) gives an equation for $\partial\phi/\partial x_j$ which can only be satisfied provided certain conditions on $D_i(\boldsymbol{x})$ and $B_{ij}(\boldsymbol{x})$ are satisfied. If $B_{ij}(\boldsymbol{x})$ has an inverse, these take the form

$$\frac{\partial \hat{Z}_i}{\partial x_j} = \frac{\partial \hat{Z}_j}{\partial x_i}, \tag{5.3.83}$$

where

$$\hat{Z}_i = \sum_k B^{-1}_{ik}(\boldsymbol{x})\left[2D_k(\boldsymbol{x}) - \sum_j \frac{\partial}{\partial x_j} B_{kj}(\boldsymbol{x})\right] \tag{5.3.84}$$

and we have

$$p_s(\boldsymbol{x}) = \exp[-\phi(\boldsymbol{x})] = \exp\left(\int^{\boldsymbol{x}} d\boldsymbol{x}' \cdot \hat{Z}\right). \tag{5.3.85}$$

Thus, as in the case of a vanishing probability current, $p_s(\boldsymbol{x})$ can be determined explicitly as an integral.

iii) *Connection between backward and forward operators* of differential Chapman-Kolmogorov equations is provided by the detailed balance. The proof of sufficient conditions amounts to showing that if $f(\boldsymbol{x}, t)$ is a solution of the forward differential Chapman-Kolmogorov equation, then

$$\tilde{f}(\boldsymbol{x}, t) = f(\varepsilon \boldsymbol{x}, -t)/p_s(\boldsymbol{x}) \tag{5.3.86}$$

is a solution of the backward differential Chapman-Kolmogorov equation. This relationship will be used in Sect.5.3.7 for the construction of eigenfunctions.

5.3.6 Examples of Detailed Balance in Fokker-Planck Equations

a) Kramers' Equation for Brownian Motion [5.10]

We take the motion of a particle in a fluctuating environment. The motion is in one dimension and the state of the particle is described by its position x and velocity v. This gives the differential equations

$$\frac{dx}{dt} = v \tag{5.3.87}$$

and

$$m\frac{dv}{dt} = -V'(x) - \beta v + \sqrt{2\beta kT}\,\xi(t) \tag{5.3.88}$$

which are essentially Langevin's equations (1.2.14) in which for brevity, we write

$$6\pi\eta a = \beta$$

and $V(x)$ is a potential whose gradient $V'(x)$ gives rise to a force on the particle. By making the assumption that the physical fluctuating force $\xi(t)$ is to be interpreted as

$$\xi(t)dt = dW(t) \tag{5.3.89}$$

as explained in Sect.4.1, we obtain SDE's

$$dx = v\ dt \tag{5.3.90}$$

$$m\ dv = -[V'(x) + \beta v]\,dt + \sqrt{2\beta\, kT}\,dW(t) \tag{5.3.91}$$

for which the corresponding FPE is

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(vp) + \frac{1}{m}\frac{\partial}{\partial v}\{[V'(x) + \beta v]p\} + \frac{\beta\, kT}{m^2}\frac{\partial^2 p}{\partial v^2}. \tag{5.3.92}$$

The equation can be slightly simplified by introducing new scaled variables

$$y = x\sqrt{m/kT} \tag{5.3.93}$$

$$u = v\sqrt{m/kT} \tag{5.3.94}$$

$$U(y) = V(x)/kT \tag{5.3.95}$$

$$\gamma = \beta/m \tag{5.3.96}$$

so that the FPE takes the form

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial y}(up) + \frac{\partial}{\partial u}[U'(y)p] + \gamma\frac{\partial}{\partial u}\left(up + \frac{\partial p}{\partial u}\right) \tag{5.3.97}$$

which we shall call *Kramers' equation.*

Here, y (the position) is an even variable and u (the velocity) an odd variable, as explained in Sect.5.3.4. The drift and diffusion can be written

$$A(y, u) = \begin{bmatrix} u \\ -U'(y) - \gamma u \end{bmatrix}, \tag{5.3.98}$$

$$\underset{\sim}{B}(y, u) = \begin{bmatrix} 0 & 0 \\ 0 & 2\gamma \end{bmatrix} \tag{5.3.99}$$

and

$$\varepsilon \begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} y \\ -u \end{bmatrix}. \tag{5.3.100}$$

We can check the conditions one by one.

The condition (5.3.53(iii)) is trivially satisfied. The condition (5.3.53(ii)) is somewhat degenerate, since $\underset{\sim}{B}$ is not invertible. It can be written

$$\varepsilon A(y, -u)p_s(y, u) = -A(y, u)p_s(y, u) + \begin{bmatrix} 0 \\ 2\gamma \dfrac{\partial p_s}{\partial u} \end{bmatrix} \tag{5.3.101}$$

or, more fully

$$\begin{bmatrix} -u \\ U'(y) - \gamma u \end{bmatrix} p_s = \begin{bmatrix} -u \\ U'(y) + \gamma u \end{bmatrix} p_s + \begin{bmatrix} 0 \\ 2\gamma \dfrac{\partial p_s}{\partial u} \end{bmatrix}. \tag{5.3.102}$$

The first line is an identity and the second states

$$-up_s(y, u) = \frac{\partial p_s}{\partial u}, \tag{5.3.103}$$

i.e.,

$$p_s(y, u) = \exp(-\tfrac{1}{2}u^2)f(y) \tag{5.3.104}$$

which means that if $p_s(y, u)$ is written in the form (5.3.104), then the detailed balance conditions are satisfied. One must now check whether (5.3.104) indeed gives a stationary solution of Kramers' equation (5.3.97) by substitution. The final bracket vanishes, leaving

$$0 = -u\frac{\partial f}{\partial y} - U'(y)uf \tag{5.3.105}$$

which means

$$f(y) = \mathcal{N} \exp[-U(y)] \tag{5.3.106}$$

and

$$p_s(y, u) = \mathcal{N} \exp\left[-U(y) - \tfrac{1}{2}u^2\right] . \tag{5.3.107}$$

In terms of the original (x, v) variables,

$$p_s(x, v) = \mathcal{N} \exp\left[-\frac{V(x)}{kT} - \frac{mv^2}{2kT}\right] \tag{5.3.108}$$

which is the familiar Boltzmann distribution of statistical mechanics. Notice that the denominators kT arise from the assumed coefficient $\sqrt{2\beta kT}$ of the fluctuating force in (5.3.88). Thus, we take the macroscopic equations and add a fluctuating force, whose magnitude is fixed by the requirement that the solution be the Boltzmann distribution corresponding to the temperature T.

But we have also achieved exactly the right distribution function. This means that the assumption that Brownian motion is described by a *Markov* process of the form (5.3.87, 88) must have considerable validity.

b) Deterministic Motion

Here we have $B_{ij}(\boldsymbol{x})$ and $W(\boldsymbol{x}|\boldsymbol{x}')$ equal to zero, so the detailed balance conditions are simply

$$\varepsilon_i A_i(\varepsilon \boldsymbol{x}) = -A_i(\boldsymbol{x}) . \tag{5.3.109}$$

Since we are now dealing with a Liouville equation (Sect.3.5.3), the motion of a point whose coordinates are $\boldsymbol{x}$ is described by the ordinary differential equation

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}[\boldsymbol{x}(t)] . \tag{5.3.110}$$

Suppose a solution of (5.3.110) which passes through the point $\boldsymbol{y}$ at $t = 0$ is

$$\boldsymbol{q}[t, \boldsymbol{y}] \tag{5.3.111}$$

which therefore satisfies

$$\boldsymbol{q}[0, \boldsymbol{y}] = \boldsymbol{y} . \tag{5.3.112}$$

Then the relation (5.3.109) implies that the reversed solution

$$\varepsilon \boldsymbol{q}(-t, \varepsilon \boldsymbol{y}) \tag{5.3.113}$$

is also a solution of (5.3.110), and since

$$\varepsilon \boldsymbol{q}(0, \varepsilon \boldsymbol{y}) = \varepsilon\varepsilon \boldsymbol{y} = \boldsymbol{y} , \tag{5.3.114}$$

i.e., the initial conditions are the same, these solutions must be identical, i.e.,

$$\varepsilon \boldsymbol{q}(-t, \varepsilon \boldsymbol{y}) = \boldsymbol{q}(t, \boldsymbol{y}) . \tag{5.3.115}$$

Now the joint probability in the stationary state can be written as

$$p_s(\boldsymbol{x}, t; \boldsymbol{x}', t') = \int d\boldsymbol{y}\, p_s(\boldsymbol{x}, t; \boldsymbol{x}', t'; \boldsymbol{y}, 0)$$
$$= \int d\boldsymbol{y}\, \delta[\boldsymbol{x} - \boldsymbol{q}(t, \boldsymbol{y})]\delta[\boldsymbol{x}' - \boldsymbol{q}(t', \boldsymbol{y})]p_s(\boldsymbol{y}) \tag{5.3.116}$$

and

$$p_s(\varepsilon\boldsymbol{x}', -t'; \varepsilon\boldsymbol{x}, -t) = \int d\boldsymbol{y}\, \delta[\varepsilon\boldsymbol{x} - \boldsymbol{q}(-t, \boldsymbol{y})]\delta[\varepsilon\boldsymbol{x}' - \boldsymbol{q}(-t', \boldsymbol{y})]p_s(\boldsymbol{y}) . \tag{5.3.117}$$

Change the variables from $\boldsymbol{y}$ to $\varepsilon\boldsymbol{y}$ and note that $p_s(\boldsymbol{y}) = p_s(\varepsilon\boldsymbol{y})$, and $d\varepsilon\boldsymbol{y} = d\boldsymbol{y}$, so that

$$(5.3.117) = \int d\boldsymbol{y}\, \delta[\boldsymbol{x} - \varepsilon\boldsymbol{q}(-t, \varepsilon\boldsymbol{y})]\delta[\boldsymbol{x}' - \varepsilon\boldsymbol{q}(-t', \varepsilon\boldsymbol{y})]p_s(\boldsymbol{y}) \tag{5.3.118}$$

and using (5.3.115),

$$= \int d\boldsymbol{y}\, \delta[\boldsymbol{x} - \boldsymbol{q}(t, \boldsymbol{y})]\delta[\boldsymbol{x}' - \boldsymbol{q}(t', \boldsymbol{y})]p_s(\boldsymbol{y}) \tag{5.3.119}$$
$$= p_s(\boldsymbol{x}, t; \boldsymbol{x}', t') . \tag{5.3.120}$$

Using the stationarity property, that p_s depends only on the time difference, we see that detailed balance is satisfied.

This direct proof is, of course, unnecessary since the original general proof is valid for this deterministic system. Furthermore, any system of deterministic first-order differential equations can be transformed into a Liouville equation, so this direct proof is in general unnecessary and it is included here merely as a matter of interest.

However, it is important to give a brief summary of the philosophy behind this demonstration of detailed balance. In physical systems, which are where detailed balance is important, we often have an unbelievably large number of variables, of the order of 10^{20} at least. These variables (say, momentum and velocity of the particles in a gas) are those which occur in the distribution function which obeys a Liouville equation for they follow deterministic equations of motion, like Newton's laws of motion.

It can be shown directly that, for appropriate forms of interaction, Newton's laws obey the principle of *microscopic reversibility* which means that they can be put in the form (5.3.110), where $\boldsymbol{A}(\boldsymbol{x})$ obeys the reversibility condition (5.3.109).

The macroscopically observable quantities in such a system are functions of these variables (for example, pressure, temperature, density of particles) and, by appropriate changes of variable, can be represented by the first few components of the vector $\boldsymbol{x}$.

Thus, we assume $\boldsymbol{x}$ can be written

$$\boldsymbol{x} = (\boldsymbol{a}, \hat{\boldsymbol{x}}) \tag{5.3.121}$$

where the vector $\boldsymbol{a}$ represents the macroscopically observable quantities and $\hat{\boldsymbol{x}}$ is all the others. Then, in practice, we are interested in

$$\tilde{p}(\boldsymbol{a}_1, t_1; \boldsymbol{a}_2, t_2; \boldsymbol{a}_3, t_3; \ldots)$$
$$= \int\int \ldots \int d\hat{\boldsymbol{x}}_1, d\hat{\boldsymbol{x}}_2 \ldots p(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2; \boldsymbol{x}_3, t_3; \ldots) . \tag{5.3.122}$$

From the microscopic reversibility, it follows from our reasoning above that p, and thus also $\tilde{p}$, both obey the detailed balance conditions but, of course, $\tilde{p}$ does not obey a Liouville equation. *If it turns out or can be proven that $\tilde{p}$ obeys, to some degree approximation, a Markov equation of motion*, then we must preserve the detailed balance property, which takes the same form for $\tilde{p}$ as for p. In this sense, the condition (5.3.43) for detailed balance may be said to be derived from microscopic reversibility of the equations of motion.

c) Ornstein-Uhlenbeck Process: Onsager Relations in Linear Systems

Most systems in which detailed balance is of interest can be approximated by an Ornstein-Uhlenbeck process, i.e., this means we assume

$$A_i(\boldsymbol{x}) = \sum_j A_{ij} x_j \tag{5.3.123}$$

$$B_{ij}(\boldsymbol{x}) = B_{ij} . \tag{5.3.124}$$

The detailed balance conditions are not trivial, however. Namely,

$$\sum_j (\varepsilon_i \varepsilon_j A_{ij} + A_{ij}) x_j = \sum_j B_{ij} \frac{\partial}{\partial x_j} \log p_s(\boldsymbol{x}) \tag{5.3.125}$$

and

$$\varepsilon_i \varepsilon_j B_{ij} = B_{ij} . \tag{5.3.126}$$

Equation (5.3.125) has the qualitative implication that $p_s(\boldsymbol{x})$ is a Gaussian since derivative of $\log p_s(\boldsymbol{x})$ is linear in $\boldsymbol{x}$. Furthermore, since the left-hand side contains no constant term, this Gaussian must have zero mean, hence, we can write

$$p_s(\boldsymbol{x}) = \mathcal{N} \exp\left(-\tfrac{1}{2} \boldsymbol{x}^{\mathrm{T}} \sigma^{-1} \boldsymbol{x}\right) . \tag{5.3.127}$$

One can now substitute (5.3.127) in the stationary Fokker-Planck equation and rearrange to obtain

$$-\sum_i A_{ii} - \tfrac{1}{2} \sum_{i,j} B_{ij} \sigma_{ij}^{-1} + \sum_{k,j} \Big(\sum_i \sigma_{ki}^{-1} A_{ij} + \tfrac{1}{2} \sum_{i,l} \sigma_{ki}^{-1} B_{il} \sigma_{lj}^{-1} \Big) x_k x_j = 0 \tag{5.3.128}$$

(we have used the symmetry of the matrix σ). The quadratic term vanishes if the symmetric part of its coefficient is zero. This condition may be written in matrix form as

$$\sigma^{-1} A + A^{\mathrm{T}} \sigma^{-1} = -\sigma^{-1} B \sigma^{-1} \tag{5.3.129}$$

or

$$A\sigma + \sigma A^{\mathrm{T}} = -B . \tag{5.3.130}$$

The constant term also vanishes if (5.3.129) is satisfied. Equation (5.3.130) is, of

course, exactly that derived by SDE techniques in Sect. 4.4.6 (4.4.51) with the substitutions

$$A \to -A \tag{5.3.131}$$
$$BB^{\mathrm{T}} \to B\,.$$

We can now write the detailed balance conditions in their most elegant form. We define the matrix ε by

$$\varepsilon = \mathrm{diag}\,(\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots) \tag{5.3.132}$$

and clearly

$$\varepsilon^2 = 1\,. \tag{5.3.133}$$

Then the conditions (5.3.125, 126) become in matrix notation

$$\varepsilon A \varepsilon + A = -B\sigma^{-1} \tag{5.3.134}$$

$$\varepsilon B \varepsilon = B\,. \tag{5.3.135}$$

The potential condition (5.3.83) is simply equivalent to the symmetry of σ.

As noted in Sect.5.3.4 (5.3.49), detailed balance requires

$$\varepsilon\sigma = \sigma\varepsilon\,. \tag{5.3.136}$$

Bearing this in mind, we take (5.3.130)

$$A\sigma + \sigma A^{\mathrm{T}} = -B$$

and from (5.3.134)

$$\varepsilon A \varepsilon \sigma + A\sigma = -B \tag{5.3.137}$$

which yield

$$\varepsilon A \varepsilon \sigma = \sigma A^{\mathrm{T}} \tag{5.3.138}$$

and with (5.3.136)

$$\varepsilon(A\sigma) = (A\sigma)^{\mathrm{T}}\varepsilon. \tag{5.3.139}$$

These are the celebrated Onsager relations; *Onsager*, [5.11]; *Casimir*, [5.12]. The derivation closely follows *van Kampen*'s [5.6] work.

The interpretation can be made simpler by introducing the *phenomenological forces* defined as the gradient of the potential $\phi = \log[p_s(\boldsymbol{x})]$:

$$\boldsymbol{F}(\boldsymbol{x}) = -\nabla\phi(\boldsymbol{x}) = \sigma^{-1}\boldsymbol{x} \tag{5.3.140}$$

(in physics, ϕ/kT is the entropy of the system). Because of the linear form of the $A_i(\boldsymbol{x})$ [(5.3.123)], the exact equations of motion for $\langle \boldsymbol{x} \rangle$ are

$$\frac{d}{dt}\langle \boldsymbol{x} \rangle = A\langle \boldsymbol{x} \rangle = A\sigma \boldsymbol{F}(\langle \boldsymbol{x} \rangle) . \tag{5.3.141}$$

Thus, if the *fluxes* $d/dt\langle \boldsymbol{x} \rangle$ are related linearly to the *forces* $\boldsymbol{F}(\langle \boldsymbol{x} \rangle)$ by a matrix L defined by

$$L = A\sigma , \tag{5.3.142}$$

then (5.3.139) says

$$\varepsilon L \varepsilon = L^{\mathrm{T}} \tag{5.3.143}$$

or

$$\boxed{\begin{array}{ll} L_{ij} = L_{ji} & (\varepsilon_i \text{ and } \varepsilon_j \text{ of the same sign}) \\ L_{ij} = -L_{ji} & (\varepsilon_i \text{ and } \varepsilon_j \text{ of different sign}) . \end{array}} \tag{5.3.144}$$

Notice also that

$$\left.\begin{array}{l} \varepsilon B \varepsilon = B \\ \text{and} \\ \varepsilon \sigma \varepsilon = \sigma \end{array}\right\} \tag{5.3.145}$$

imply that B_{ij} and σ_{ij} vanish if ε_i and ε_j have opposite signs.

In the special case that all the ε_i have the same sign, we find that

$$L = L^{\mathrm{T}} \tag{5.3.146}$$

and noting that, since σ is symmetric and positive definite it has a *real* square root $\sigma^{1/2}$, we find that

$$\hat{A} \equiv \sigma^{-1/2} A \sigma^{1/2} . \tag{5.3.147}$$

is symmetric, so that A is similar to a symmetric matrix. Hence, all the eigenvalues of A are real.

d) Significance of the Onsager Relations: Fluctuation Dissipation Theorem

The Onsager relations are for a set of macroscopically observable quantities and thus provide an easily observed consequence of detailed balance, which is itself a consequence of the reversibility of microscopic equations of motion, as outlined in (b) above. However, to check the validity in a given situation requires a knowledge of the covariance matrix σ.

Fortunately, in such situations, statistical mechanics gives us the form of the stationary distribution, provided this is thermodynamic equilibrium in which de-

tailed balance is always satisfied. The principle is similar to that used by Langevin (Sect.1.2.2). We shall illustrate with an example; the derivation of Nyquist's formula

We assume an electric circuit as in Fig. 5.4 in which there is assumed to be a fluctuating voltage arising from the system having a nonzero temperature, and a fluctuating charge, which arises because the system is attached to a large neutral charge reservoir, e.g., an ionic solution. The electrical equations arise from conservation of an electric charge q and Kirchoff's voltage law. The charge equation takes the form

$$\frac{dq}{dt} = i - \gamma q + \Delta q(t) \tag{5.3.148}$$

in which we equate the rate of gain of charge on the capacitor to the current i, less a leakage term γq into the reservoir, plus a fluctuating term $\Delta q(t)$, which arises from the reservoir, and whose magnitude will be shortly calculated.

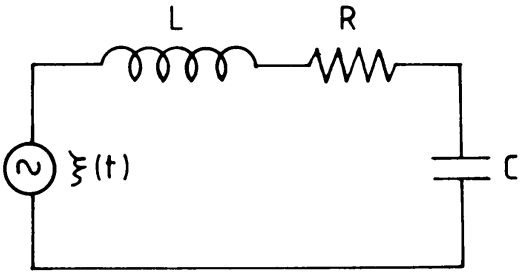


Fig. 5.4. Electric circuit used in the derivation of Nyquist's formula

Kirchoff's voltage law is obtained by adding up all the voltages around the circuit, including a possible fluctuating voltage $\Delta V(t)$:

$$\frac{di}{dt} = \frac{1}{L}\left[-\frac{q}{C} - iR + \Delta V(t)\right]. \tag{5.3.149}$$

We now assume that $\Delta q(t)$ and $\Delta V(t)$ are white noise. We can write in the most general case,

$$\Delta q(t) = b_{11}\xi_1(t) + b_{12}\xi_2(t) \tag{5.3.150}$$

$$\frac{1}{L}\Delta V(t) = b_{21}\xi_1(t) + b_{22}\xi_2(t) \tag{5.3.151}$$

in which $\xi_1(t)$ and $\xi_2(t)$ are uncorrelated Langevin sources, i.e.,

$$\xi_1(t)dt = dW_1(t) \tag{5.3.152}$$

$$\xi_2(t)dt = dW_2(t)\,, \tag{5.3.153}$$

where $W_1(t)$ and $W_2(t)$ are independent Wiener processes.

Thus, here

$$A = -\begin{bmatrix} \gamma & -1 \\ \frac{1}{LC} & \frac{R}{L} \end{bmatrix}. \tag{5.3.154}$$

The total energy in the system is

$$E = \frac{L}{2} i^2 + \frac{1}{2C} q^2 . \tag{5.3.155}$$

From statistical mechanics, we know that $p_s(q, i)$ is given by the *Boltzmann Distribution* at temperature T, i.e.,

$$\begin{aligned} p_s(q, i) &= \mathcal{N} \exp(-E/kT) \\ &= \mathcal{N} \exp\left(-\frac{Li^2}{2kT} - \frac{q^2}{2CkT}\right), \end{aligned} \tag{5.3.156}$$

where k is the Boltzmann constant, so that the covariance matrix is

$$\sigma = \begin{bmatrix} kTC & 0 \\ 0 & kT/L \end{bmatrix}. \tag{5.3.157}$$

The Onsager symmetry can now be checked:

$$A\sigma = -\begin{bmatrix} \gamma kTC & -kT/L \\ kT/L & RkT/L^2 \end{bmatrix}. \tag{5.3.158}$$

For this system q, the total charge is even under time inversion and i, the current, is odd. Thus,

$$\varepsilon = \text{diag}\,(1, -1) \tag{5.3.159}$$

and it is clear that

$$(A\sigma)_{12} = -(A\sigma)_{21} \tag{5.3.160}$$

is the only consequence of the symmetry and is satisfied by (5.3.158).

Here, also,

$$B = -(A\sigma + \sigma A^{\mathrm{T}}) = 2kT \begin{bmatrix} \gamma C & 0 \\ 0 & R/L^2 \end{bmatrix} \tag{5.3.161}$$

so that

$$b_{12} = b_{21} = 0 \tag{5.3.162}$$

$$b_{11} = \sqrt{2kT\gamma C} \tag{5.3.163}$$

$$b_{22} = \sqrt{2kTR/L}$$

and we see that $B_{12} = B_{21} = 0$, as required by physical intuition, which suggests that the two sources of fluctuations arise from different causes and should be independent.

These results are *fluctuation dissipation* results. The magnitudes of the fluctuations b_{ij} are determined by the *dissipative terms* γ and R. In fact, the result (5.3.163) is precisely Nyquist's theorem, which we discussed in Sect. 1.4.4. The noise voltage in the circuit is given by

$$\Delta V(t) = \sqrt{2kTR}\,\xi_1(t) \tag{5.3.164}$$

so that

$$\langle \Delta V(t+\tau)\Delta V(t)\rangle = 2kTR\delta(\tau) \tag{5.3.165}$$

which is Nyquists' theorem in the form quoted in (1.4.49).

The terms γ and R are called dissipative because they give rise to energy dissipation; in fact, deterministically (i.e. setting noise equal to 0),

$$\frac{dE}{dt} = -i^2R - \frac{\gamma q^2}{C} \tag{5.3.166}$$

which explicitly exhibits the dissipation.

5.3.7 Eigenfunction Methods in Many Variables—Homogeneous Processes

Here we shall proceed similarly to Sect.5.2.5. We *assume* the existence of a complete set of eigenfunctions $P_\lambda(\boldsymbol{x})$ of the forward Fokker-Planck and a set $Q_\lambda(\boldsymbol{x})$ of the backward Fokker-Planck equation. Thus,

$$-\sum_i \partial_i[A_i(\boldsymbol{x})P_\lambda(\boldsymbol{x})] + \tfrac{1}{2}\sum_{i,j}\partial_i\partial_j[B_{ij}(\boldsymbol{x})P_\lambda(\boldsymbol{x})] = -\lambda P_\lambda(\boldsymbol{x}) \tag{5.3.167}$$

$$\sum A_i(\boldsymbol{x})\partial_i Q_{\lambda'}(\boldsymbol{x}) + \tfrac{1}{2}\sum B_{ij}(\boldsymbol{x})\partial_i\partial_j Q_{\lambda'}(\boldsymbol{x}) = -\lambda' Q_{\lambda'}(\boldsymbol{x})\,. \tag{5.3.168}$$

Whether $Q_{\lambda'}(\boldsymbol{x})$ and $P_\lambda(\boldsymbol{x})$ satisfy absorbing or reflecting boundary conditions, one can show, in a manner very similar to that used in Sect.5.2.4, that

$$-(\lambda - \lambda')\int_R d\boldsymbol{x}\, P_\lambda(\boldsymbol{x})Q_{\lambda'}(\boldsymbol{x}) = 0 \tag{5.3.169}$$

so that the $P_\lambda(\boldsymbol{x})$ and $Q_{\lambda'}(\boldsymbol{x})$ form a bi-orthogonal set. However, the functional relationship between the P_λ and the Q_λ only exists if detailed balance with all $\varepsilon_i = 1$ is valid. We assume

$$\int d\boldsymbol{x}\, P_\lambda(\boldsymbol{x})Q_\lambda(\boldsymbol{x}) = \delta_{\lambda\lambda'}\,, \tag{5.3.170}$$

if the spectrum of eigenvalues λ is discrete. [The Kronecker $\delta_{\lambda\lambda'}$ is to be replaced

by $\delta(\lambda - \lambda')$ when the spectrum is continuous, except where we have reflecting boundary conditions, and $\lambda = 0$ is stationary. The normalisation of $p_s(\boldsymbol{x})$ then gives

$$\int d\boldsymbol{x}\, P_0(\boldsymbol{x})Q_0(\boldsymbol{x}) = \int d\boldsymbol{x}\, P_0(\boldsymbol{x}) = 1 \tag{5.3.171}$$

so that there is also a discrete point with zero eigenvalue in the spectrum then.]

For, if detailed balance is valid, we have already noted in Sect.5.3.5 (iii) that $p(\varepsilon \boldsymbol{x}, -t)/p_s(\boldsymbol{x})$ is a solution of the backward Fokker-Planck equation so that, from the uniqueness of solutions, we can say

$$Q_\lambda(\boldsymbol{x}) = \eta_\lambda P_\lambda(\varepsilon \boldsymbol{x})/p_s(\boldsymbol{x}) . \tag{5.3.172}$$

Here, $\eta_\lambda = \pm 1$ but is otherwise undetermined. For, if $\lambda = \lambda'$, we can then write (5.3.170) in the form

$$\int d\boldsymbol{x}\, \frac{P_\lambda(\boldsymbol{x})\, P_\lambda(\varepsilon \boldsymbol{x})}{P_s(\boldsymbol{x})} = \eta_\lambda . \tag{5.3.173}$$

We cannot determine η_λ *a priori*, but by suitable normalisation it may be chosen ± 1. If, for example, all ε_i are -1 and $P_\lambda(\boldsymbol{x})$ is an odd function of x, it is clear that η_λ must be -1.

i) **Even Variables Only: Negativity of Eigenvalues.** If all the ε_i are equal to one, then we can write

$$Q_\lambda(\boldsymbol{x}) = P_\lambda(\boldsymbol{x})/p_s(\boldsymbol{x}) \tag{5.3.174}$$

and η_λ can always be set equal to one.

Hence, the expansions in eigenfunctions will be much the same as for the one-variable case.

The *completeness* of the eigenfunctions is a matter for proof in each individual case. If all the ε_i are equal to one, then we can show that the Fokker-Planck operator is self adjoint and negative semi-definite in a certain Hilbert space.

To be more precise, let us write the forward Fokker-Planck operator as

$$\mathscr{L}(\boldsymbol{x}) = -\sum_i \partial_i A_i(\boldsymbol{x}) + \tfrac{1}{2} \sum_{i,j} \partial_i \partial_j B_{ij}(\boldsymbol{x}) \tag{5.3.175}$$

and the backward Fokker-Planck operator as

$$\mathscr{L}^*(\boldsymbol{x}) = \sum_i A_i(\boldsymbol{x})\partial_i + \tfrac{1}{2} \sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i \partial_j . \tag{5.3.176}$$

Then the fact that, if all $\varepsilon_i = 1$, we can transform a solution of the forward FPE to a solution of the backward FPE by dividing by $p_s(\boldsymbol{x})$ arises from the fact that for any $f(\boldsymbol{x})$,

$$\mathscr{L}(\boldsymbol{x})[f(\boldsymbol{x})p_s(\boldsymbol{x})] = p_s(\boldsymbol{x})\mathscr{L}^*(\boldsymbol{x})[f(\boldsymbol{x})] . \tag{5.3.177}$$

Define a Hilbert space of functions $f(\boldsymbol{x})$, $g(\boldsymbol{x})$ ... by the scalar product

$$(f, g) = \int_R d\boldsymbol{x} \frac{f(\boldsymbol{x})g(\boldsymbol{x})}{p_s(\boldsymbol{x})}. \tag{5.3.178}$$

Then from (5.3.177),

$$(f, \mathscr{L}g) = \int_R d\boldsymbol{x} \frac{f(\boldsymbol{x})}{p_s(\boldsymbol{x})} \mathscr{L}(\boldsymbol{x}) \left[\frac{g(\boldsymbol{x})}{p_s(\boldsymbol{x})} p_s(\boldsymbol{x}) \right] \tag{5.3.179}$$

$$= \int_R dx\, f(\boldsymbol{x}) \mathscr{L}^*(x) \left[\frac{g(\boldsymbol{x})}{p_s(\boldsymbol{x})} \right] \tag{5.3.180}$$

and integrating by parts, discarding surface terms by use of either reflecting or absorbing boundary conditions

$$= \int_R d\boldsymbol{x} \frac{g(\boldsymbol{x})}{p_s(\boldsymbol{x})} \mathscr{L}(\boldsymbol{x})[f(\boldsymbol{x})]. \tag{5.3.181}$$

Thus, in this Hilbert space,

$$(f, \mathscr{L}g) = (g, \mathscr{L}f). \tag{5.3.182}$$

This condition is sufficient to ensure that the eigenvalues of $\mathscr{L}(\boldsymbol{x})$ are real. To prove negative semi-definiteness, notice that for even variables only, [see(5.3.75)]

$$A_i(\boldsymbol{x}) = \sum_j \partial_j [B_{ij}(\boldsymbol{x}) p_s(\boldsymbol{x})]/2p_s(\boldsymbol{x}) \tag{5.3.183}$$

so that for any $p(\boldsymbol{x})$,

$$\mathscr{L}(\boldsymbol{x})p(\boldsymbol{x}) = \sum_{i,j} \partial_i \left\{ - \frac{\partial_j [B_{ij}(\boldsymbol{x}) p_s(\boldsymbol{x})] p(\boldsymbol{x})}{2p_s(\boldsymbol{x})} + \tfrac{1}{2} \partial_j [B_{ij}(\boldsymbol{x}) p(\boldsymbol{x})] \right\}$$

$$= \frac{1}{2} \sum_{i,j} \partial_i \{ B_{ij}(\boldsymbol{x}) p_s(\boldsymbol{x}) \partial_j [p(\boldsymbol{x})/p_s(\boldsymbol{x})] \}. \tag{5.3.184}$$

Hence,

$$(p, \mathscr{L}p) = \frac{1}{2} \int d\boldsymbol{x}\, p(\boldsymbol{x})/p_s(\boldsymbol{x}) \sum_{i,j} \partial_i \{ B_{ij}(\boldsymbol{x}) p_s(\boldsymbol{x}) \partial_j [p(\boldsymbol{x})/p_s(\boldsymbol{x})] \}$$

and integrating by parts (discarding surface terms),

$$(p, \mathscr{L}p) = - \int d\boldsymbol{x}\, B_{ij}(\boldsymbol{x}) p_s(\boldsymbol{x}) \partial_i [p(\boldsymbol{x})/p_s(\boldsymbol{x})] \partial_j [p(\boldsymbol{x})/p_s(\boldsymbol{x})] \tag{5.3.185}$$

$$\leqslant 0$$

since $B_{ij}(\boldsymbol{x})$ is positive semi-definite.

Hence, we conclude for any eigenfunction $P_\lambda(\boldsymbol{x})$ that λ is real, and

$$\lambda \geqslant 0 \tag{5.3.186}$$

(remember that $-\lambda$ is the eigenvalue).

ii) **A Variational Principle.** Combined with this property is a variational principle. For, suppose we choose any real function $f(\boldsymbol{x})$ and expand it in eigenfunctions

$$f(\boldsymbol{x}) = \sum_{\lambda} \alpha_\lambda P_\lambda(\boldsymbol{x}) . \tag{5.3.187}$$

Further, let us fix $(f,f) = \int d\boldsymbol{x}\, f(\boldsymbol{x})^2/p_s(\boldsymbol{x}) = 1$.
Then,

$$-(f, \mathscr{L}f) = \sum_{\lambda} \lambda \alpha_\lambda^2$$

and (5.3.188)

$$(f,f) = \sum_{\lambda} \alpha_\lambda^2 .$$

Clearly $-(f, \mathscr{L}f)$ has its minimum of zero only if only the term $\lambda = 0$ exists, i.e.,

$$\begin{aligned} &\alpha_0 \neq 0 \\ &\alpha_\lambda = 0 \qquad \text{for} \qquad \lambda \neq 0 . \end{aligned} \tag{5.3.189}$$

Now choose $f(\boldsymbol{x})$ orthogonal to $P_0(x)$ so $\alpha_0 = 0$, and we see that the minimum of $(f, \mathscr{L}f)$ occurs when

$$\alpha_{\lambda_1} = 1$$

and (5.3.190)

$$\alpha_\lambda = 0 \qquad \text{for all other } \lambda,$$

where λ_1 is the next smallest eigenvalue. This means that $P_{\lambda_1}(\boldsymbol{x})$ is obtained by minimising $-(p, \mathscr{L}p)$ [which can be put in the form of (5.3.185)], subject to the condition that

$$\begin{aligned} &(p,p) = 1 \\ &(p, P_0) = 0 . \end{aligned} \tag{5.3.191}$$

This method can yield a numerical way of estimating eigenfunctions in terms of trial solutions and is useful particularly in bistability situations. It is, however, limited to situations in which detailed balance is valid.

iii) **Conditional Probability.** Assuming completeness, we find that the conditional probability can be written

$$p(\boldsymbol{x}, t \,|\, \boldsymbol{x}_0, 0) = \sum_{\lambda} P_\lambda(\boldsymbol{x}) Q_\lambda(\boldsymbol{x}_0) \mathrm{e}^{-\lambda t} . \tag{5.3.192}$$

iv) **Autocorrelation Matrix.** If detailed balance is valid, for the stationary autocorrelation matrix [using (5.3.172)] we have

$$G(t) = \langle \boldsymbol{x}(t)\boldsymbol{x}^{\mathrm{T}}(0)\rangle = \sum \eta_\lambda e^{-\lambda t}\left[\int_R d\boldsymbol{x}\, \boldsymbol{x}P_\lambda(\boldsymbol{x})\right]\left[\int_R d\boldsymbol{x}\, \boldsymbol{x}P_\lambda(\boldsymbol{x})\right]^{\mathrm{T}}\boldsymbol{\varepsilon} \tag{5.3.193}$$

which explicitly satisfies the condition

$$G(t) = \boldsymbol{\varepsilon}G^{\mathrm{T}}(t)\boldsymbol{\varepsilon} \tag{5.3.194}$$

derived in Sect. 5.3.4b.

v) **Spectrum Matrix.** The spectrum matrix is

$$\begin{aligned} S(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega t}G(t)dt \\ &= \frac{1}{2\pi}\left\{\int_0^{\infty} e^{-i\omega t}G(t)\,dt + \int_0^{\infty} e^{i\omega t}\,G^{\mathrm{T}}(t)dt\right\}. \end{aligned} \tag{5.3.195}$$

If we define, for convenience,

$$\boldsymbol{U}_\lambda = \int_R d\boldsymbol{x}\, \boldsymbol{x}P_\lambda(\boldsymbol{x}), \tag{5.3.196}$$

then

$$S(\omega) = \frac{1}{2\pi}\sum_\lambda\left[\frac{\eta_\lambda}{\lambda + i\omega}\boldsymbol{U}_\lambda\boldsymbol{U}_\lambda^{\mathrm{T}}\boldsymbol{\varepsilon} + \frac{\eta_\lambda}{\lambda - i\omega}\boldsymbol{\varepsilon}\boldsymbol{U}_\lambda\boldsymbol{U}_\lambda^{\mathrm{T}}\right]. \tag{5.3.197}$$

If any of the λ are complex, from the reality of $\mathscr{L}(\boldsymbol{x})$ we find that the eigenfunction belonging to λ^* is $[P_\lambda(\boldsymbol{x})]^*$, $\eta_\lambda = \eta_\lambda^*$ and $[\boldsymbol{U}_\lambda]^* = \boldsymbol{U}_\lambda *$. The spectrum is then obtained by adding the complex conjugate of those terms involving complex eigenvalues to (5.3.197).

In the case where $\boldsymbol{\varepsilon} = 1$ and hence $\eta_\lambda = 1$, the spectrum matrix has the simpler form

$$S(\omega) = \frac{1}{\pi}\sum_\lambda \frac{\lambda \boldsymbol{U}_\lambda\boldsymbol{U}_\lambda^{\mathrm{T}}}{\lambda^2 + \omega^2} \tag{5.3.198}$$

which is explicitly a positive definite matrix. The spectrum of a single variable q made up as a linear combination of the $\boldsymbol{x}$ by a formula such as

$$q = \boldsymbol{m}\cdot\boldsymbol{x} \tag{5.3.199}$$

is given by

$$S_q(\omega) = \boldsymbol{m}^{\mathrm{T}}S(\omega)\boldsymbol{m} = \frac{1}{\pi}\sum_\lambda \frac{\lambda(\boldsymbol{m}\cdot\boldsymbol{U}_\lambda)^2}{\lambda^2 + \omega^2} \tag{5.3.200}$$

and is a strictly decreasing function of ω.

In the case where $\boldsymbol{\varepsilon} \neq 1$, the positivity of the spectrum is no longer obvious, though general considerations such as the result (3.7.19) show that it must be.

An important difference arises because the λ now need not all be real, and this means that denominators of the form

$$1/[\lambda_R^2 + (\omega - \lambda_I)^2] \tag{5.3.201}$$

can occur, giving rise to peaks in the spectrum away from $\omega = 0$.

5.4 First Exit Time from a Region (Homogeneous Processes)

We wish to treat here the multidimensional analogue of the first passage time problem in one dimension, treated in Sect.5.2.7. As in that section, we will restrict ourselves to homogeneous processes.

The analogous problem here is to compute the earliest time at which a particle, initially inside a region R with boundary S, leaves that region.

As in the one-dimensional case, we consider the problem of solving the backward Fokker-Planck equation with an absorbing boundary condition on S, namely,

$$p(\boldsymbol{x}', t|\boldsymbol{x}, 0) = 0 \qquad (\boldsymbol{x} \in S)\,. \tag{5.4.1}$$

The probability that the particle, initially at $\boldsymbol{x}$, is somewhere within R after a time t is

$$G(\boldsymbol{x}, t) = \int_R d\boldsymbol{x}' p(\boldsymbol{x}', t|\boldsymbol{x}, 0) \tag{5.4.2}$$

and if T is the time at which the particle leaves R, then

$$\text{Prob}(T \geqslant t) = G(\boldsymbol{x}, t)\,. \tag{5.4.3}$$

Since the process is homogeneous, we find that $G(\boldsymbol{x}, t)$ obeys the backward Fokker-Planck equation

$$\partial_t G(\boldsymbol{x}, t) = \sum_i A_i(\boldsymbol{x})\partial_i G(\boldsymbol{x}, t) + 2\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\partial_j G(\boldsymbol{x}, t)\,. \tag{5.4.4}$$

The initial conditions on (5.4.4) will arise from:

i) $$p(\boldsymbol{x}', 0|\boldsymbol{x}, 0) = \delta(\boldsymbol{x} - \boldsymbol{x}') \tag{5.4.5}$$

so that

$$\begin{aligned} G(\boldsymbol{x}, 0) &= 1 \qquad \boldsymbol{x} \in R \\ &= 0 \qquad \text{elsewhere;} \end{aligned} \tag{5.4.6}$$

ii) the boundary condition (5.4.1) requires

$$G(\boldsymbol{x}, t) = 0 \qquad \boldsymbol{x} \in S\,. \tag{5.4.7}$$

As in Sect. 5.2.7, we find that these imply that the *mean exit time* from R starting at $\boldsymbol{x}$, for which we shall use the symbol $T(\boldsymbol{x})$, satisfies

$$\sum_i A_i(\boldsymbol{x})\partial_i T(\boldsymbol{x}) + 2\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\partial_j T(\boldsymbol{x}) = -1 \tag{5.4.8}$$

with the boundary condition

$$T(\boldsymbol{x}) = 0 \qquad \boldsymbol{x} \in S \tag{5.4.9}$$

and the nth moments

$$T_n(\boldsymbol{x}) = \langle T^n \rangle = \int_0^\infty t^{n-1} G(\boldsymbol{x}, t)dt \tag{5.4.10}$$

satisfy

$$-n\, T_{n-1}(\boldsymbol{x}) = \sum_i A_i(\boldsymbol{x})\partial_i T_n(\boldsymbol{x}) + \sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\partial_j T_n(\boldsymbol{x}) \tag{5.4.11}$$

with the boundary conditions

$$T_n(\boldsymbol{x}) = 0 \qquad \boldsymbol{x} \in S\,. \tag{5.4.12}$$

Inclusion of Reflecting Regions. It is possible to consider that S, the boundary of R, is divided into two regions S_r and S_a such that the particle is reflected when it meets S_r and is absorbed when it meets S_a. The boundary conditions on $G(\boldsymbol{x}, t)$ are then, from those derived in Sect. 5.2.4,

$$\sum_{i,j} n_i B_{ij}(\boldsymbol{x})\partial_j G(\boldsymbol{x}, t) = 0 \qquad (\boldsymbol{x} \in S_r) \tag{5.4.13}$$

$$G(\boldsymbol{x}, t) = 0 \qquad (\boldsymbol{x} \in S_a) \tag{5.4.14}$$

and hence,

$$\sum_{i,j} n_i B_{ij}(\boldsymbol{x})\partial_j T_n(\boldsymbol{x}) = 0 \qquad (\boldsymbol{x} \in S_r) \tag{5.4.15}$$

$$T_n(\boldsymbol{x}) = 0 \qquad (\boldsymbol{x} \in S_a)\,. \tag{5.4.16}$$

5.4.1 Solutions of Mean Exit Time Problems

The basic partial differential equation for the mean first passage time is only simple to solve in one dimension or in situations where there is a particular symmetry available. Asymptotic approximations can provide very powerful results, but these will be dealt with in Chap.9. We will illustrate some methods here with some examples.

a) Ornstein-Uhlenbeck Process in Two Dimensions (Rotationally Symmetric)
we suppose that a particle moves according to

$$\begin{aligned} dx &= -kx\,dt + \sqrt{D}\,dW_1(t) \\ dy &= -ky\,dt + \sqrt{D}\,dW_2(t) \end{aligned} \tag{5.4.17}$$

and want to know the mean exit time from the region

$$x^2 + y^2 \leqslant a^2 \tag{5.4.18}$$

given the initial position is at x_0, y_0.

This is easily reduced to the one-dimensional problem for the variable

$$r = \sqrt{x^2 + y^2} \tag{5.4.19}$$

by changing variables as in Sect.4.4.5, namely,

$$dr = (-kr + \tfrac{1}{2}\,D/r)dt + \sqrt{D}\,dW(t) \tag{5.4.20}$$

and we want to know the mean exit time from the region $(0, a)$. This can be solved by using (5.2.165) with the replacements

$$\left.\begin{aligned} U(x) &\to \tfrac{1}{2}kr^2 - \tfrac{1}{2}D\log r \\ D &\to \tfrac{1}{2}D \\ x_0 &\to a \\ a &\to \sqrt{x_0^2 + y_0^2} \equiv r_0 \\ -\infty &\to 0\,. \end{aligned}\right\} \tag{5.4.21}$$

Thus, $T(r_0 \to a)$

$$= \frac{1}{D}\int_{r_0}^{a} y^{-1}\exp[ky^2/D]dy \int_0^y z\exp(-kz^2/D)dz\,. \tag{5.4.22}$$

The problem is thus essentially one dimensional. This does not often happen.

b) Application of Eigenfunctions
Suppose we use the eigenfunctions $Q_\lambda(\boldsymbol{x})$ and $P_\lambda(\boldsymbol{x})$ to expand the mean first passage time as

$$T(\boldsymbol{x}) = \sum t_\lambda Q_\lambda(\boldsymbol{x}). \tag{5.4.23}$$

We suppose that the $P_\lambda(\boldsymbol{x})$ and $Q_\lambda(\boldsymbol{x})$ satisfy the boundary conditions required for the particular exit problem being studied, so that $T(\boldsymbol{x})$ as written in (5.4.23) satisfies the appropriate boundary conditions.

We then expand

$$-1 = \sum_\lambda I_\lambda Q_\lambda(\boldsymbol{x})\,, \tag{5.4.24}$$

where

$$I_\lambda = -\int_R dx\, P_\lambda(x)\,. \tag{5.4.25}$$

Inserting (5.4.23) into (5.4.8) yields

$$\lambda t_\lambda = -I_\lambda \tag{5.4.26}$$

so that

$$T(x) = \sum_\lambda \frac{1}{\lambda} Q_\lambda(x) \int_R dx'\, P_\lambda(x')\,. \tag{5.4.27}$$

The success of the method depends on the knowledge of the eigenvalues satisfying the correct boundary conditions on S and normalised on R.

c) Asymptotic Result

If the first eigenvalue λ_1 is very much less than all other eigenvalues, the series may be approximated by its first term. This will mean that the eigenfunction Q_1 will be very close to a solution of

$$\sum_i A_i(x)\partial_i f(x) + \frac{1}{2}\sum_{i,j} B_{ij}(x)\partial_i\partial_j f(x) = 0 \tag{5.4.28}$$

since λ_1 is very small. Hence,

$$Q_1(x) \sim K \tag{5.4.29}$$

where K is a constant. Taking account of the bi-orthonormality of P_λ and Q_λ, we see

$$1 = \int dx\, P_1(x)Q_1(x) \sim K \int dx\, P_1(x) \tag{5.4.30}$$

so that

$$T(x) \sim 1/\lambda_1\,. \tag{5.4.31}$$

The reasoning given here is rather crude. It can be refined by the asymptotic methods of Chap.9.

d) Application of the Eigenfunction Method

Two-dimensional Brownian motion: the particle moves in the $x\,y$ plane within a square whose corners are (0, 0), (0, 1), (1, 0), (1, 1). The sides of this square are absorbing barriers. $T(x, y)$ obeys

$$\frac{D}{2}\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) = -1\,. \tag{5.4.32}$$

The eigenfunctions satisfying the boundary condition $T = 0$ on the edges of the square are

$$P_{n,m}(x, y) = \sin(n\pi x)\sin(m\pi x) \tag{5.4.33}$$

with

$$Q_{n,m}(x, y) = 4 \sin(n\pi x) \sin(m\pi x) \tag{5.4.34}$$

and n, m positive and integral. The eigenvalues are

$$\lambda_{n,m} = \frac{\pi^2 D}{2}(n^2 + m^2) . \tag{5 4.35}$$

The coefficient

$$\begin{aligned} \int_R dx\, dy\, P_{n,m}(x, y) &= 0 \qquad \text{(either } n \text{ or } m \text{ even)} \\ &= \frac{4}{mn\pi^2} \quad (m \text{ and } n \text{ both odd}) . \end{aligned} \tag{5.4.36}$$

Hence,

$$T(x, y) = \frac{1}{D} \sum_{\substack{n,m \\ \text{odd}}} \frac{32}{\pi^2\, nm(m^2 + n^2)} \sin(n\pi x) \sin(m\pi y) . \tag{5.4.37}$$

5.4.2 Distribution of Exit Points

This problem is the multidimensional analogue of that treated in Sect.5.2.8. Namely, what is the probability of exiting through an element $d\boldsymbol{S}(\boldsymbol{a})$ at $\boldsymbol{a}$ of the boundary S of the region R. We assume absorption on all S.

The probability that the particle exits through $d\boldsymbol{S}(\boldsymbol{a})$ after time t is

$$\boldsymbol{g}(\boldsymbol{a}, \boldsymbol{x}, t)|d\boldsymbol{S}(\boldsymbol{a})| = -\int_t^\infty dt' \boldsymbol{J}(\boldsymbol{a}, t' | \boldsymbol{x}, 0) \cdot d\boldsymbol{S}(a) . \tag{5.4.38}$$

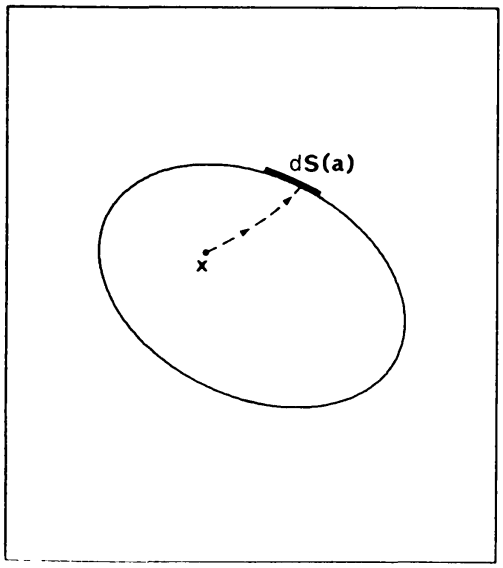


Fig. 5.5. Region and surface considered in Sect. 5.4.2

Similar reasoning to that in Sect.5.2.8 shows that $g(\boldsymbol{a}, \boldsymbol{x}, t)$ obeys the backward Fokker-Planck equation

$$\sum_i A_i(\boldsymbol{x})\partial_i g(\boldsymbol{a}, \boldsymbol{x}, t) + \tfrac{1}{2}\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\partial_j g(\boldsymbol{a}, \boldsymbol{x}, t) = \partial_t g(\boldsymbol{a}, \boldsymbol{x}, t)\,. \tag{5.4.39}$$

The boundary conditions follow by definition. Initially we have

$$g(\boldsymbol{a}, \boldsymbol{x}, 0) = 0 \qquad \text{for} \quad \boldsymbol{x} \neq \boldsymbol{a}, \boldsymbol{x} \in R \tag{5.4.40}$$

and at all times

$$g(\boldsymbol{a}, \boldsymbol{x}, t) = 0 \qquad \text{for} \quad \boldsymbol{x} \neq \boldsymbol{a}, \boldsymbol{x} \in S\,. \tag{5.4.41}$$

If $\boldsymbol{x} = \boldsymbol{a}$, then exit through $d\boldsymbol{S}(\boldsymbol{a})$ is certain, hence,

$$g(\boldsymbol{a}, \boldsymbol{a}, t)\, d\boldsymbol{S}(\boldsymbol{a}) = 1 \qquad \text{for all } t \tag{5.4.42}$$

or effectively

$$g(\boldsymbol{a}, \boldsymbol{x}, t) = \delta_s(\boldsymbol{a} - \boldsymbol{x}) \qquad \boldsymbol{x} \in \boldsymbol{S}, \quad \text{for all } t, \tag{5.4.43}$$

where $\delta_s(\boldsymbol{a} - \boldsymbol{x})$ is an appropriate surface delta function such that

$$\int |d\boldsymbol{S}(\boldsymbol{a})|\, \delta_s(\boldsymbol{a} - \boldsymbol{x}) = 1\,. \tag{5.4.44}$$

The probability of ultimate exit through $d\boldsymbol{S}(\boldsymbol{a})$ is

$$\pi(\boldsymbol{a}, \boldsymbol{x})\,|d\boldsymbol{S}(\boldsymbol{a})| = g(\boldsymbol{a}, \boldsymbol{x}, 0)\,|\,dS(\boldsymbol{a})\,|\,. \tag{5.4.45}$$

The mean exit time given that exit occurred at $\boldsymbol{a}$ is

$$T(\boldsymbol{a}, \boldsymbol{x}) = \int_0^\infty dt\; g(\boldsymbol{a}, \boldsymbol{x}, t)/\pi(\boldsymbol{a}, \boldsymbol{x}) \tag{5.4.46}$$

and in the same way as in Sect.5.2.8, we show that this satisfies

$$\boxed{\sum_i A_i(\boldsymbol{x})\partial_i[\pi(\boldsymbol{a}, \boldsymbol{x})T(\boldsymbol{a}, \boldsymbol{x})] + \tfrac{1}{2}\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\partial_j[\pi(\boldsymbol{a}, \boldsymbol{x})T(\boldsymbol{a}, \boldsymbol{x})] = -\pi(\boldsymbol{a}, \boldsymbol{x})} \tag{5.4.47}$$

and the boundary conditions is

$$\pi(\boldsymbol{a}, \boldsymbol{x})T(\boldsymbol{a}, \boldsymbol{x}) = 0, \qquad \boldsymbol{x} \in S\,. \tag{5.4.48}$$

Further, by letting $t \to \infty$ in the corresponding Fokker-Planck equation for $g(\boldsymbol{a}, \boldsymbol{x}, t)$, we obtain the equation for $\pi(\boldsymbol{a}, \boldsymbol{x})$:

$$\boxed{\sum A_i(\boldsymbol{x})\partial_i[\pi(\boldsymbol{a}, \boldsymbol{x})] + \tfrac{1}{2}\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\partial_j[\pi(\boldsymbol{a}, \boldsymbol{x})] = 0\,.} \tag{5.4.49}$$

The boundary condition for (5.4.49) is

$$\pi(\boldsymbol{a}, \boldsymbol{x}) = 0 \qquad \boldsymbol{a} \neq \boldsymbol{x}, \qquad \boldsymbol{x} \in S \tag{5.4.50}$$

and

$$\int |dS(\boldsymbol{a})|\,\pi(\boldsymbol{a}, \boldsymbol{x}) = 1\,. \tag{5.4.51}$$

Thus, we can summarise this as

$$\pi(\boldsymbol{a}, \boldsymbol{x}) = \delta_s(\boldsymbol{a} - \boldsymbol{x}) \qquad \boldsymbol{x} \in S\,, \tag{5.4.52}$$

where $\delta_s(\boldsymbol{a} - \boldsymbol{x})$ is the surface delta function for the boundary $\boldsymbol{S}$.

6. Approximation Methods for Diffusion Processes

The methods described in the previous two chapters have concentrated on exact results, and of course the results available are limited in their usefulness. Approximation methods are the essence of most applications, where some way of reducing a problem to an exactly soluble one is always sought. It could even be said that most work on applications is concerned with the development of various approximations.

There are two major approximation methods of great significance. The first is the small noise expansion theory which gives solutions linearised about a deterministic equation. Since noise is often small, this is a method of wide practical application, the equations are reduced into a sequence of time-dependent Ornstein-Uhlenbeck processes. Mostly the first order is used.

Another large class of methods is given by adiabatic elimination, in which different time scales are identified and fast variables are eliminated completely. This forms the basis of the second half of the chapter.

6.1 Small Noise Perturbation Theories

In many physical and chemical problems, the stochastic element in a dynamical system arises from thermal fluctuations, which are always very small. Unless one measures very carefully, it is difficult to detect the existence of fluctuations. In such a case, the time development of the system will be almost deterministic and the fluctuations will be a small perturbation.

With this in mind, we consider a simple linear example which is exactly soluble: a one-variable Ornstein-Uhlenbeck process described by the stochastic differential equation:

$$dx = -kx\,dt + \varepsilon\,dW(t) \tag{6.1.1}$$

for which the Fokker-Planck equation is

$$\partial_t p = \partial_x(kx\,p) + \tfrac{1}{2}\varepsilon^2\partial_x^2 p. \tag{6.1.2}$$

The solutions of these have been previously investigated in Sects. 3.8.4, 4.4.4. Here ε is a small parameter which is zero in the deterministic limit. However, the limit $\varepsilon \to 0$ is essentially different in the two cases.

In the stochastic differential equation (SDE) (6.1.1), as $\varepsilon \to 0$, the differential equation becomes nonstochastic but remains of first order in t, and the limit $\varepsilon \to 0$ is therefore not singular. In contrast, in the Fokker-Planck equation (FPE) (6.1.2),

the limit $\varepsilon \to 0$ reduces a second-order differential equation to one of first order. This limit is singular and any perturbation theory is a singular perturbation theory.

The solution to (6.1.1) is known exactly—it is (4.4.26)

$$x_\varepsilon(t) = c\,e^{-kt} + \varepsilon \int_0^t e^{-k(t-t')}\,dW(t') \tag{6.1.3}$$

which can be written

$$x_\varepsilon(t) = x_0(t) + \varepsilon x_1(t) \tag{6.1.4}$$

and this is generic; that is, we can normally solve in a power series in the small parameter ε. Furthermore, the zero-order term $x_0(t)$ is the solution of the equation obtained by setting $\varepsilon = 0$, i.e., of

$$dx = -kx\,dt\,. \tag{6.1.5}$$

The situation is by no means so simple for the Fokker-Planck Equation (6.1.2). Assuming the initial condition c is a nonstochastic variable, the exact solution is the Gaussian with mean and variance given by

$$\langle x(t)\rangle = \alpha(t) \equiv c\,e^{-kt} \tag{6.1.6}$$

$$\operatorname{var}\{x(t)\} = \varepsilon^2\beta(t) \equiv \varepsilon^2(1 - e^{-2kt})/2k \tag{6.1.7}$$

so that

$$p_\varepsilon(x, t\,|\,c, 0) = \frac{1}{\varepsilon}\frac{1}{\sqrt{2\pi\beta(t)}}\exp\left\{-\frac{1}{\varepsilon^2}\frac{[x - \alpha(t)]^2}{2\beta(t)}\right\}. \tag{6.1.8}$$

The solution for the conditional probability has the limiting form as $\varepsilon \to 0$ of

$$p_\varepsilon(x, t\,|\,c, 0) \to \delta[x - \alpha(t)] \tag{6.1.9}$$

which corresponds exactly to the first-order solution of the SDE, which is a deterministic trajectory along the path $x(t) = c\exp(-kt)$. However, p_ε cannot be expanded as a simple power series in ε. To carry out a power series expansion, one must define *a scaled variable* at each time;

$$y = [x - \alpha(t)]/\varepsilon \tag{6.1.10}$$

so that a probability density for y is

$$\bar{p}_\varepsilon(y, t\,|\,0, 0) = p_\varepsilon(x, t\,|\,c, 0)\frac{dx}{dy} \tag{6.1.11}$$

$$= \frac{1}{\sqrt{2\pi\beta(t)}}\exp\left[-\frac{y^2}{2\beta(t)}\right]. \tag{6.1.12}$$

The probability density for the scaled variable y has no singularity; indeed, we have no ε dependence.

The transformation (6.1.10) can be rewritten as

$$x = \alpha(t) + \varepsilon y \tag{6.1.13}$$

and is to be interpreted as follows. From (6.1.12) we see that the distribtuion of y is Gaussian with mean zero, variance $\beta(t)$. Equation (6.1.13) says that the deviation of x from $\alpha(t)$, the deterministic path, is of order ε as $\varepsilon \to 0$, and the coefficient is the Gaussian random variable y. This is essentially the same conclusion as was reached by the SDE method.

The general form of these results is the following. We have a system described by the SDE

$$dx = a(x)dt + \varepsilon b(x)\, dW(t). \tag{6.1.14}$$

Then we can write the solution as

$$x_\varepsilon(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \ldots \tag{6.1.15}$$

and solve successively for the $x_n(t)$. In particular, $x_0(t)$ is the solution of the deterministic equation

$$dx = a(x)dt\,. \tag{6.1.16}$$

Alternatively, we consider the Fokker-Planck equation

$$\partial_t p = -\partial_x[a(x)p] + \tfrac{1}{2}\varepsilon^2\partial_x^2[b(x)^2 p]\,. \tag{6.1.17}$$

Then by changing the variable to the *scaled variable* and thus writing

$$y = [x - x_0(t)]/\varepsilon \tag{6.1.18}$$

$$\bar{p}_\varepsilon(y, t) = \varepsilon p(x, t|c, 0)\,, \tag{6.1.19}$$

we can write the perturbation expansion

$$\bar{p}_\varepsilon(y, t) = \bar{p}_0(y, t) + \varepsilon\bar{p}_1(y, t) + \varepsilon^2\bar{p}_2(y, t) + \ldots\,. \tag{6.1.20}$$

Here we will find that $\bar{p}_0(y, t)$ is indeed a genuine probability density, i.e., is positive and normalised, while the higher-order terms are negative in some regions. Thus, it can be said that the Fokker-Planck perturbation theory is not probabilistic. In contrast, the SDE theory expands in a series of random variables $x_n(t)$, each of which has its own probability distribution. At every stage, the system is probabilistic.

And finally, the most noticeable difference. The first term in the SDE perturbation theory is $x_0(t)$ which is the solution of the SDE obtained by setting $\varepsilon = 0$ in (6.1.1). In contrast, the first term $\bar{p}_0(y, t)$ in (6.1.20) is *not* the solution of the equa-

tion obtained by setting $\varepsilon = 0$ in (6.1.2). In general, it is a limiting form of a FPE for $\bar{p}_\varepsilon(y, t)$ obtained by setting $\varepsilon = 0$ in the FPE for the scaled variable, y.

6.2 Small Noise Expansions for Stochastic Differential Equations

We consider a stochastic differential equation of the form

$$dx = a(x)dt + \varepsilon b(x)dW(t) \tag{6.2.1}$$

in which ε is a small parameter. At this stage we exclude time dependence from $a(x)$ and $b(x)$, which is only necessary in order to simplify the algebra. The results and methods are exactly the same. We then assume that the solution $x(t)$ of (6.2.1) can be written

$$x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \dots . \tag{6.2.2}$$

We also assume that we can write

$$a(x) = a(x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \dots) \tag{6.2.3}$$

$$= a_0(x_0) + \varepsilon a_1(x_0, x_1) + \varepsilon^2 a_2(x_0, x_1, x_2) + \dots . \tag{6.2.4}$$

The particular functional dependence in (6.2.4) is important and is easy to demonstrate, for

$$a(x) = a(x_0 + \sum_{m=1}^{\infty} \varepsilon^m x_m)$$

$$= \sum_{p=0}^{\infty} \frac{1}{p!} \frac{d^p a(x_0)}{dx_0^p} (\sum_{m=1}^{\infty} \varepsilon^m x_m)^p . \tag{6.2.5}$$

Formally resumming is not easy, but it is straightforward to compute the first few powers of ε and to obtain

$$\begin{aligned} a_0(x_0) &= a(x_0) \\ a_1(x_0, x_1) &= x_1 \frac{da(x_0)}{dx_0} \\ a_2(x_0, x_1, x_2) &= x_2 \frac{da(x_0)}{dx_0} + \frac{1}{2} x_1^2 \frac{d^2 a(x_0)}{dx_0^2} \\ a_3(x_0, x_1, x_2, x_3) &= x_3 \frac{da(x_0)}{dx_0} + x_1 x_2 \frac{d^2 a(x_0)}{dx_0^2} + \frac{1}{6} x_1^3 \frac{d^3 a(x_0)}{dx_0^3} . \end{aligned} \tag{6.2.6}$$

Although it is not easy to write explicitly the full set of terms in general, it is easy to see that we can always write for $n \geqslant 1$,

$$a_n(x_0, x_1, \dots x_n) = x_n \frac{da(x_0)}{dx_0} + A_n(x_0, x_1, \dots x_{n-1}) , \tag{6.2.7}$$

where it is seen that A_n is independent of x_n. In fact, it is easy to see directly from (6.2.5) that the coefficient of ε^n can only involve x_n if this is contributed from the term $p = 1$, and the only other possibility for ε^n to arise is from terms with $m < n$, which do not involve x_n.

The form (6.2.7) for a_n is very important. It is clear that if we also require

$$b(x) = b_0(x_0) + \varepsilon\, b_1(x_0, x_1) + \varepsilon^2 b_2(x_0, x_1, x_2) + \dots , \tag{6.2.8}$$

then all the same conclusions apply. However, they are not so important in the perturbation expansion.

We now substitute the expansions (6.2.2, 7, 8) in the stochastic differential equation (6.2.1) and equate coefficients of like powers of ε. We then obtain an infinite set of SDE's. In these we use the notation

$$k(x_0) = -\frac{da(x_0)}{dx_0} \tag{6.2.9}$$

which simplifies the notation.
We obtain

$$dx_0 = a(x_0)dt \tag{6.2.10a}$$

$$dx_n = [-k(x_0)x_n + A_n(x_0, \dots x_{n-1})]dt + b_{n-1}(x_0, \dots x_{n-1})dW(t)\,. \tag{6.2.10b}$$

These equations can now be solved sequentially. Equation (6.2.10a) is a (possibly non-linear) ordinary differential equation whose solution is assumed to be known, subject to an initial condition. It is of course possible to assume independent nontrivial initial conditions for all the x_n, but this is unnessarily complicated. It is simplest to write (setting $t = 0$ as the initial time)

$$\begin{aligned} x_0(0) &= x(0) \\ x_n(0) &= 0 \qquad n \geqslant 1\,. \end{aligned} \tag{6.2.11}$$

Assuming that the solution of (6.2.9) is given by

$$x_0(t) = \alpha(t)\,, \tag{6.2.12}$$

the equation (6.2.10a) for $n = 1$ can be written as

$$\boxed{dx_1 = -k[\alpha(t)]x_1 dt + b[\alpha(t)]dW(t)\,,} \tag{6.2.13}$$

where we have noted from (6.2.5) that A_0 vanishes and $b_0 = b$.

This equation, the first of a perturbation theory, is a time-dependent Ornstein-Uhlenbeck process whose solution can be obtained straightforwardly by the methods of Sect.4.4.9, reduced to one dimension. The solution is obtained simply by multiplying by the integrating factor

$$\exp\{\int_0^t dt' k[\alpha(t')]\}$$

and is

$$x_1(t) = \int_0^t b[\alpha(t')] \exp\{-\int_{t'}^t k[\alpha(s)ds\} dW(t') , \tag{6.2.14}$$

where the initial condition $x_1(0) = 0$ has been included.

For many purposes this form is quite adequate and amounts to a linearisation of the original equation about the deterministic solution. Higher-order terms are more complex because of the more complicated form of (6.2.10b) but are, in essence, treated in exactly the same way. In order to solve the equation for $x_N(t)$, we assume we know all the $x_n(t)$ for $n < N$ so that A_n and b_{n-1} become known (stochastic) functions of t after substituting these solutions. Then (6.2.10b) becomes

$$dx_N = \{-k[\alpha(t)]x_N + A_N(t)\} dt + b_{N-1}(t)dW(t) \tag{6.2.15}$$

whose solution is obtained directly, or from Sect.4.4.9, as

$$x_N(t) = \int_0^t [A_N(t')dt' + b_{N-1}(t')dW(t')] \exp\{-\int_{t'}^t k[\alpha(s)]ds\} . \tag{6.2.16}$$

Formally, the procedure is now complete. The range of validity of the method and its practicability are yet unanswered. Like all perturbation theories, terms rapidly become unwieldy with increasing order.

6.2.1 Validity of the Expansion

The expansion will not normally be a convergent power series. For (6.2.14) shows that $x_1(t)$ is a Gaussian variable, being simply an Ito integral with nonstochastic coefficients, and hence $x_1(t)$ can with finite probability assume a value greater than any fixed value. Thus, only if all power series involved in the derivation of (6.2.5) are convergent for all arguments, no matter how large, can we expect the method to yield a convergent expansion.

We can, in fact, show that the expansion is *asymptotic* by using the results on dependence on a parameter given in Sect.4.3.7.

We define a remainder by

$$y_n(\varepsilon, t) = [x(t) - \sum_{r=0}^{n} \varepsilon^r x_r(t)]/\varepsilon^{n+1} , \tag{6.2.17}$$

where the $x_r(t)$ are solutions of the set of stochastic differential equations (6.2.10) with initial conditions (6.2.11).

We then derive an equation for $y_n(t)$. We can write

$$a[x(t)] = a[\sum_{r=0}^{n} \varepsilon^r x_r(t) + y_n(\varepsilon, t)\varepsilon^{n+1}] \tag{6.2.18}$$

and we define a function $\hat{a}_{n+1}[X_0, X_1, X_2 \ldots X_n, Y, \varepsilon]$ by

$$\hat{a}_{n+1}[X_0, X_1, \ldots X, Y, \varepsilon] = \varepsilon^{-n-1}\{a[\sum_{r=0}^{n} \varepsilon^r X_r + \varepsilon^{n+1} Y] - \sum_{r=0}^{n} \varepsilon^r a_r(X_0, X_1, \ldots X_r)\} . \tag{6.2.19}$$

We require that for all fixed $X_0, X_1, \ldots X_N, Y$,

$$\lim_{\varepsilon \to 0} \hat{a}_{n+1}[X_0, X_1, \ldots X_n, Y, \varepsilon] \tag{6.2.20}$$

exists. We similarly define $\hat{b}_n[X_0, X_1, \ldots X_n, Y, \varepsilon]$ and impose the same condition on it.

This condition is not probabilistic, but expresses required analytic properties of the functions $a(x)$ and $b(x)$; it requires, in fact, that the expansions (6.2.4, 8) be merely *asymptotic* expansions.

Now we can write the differential equation for $y_n(\varepsilon, t)$ as

$$dy_n = \hat{a}_{n+1}[x_0(t), x_1(t), \ldots x_n(t), y_n, \varepsilon]\, dt + \hat{b}_n[x_0(t), x_1(t), \ldots x_{n-1}(t), y_n, \varepsilon]\, dW(t) . \tag{6.2.21}$$

The coefficients of dt and $dW(t)$ are now *stochastic* functions because the $x_r(t)$ are stochastic. However, the requirement (6.2.20) is now an almost certain limit, and hence implies the existence of the stochastic limits

$$\text{st-lim}_{\varepsilon \to 0} \hat{a}_{n+1}[x_0(t), x_1(t), \ldots x_n(t), y_n, \varepsilon] \equiv \tilde{a}_{n+1}(t, y_n) \tag{6.2.22}$$

and

$$\text{st-lim}_{\varepsilon \to 0} \hat{b}_n [x_0(t), x_1(t), \ldots x_{n-1}(t), y_n, \varepsilon] \equiv \tilde{b}_n(t, y_n) \tag{6.2.23}$$

which is sufficient to satisfy the result of Sect.4.3.7 on the continuity of solutions of the SDE (6.2.21) with respect to the parameter ε, provided the appropriate Lipschitz conditions (ii) and (iii) of Sect.4.3.7 are satisfied. Thus, $y_n(0, t)$ exists as a solution of the SDE

$$dy_n(0, t) = \tilde{a}_{n+1}[t, y_n(0, t)] + \tilde{b}_n[t, y_n(0, t)]dW(t) \tag{6.2.24}$$

which, from the definition (6.2.17) shows that

$$x(t) - \sum_{r=0}^{n} \varepsilon^r x_r(t) \sim \varepsilon^{n+1} . \tag{6.2.25}$$

Hence, the expansion in power of ε is an *asymptotic expansion.*

6.2.2 Stationary Solutions (Homogeneous Processes)

A stationary solution is obtained by letting $t \to \infty$. If the process is, as written, homogeneous and ergodic, it does not matter what the initial condition is. In this

case, one chooses $x_0(0)$ so that $a[x_0(0)]$ vanishes and the solution to (6.2.10a) is

$$x_0(t) = x_0(0) \equiv \alpha \tag{6.2.26}$$

[where we write simply α instead of $\alpha(t)$].

Because of the initial condition, at $t = 0$ the solution (6.2.14) to the equation of order one is not a stationary process. One must either let $t \to \infty$ or set the initial condition not at $t = 0$, but at $t = -\infty$. Choosing the latter, we have

$$x_1^s(t) = \int_{-\infty}^{t} b(\alpha) \exp[-(t - t')k(\alpha)]dW(t') . \tag{6.2.27}$$

Similarly,

$$x_n^s(t) = \int_{-\infty}^{t} [A_n^s(t')dt' + b_{n-1}^s(t')dW(t')] \exp[-(t - t')k(\alpha)] , \tag{6.2.28}$$

where by A_n^s and b_{n-1}^s we mean the values of A_n and b_{n-1} obtained by inserting the stationary values of all arguments. From (6.2.28) it is clear that $x_n^s(t)$ is, by construction, stationary. Clearly the integrals in (6.2.27, 28) converge only if $k(\alpha) > 0$, which will mean that only a *stable* stationary solution of the deterministic process generates a stationary solution by this method. This is rather obvious—the addition of fluctuations to an unstable state derives the system away from that state.

6.2.3 Mean, Variance, and Time Correlation Function

If the series expansion in ε is valid in some sense, it is useful to know the expansion for mean and variance. Clearly

$$\langle x(t) \rangle = \sum_{n=0}^{\infty} \varepsilon^n \langle x_n(t) \rangle \tag{6.2.29}$$

$$\text{var}\{x(t)\} = \sum_{s=0}^{\infty} \varepsilon^n \sum_{m=0}^{\infty} \langle x_m(t) x_{n-m}(t) \rangle - \langle x_m(t) \rangle \langle x_{n-m}(t) \rangle] . \tag{6.2.30}$$

Since, however, we assume a deterministic initial condition and $x_0(t)$ is hence deterministic, all terms involving $x_0(t)$ vanish. We can then work out that

$$\begin{aligned} \text{var}\{x(t)\} &= \varepsilon^2 \text{var}\{x_1(t)\} + 2\varepsilon^3 \langle x_1(t), x_2(t) \rangle \\ &\quad + \varepsilon^4[2\langle x_1(t), x_3(t) \rangle + \text{var}\{x_2(t)\}] + \ldots \end{aligned} \tag{6.2.31}$$

and similarly,

$$\begin{aligned} \langle x(t), x(s) \rangle &= \varepsilon^2 \langle x_1(t), x_1(s) \rangle + \varepsilon^3[\langle x_1(t), x_2(s) \rangle + \langle x_1(s), x_2(t) \rangle] \\ &\quad + \varepsilon^4[\langle x_1(t), x_3(s) \rangle + \langle x_1(s), x_3(t) \rangle + \langle x_2(t), x_2(s) \rangle] \\ &\quad + \ldots . \end{aligned} \tag{6.2.32}$$

6.2.4 Failure of Small Noise Perturbation Theories

a) Example: Cubic Process and Related Behaviour

Consider the stochastic differential equation

$$dx = -x^3 dt + \varepsilon\, dW(t)\,. \tag{6.2.33}$$

It is not difficult to see that the expansion conditions (6.2.20) are trivially satisfied for the coefficients of both dt and $dW(t)$, and in fact for any *finite* t, an asymptotic expansion with terms $x_N(t)$ given by (6.2.16) is valid. However, at $x = 0$, it is clear that

$$\frac{d}{dx}(x^3)\bigg|_{x=0} = k(0) = 0\,, \tag{6.2.34}$$

and because $x = 0$ is the stationary solution of the deterministic equation, the perturbation series for stationary solutions is not likely to converge since the exponential time factors are all constant. For example, the first-order term in the stationary expansion is, from (6.2.27),

$$x_1^s(t) = \int_{-\infty}^{t} dW(t') = W(t) - W(-\infty) \tag{6.2.35}$$

which is infinite with probability one (being a Gaussian variable with infinite variance).

The problem is rather obvious. Near $x = 0$, the motion described by (6.2.33) is simply not able to be approximated by an Ornstein-Uhlenbeck process. For example, the stationary probability distribution, which is the stationary solution of the FPE

$$\partial_t p = \partial_x(x^3 p) + \tfrac{1}{2}\varepsilon^2 \partial_x^2 p\,, \tag{6.2.36}$$

is given by

$$p_s(x) = \mathcal{N} \exp\left(-x^4/2\varepsilon^2\right) \tag{6.2.37}$$

and the moments are

$$\begin{aligned} \langle x^n \rangle &= (2\varepsilon^2)^{n/4}\, \Gamma\left(\frac{n+1}{4}\right)/\Gamma\left(\frac{1}{4}\right) \qquad (n \text{ even}) \\ &= 0 \qquad\qquad\qquad\qquad\qquad\quad (n \text{ odd}). \end{aligned} \tag{6.2.38}$$

The lowest-order term of the expansion of the variance is proportional to ε to the first power, not ε^2 as in (6.2.31).

In this case, we must simply regard the *cubic process* described by (6.2.33) as a fundamental process. If we introduce the new scaled variables through the definitions

$$x = \sqrt{\varepsilon}\, y$$
$$t = \tau/\varepsilon \tag{6.2.39}$$

and use

$$dW(\tau/\varepsilon) = dW(\tau)/\sqrt{\varepsilon}, \tag{6.2.40}$$

then the cubic process can be recuced to a parameterless form

$$dy = -y^3 d\tau + dW(\tau)\,. \tag{6.2.41}$$

Regarding the solution of (6.2.41) as a known quantity, we can write

$$x(t) = \sqrt{\varepsilon}\; y(\varepsilon t) \tag{6.2.42}$$

so that the limit $\varepsilon \to 0$ is approached like $\sqrt{\varepsilon}$, and also with a slower time scale. This kind of scaling result is the basis of many critical phenomena.

A successful perturbation theory in the case where $a(x)$ behaves like x^3 near $x = 0$ must involve firstly the change of variables (6.2.39) then a similar kind of perturbation theory to that already outlined—but in which the zero-order solution is the cubic process. Thus, let us assume that we can write

$$a(x) = -x^3 c(x)\,, \tag{6.2.43}$$

where $c(x)$ is a smooth function with $c(0) \neq 0$. Then, using the transformations (6.2.39,40), we can rewrite the SDE as

$$dy = -y^3 c(y\sqrt{\varepsilon})d\tau + b(y\sqrt{\varepsilon})dW(\tau)\,. \tag{6.2.44}$$

If we expand $y(t)$, $c(y\sqrt{\varepsilon})$, $b(y\sqrt{\varepsilon})$ as series in $\sqrt{\varepsilon}$, we obtain a perturbation theory. If we write

$$y(t) = \sum_{n=0}^{\infty} \varepsilon^{n/2} y_n(t)\,, \tag{6.2.45}$$

then we get for the first two terms

$$dy_0 = -y_0^3 c(0)d\tau + b(0)dW(\tau) \tag{6.2.46}$$

$$dy_1 = -y_1\left[3y_0^2 c(0) + y_0^3 \frac{dc}{dx}(0)\right]d\tau + \left[\frac{db}{dx}(0)y_0\right]dW(\tau)\,. \tag{6.2.47}$$

We see that the equation for y_1 is in fact that of a time-dependent Ornstein-Uhlenbeck process with stochastic coefficients. Thus, *in principle*, as long as the cubic process is known, the rest is easily computed. In practice, not a great deal is in fact known about the cubic process and this kind of perturbation is not very practical.

b) Order ν Processes (ν odd)

If instead we have the Stochastic differential equation

$$dx = -x^{\nu}dt + \varepsilon\, dW(t)\,, \tag{6.2.48}$$

we find a stable deterministic state only if ν is an odd integer. In this case we make the transformation to scaled variables

$$x = y\varepsilon^{2/(1+\nu)} \tag{6.2.49}$$

$$t = \tau\, \varepsilon^{2(1-\nu)/(1+\nu)} \tag{6.2.50}$$

and follow very similar procedures to those used for the cubic process.

c) Bistable Systems

Suppose we have a system described by

$$dx = (x - x^3)dt + \varepsilon\, dW(t) \tag{6.2.51}$$

for which there are three deterministic stationary states, at $x = 0, \pm 1$. The state at $x = 0$ is deterministically unstable, and we can see directly from the perturbation theory that no stationary process arises from it, since the exponentials in the perturbation series integrals (6.2.27, 28) have increasing arguments.

The solutions $x_0(t)$ of the deterministic differential equation

$$dx/dt = x - x^3 \tag{6.2.52}$$

divide into three classes depending on their behaviour as $t \to \infty$. Namely,

i) $x_0(t) < 0 \Rightarrow x_0(t) \to -1$

ii) $x_0(0) = 0 \Rightarrow x_0(t) = 0$ for all t

iii) $x_0(t) > 0 \Rightarrow x_0(t) \to 1$.

Thus, depending on the initial condition, we get two different asymptotic expansions, whose stationary limits represent the fluctuations about the two deterministic stationary states. There is no information in these solutions about the possible jump from the branch $x = 1$ to the branch $x = -1$, or conversely – at least not in any obvious form. In this sense the asymptotic expansion fails, since it does not give a picture of the overall behaviour of the stationary state. We will see in Chap. 9 that this results because an asymptotic expansion of behaviour characteristic of jumps from one branch to the other is typically of the order of magnitude of $\exp(-1/\varepsilon^2)$, which approaches zero faster than any power as $\varepsilon \to 0$, and thus is not represented in an expansion in powers of ε.

6.3 Small Noise Expansion of the Fokker-Planck Equation

As mentioned in Sect. 6.1, a small noise expansion of a Fokker-Planck equation is a singular expansion involving the introduction of scaled variables. Let us consider how this is done.

We consider the Fokker-Planck equation

$$\partial_t p = -\partial_x[A(x)p] + \tfrac{1}{2}\varepsilon^2\partial_x^2[B(x)p]\,. \tag{6.3.1}$$

We assume the solution of the deterministic equation to be $\alpha(t)$ so that

$$d_t\alpha(t) = A[\alpha(t)]\,. \tag{6.3.2}$$

and introduce new variables (y, s) by

$$y = [x - \alpha(t)]/\varepsilon \tag{6.3.3}$$

$$s = t \tag{6.3.4}$$

and $$\hat{p}(y, s) = \varepsilon\, p(x, t)\,. \tag{6.3.5}$$

We note that

$$\partial_t\hat{p}(y, s) = \frac{\partial\hat{p}}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial\hat{p}}{\partial s}\frac{\partial s}{\partial t} = -\frac{\dot{\alpha}(t)}{\varepsilon}\frac{\partial\hat{p}}{\partial y} + \frac{\partial\hat{p}}{\partial s} \tag{6.3.6}$$

$$\partial_x\hat{p}(y, s) = \frac{1}{\varepsilon}\frac{\partial\hat{p}}{\partial y}, \tag{6.3.7}$$

so that substituting into (6.3.1) we get, with the help of the equation of motion (6.3.2) for $\alpha(t)$

$$\frac{\partial\hat{p}}{\partial s} = -\frac{\partial}{\partial y}\left\{\frac{A[\alpha(s) + \varepsilon y] - A[\alpha(s)]}{\varepsilon}\hat{p}\right\} + \frac{1}{2}\frac{\partial^2}{\partial y^2}\{B[\alpha(s) + \varepsilon y]\hat{p}\}\,. \tag{6.3.8}$$

We are now in a position to make an expansion in powers of ε. We assume that A and B have an expansion in powers of ε of the form

$$A[\alpha(s) + \varepsilon y] = \sum_{n=0}^{\infty}\tilde{A}_n(s)\varepsilon^n y^n \tag{6.3.9}$$

$$B[\alpha(s) + \varepsilon y] = \sum_{n=0}^{\infty}\tilde{B}_n(s)\varepsilon^n y^n \tag{6.3.10}$$

and expand $\hat{p}$ in powers of ε:

$$\hat{p} = \sum_{n=0}^{\infty}\hat{p}_n\varepsilon^n\,. \tag{6.3.11}$$

Substituting these expansions into the FPE (6.3.8), we get by equating coefficients

$$\frac{\partial\hat{p}_0}{\partial s} = -\tilde{A}_1(s)\frac{\partial}{\partial y}(y\hat{p}_0) + \frac{1}{2}B_0(s)\frac{\partial^2\hat{p}_0}{\partial y^2} \tag{6.3.12}$$

$$\frac{\partial \hat{p}_r}{\partial s} = -\frac{\partial}{\partial y}[\tilde{A}_1(s)y\hat{p}_1 + \tilde{A}_2(s)y^2\hat{p}_0] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[\tilde{B}_0(s)\hat{p}_1 + \tilde{B}_1(s)y\hat{p}_0] \tag{6.3.13}$$

and, in general,

$$\frac{\partial \hat{p}_r}{\partial s} = -\frac{\partial}{\partial y}\left[\sum_{m=0}^{r} y^{r-m+1}\tilde{A}_{r-m+1}(s)\hat{p}_m\right] + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[\sum_{m=0}^{r} y^{r-m}\tilde{B}_{r-m}(s)\hat{p}_m\right]. \tag{6.3.14}$$

Only the equation for $\hat{p}_0$ is a FPE and, as mentioned in Sect.6.1, only $\hat{p}_0$ is a probability. The first equation in the hierarchy, (6.3.12), is a time-dependent Ornstein-Uhlenbeck process which corresponds exactly to (6.2.13), the first equation in the hierarchy for the stochastic differential equation. Thereafter the correspondence ceases.

The boundary conditions on the $\hat{p}_r$ do present technical difficulties since the transformation from x to y is time dependent, and a boundary at a fixed position in the x variable corresponds to a moving boundary in the y variable. Further, a boundary at $x = a$ corresponds to one at

$$y = \frac{[a - \alpha(s)]}{\varepsilon} \tag{6.3.15}$$

which approaches $\pm\infty$ as $\varepsilon \to 0$. There does not seem to be any known technique of treating such boundaries, except when $a = \pm\infty$, so that the y boundary is also at $\pm\infty$ and hence constant. Boundary conditions then assume the same form in the y variable as in the x variable.

In the case where the boundaries are at infinity, the result of the transformation (6.3.3) is to change a singular perturbation problem (6.3.1). (in which the limit $\varepsilon \to 0$ yields an equation of lower order) into an ordinary perturbation problem (6.3.8) in which the coefficients of the equation depend smoothly on ε, and the limit $\varepsilon \to 0$ is an equation of 2nd order. The validity of the expansion method will depend on the form of the coefficients.

6.3.1 Equations for Moments and Autocorrelation Functions

The hierarchy (6.3.14) is not very tractable, but yields a relatively straightforward procedure for computing the moments perturbatively. We assume that the boundaries are at $\pm\infty$ so that we can integrate by parts and discard surface terms. Then we define

$$\langle [y(t)]^n \rangle = \sum_{r=0}^{\infty} \varepsilon^r M_r^n(t)\,. \tag{6.3.16}$$

Then clearly

$$M_r^n(t) = \int dy\, y^n \hat{p}_r(y, t) \tag{6.3.17}$$

Then using (6.3.12–14), we easily derive by integrating by parts

$$\frac{dM_r^n(t)}{dt} = \sum_{m=0}^{r} \left[n\, \tilde{A}_{r-m+1}(t)\, M_m^{n+r-m}(t) + \frac{n(n-1)}{2}\, \tilde{B}_{r-m}(t)\, M_m^{n+r-m-2}(t) \right] \quad (6.3.18)$$

which is a closed hierarchy of equations, since the equation for $M_r^n(t)$ can be solved it all $M_m^p(t)$ are known for $m < r$ and $p < n + r$.

Writing out the first few for the mean $M_r^1(t)$ and mean square $M_r^2(t)$, we find

$$\frac{dM_0^1(t)}{dt} = \tilde{A}_1(t) M_0^1(t) \quad (6.3.19)$$

$$\frac{dM_1^1(t)}{dt} = \tilde{A}_1(t) M_1^1(t) + \tilde{A}_2(t) M_0^2(t) \quad (6.3.20)$$

$$\frac{dM_2^1(t)}{dt} = \tilde{A}_1(t) M_2^1(t) + \tilde{A}_2(t) M_1^2(t) + \tilde{A}_3(t) M_0^3(t) \quad (6.3.21)$$

$$\frac{dM_0^2(t)}{dt} = 2\tilde{A}_1(t) M_0^2(t) + \tilde{B}_0(t) \quad (6.3.22)$$

$$\frac{dM_1^2(t)}{dt} = 2\tilde{A}_1(t) M_1^2(t) + 2\tilde{A}_2(t) M_0^3(t) + \tilde{B}_1(t) M_0^1(t) \quad (6.3.23)$$

$$\frac{dM_0^3(t)}{dt} = 3\tilde{A}_1(t) M_0^3(t) + 3\tilde{B}_0(t) M_0^1(t)\,. \quad (6.3.24)$$

In deriving the last two equations we note that

$$M_r^0(t) = \int dy\, \hat{p}_r(y, t) \quad (6.3.25)$$

and using

$$\int dy\, \hat{p}(y, t) = 1 = \sum_r \varepsilon^r M_r^0(t) \quad (6.3.26)$$

we see that

$$M_0^0(t) = 1 \quad (6.3.27)$$

$$M_r^0(t) = 0 \qquad r \neq 1\,. \quad (6.3.28)$$

The equations are linear ordinary differential equations with inhomogenieties that are computed from lower equations in the hierarchy.

a) Stationary Moments

These are obtained by letting $t \to \infty$ and setting the left-hand side of (6.3.18) equal to zero. (All coefficients, $\tilde{A}$, $\tilde{B}$, etc. are taken time independent.)

From (6.3.19–22) we find

$$M_0^1(\infty) = 0$$

$$M_0^2(\infty) = -\tfrac{1}{2}\tilde{B}_0/\tilde{A}_1$$

$$
\begin{aligned}
&M_0^3(\infty) = 0 \\
&M_1^1(\infty) = -\tilde{A}_2 M_0^2(\infty)/\tilde{A}_1 = \tfrac{1}{2}\,\tilde{A}_2\tilde{B}_0/(\tilde{A}_1)^2 \\
&M_1^2(\infty) = -\tilde{A}_2 M_0^3(\infty)/\tilde{A}_1 = 0 \\
&M_2^1(\infty) = -[\tilde{A}_2 M_1^2(\infty) + \tilde{A}_3 M_0^3(\infty)]/\tilde{A}_1 = 0\,.
\end{aligned}
\tag{6.3.29}
$$

Thus, the stationary mean, variance, etc., are

$$
\begin{aligned}
\langle x\rangle_s &= \alpha + \varepsilon[M_0^1(\infty) + \varepsilon M_1^1(\infty) + \varepsilon^2 M_2^1(\infty)] \\
&= \alpha + \tfrac{1}{2}\varepsilon^2\tilde{A}_2\tilde{B}_0/(\tilde{A}_1)^2
\end{aligned}
\tag{6.3.30}
$$

$$
\begin{aligned}
\mathrm{var}\{x\}_s &= \langle x^2\rangle_s - \langle x_s\rangle^2 \\
&= \langle(\alpha + \varepsilon y)^2\rangle_s - \langle\alpha + \varepsilon y\rangle_s^2 = \varepsilon^2\mathrm{var}\{y\}_s \\
&= -\frac{\varepsilon^2}{2}\tilde{B}_0/\tilde{A}_1 \qquad \text{to order } \varepsilon^2\,.
\end{aligned}
\tag{6.3.31}
$$

The procedure can clearly be carried on to arbitrarily high order. Of course in a one-variable system, the stationary distribution can be evaluated exactly and the moments found by integration. But in many variable systems this is not always possible, whereas the multivariate extension of this method is always able to be carried out.

b) Stationary Autocorrelation Function

The autocorrelation function of x is simply related to that of y in a stationary state by

$$
\langle x(t)x(0)\rangle_s = \alpha^2 + \varepsilon^2\langle y(t)y(0)\rangle_s \tag{6.3.32}
$$

and a hierarchy of equations for $\langle y(t)y(0)\rangle$ is easily developed. Notice that

$$
\begin{aligned}
\frac{d}{dt}\langle y(t)^n y(0)\rangle_s = \Big\langle\Big\{&\frac{A[\alpha + \varepsilon y(t)] - A(\alpha)}{\varepsilon}\, n y(t)^{n-1} \\
&+ \tfrac{1}{2}n(n-1)B[\alpha + \varepsilon y(t)]y(t)^{n-2}\Big\}\, y(0)\Big\rangle_s
\end{aligned}
\tag{6.3.33}
$$

which can be derived by using the FPE (6.3.1) for $p(y, t\,|\,y_0, t_0)$ and integrating by parts, or by using Ito's formula for the corresponding SDE.

Using the definition of $\tilde{A}_r$, $\tilde{B}_r$ in (6.3.9.10) and expanding A and B in a power series, we get

$$
\begin{aligned}
\frac{d}{dt}\langle y(t)^n y(0)\rangle_s = \sum_{q=0}^{\infty}\varepsilon^q\Big[&n\tilde{A}_{q+1}\langle y(t)^{q+n}y(0)\rangle_s \\
&+ \frac{n(n-1)}{2}\,\tilde{B}_q\langle y(t)^{q+n-2}y(0)\rangle_s\Big]\,.
\end{aligned}
\tag{6.3.34}
$$

These equations themselves form a hierarchy which can be simply solved in a power

series in ε. Normally one is most interested in $\langle y(t)y(0)\rangle_s$, which can be calculated to order ε^q, provided one knows $\langle y(t)^p y(0)_s\rangle$ for $p \leqslant q+1$. We have the initial condition

$$\langle y(0)^n y(0)\rangle_s = \langle y^{n+1}\rangle_s \tag{6.3.35}$$

and the stationary moments can be evaluated from the stationary distribution, or as has just been described.

6.3.2 Example

We consider the Fokker-Planck equation

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}[(x + x^3)p] + \frac{\varepsilon^2}{6}\frac{\partial^2 p}{\partial x^2} \tag{6.3.36}$$

for which we have [in the stationary state $\alpha(t) = 0$]

$$\begin{aligned}
&\tilde{A}_1 = -1\\
&\tilde{A}_2 = 0\\
&\tilde{A}_3 = -1\\
&\tilde{A}_q = 0 \ (q > 3)\\
&\tilde{B}_q = \tfrac{1}{3}\delta_{q,0}\\
&\alpha \ = 0\,.
\end{aligned} \tag{6.3.37}$$

Using (6.3.30, 31) we have

$$\begin{aligned}
&\langle x\rangle_s = 0\\
&\operatorname{var}\{x\}_s = \varepsilon^2/6\,.
\end{aligned} \tag{6.3.38}$$

For convenience, let us use the notation

$$c_n(t) = \langle y^n(t)y(0)\rangle_s \tag{6.3.39}$$

so that the equations for the c_1 and c_3 are

$$\begin{bmatrix} \dfrac{dc_1}{dt} \\ \dfrac{dc_3}{dt} \end{bmatrix} = \begin{bmatrix} -1 & -\varepsilon^2 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_3 \end{bmatrix} \tag{6.3.40}$$

[the equations for the c_{2n} decouple from those for c_{2n+1} because $B(x)$ is constant and $A(x)$ is an odd function of x].

It is simpler to solve (6.3.40) exactly than to perturb. The eigenvalues of the matrix are

$$\lambda_1 = -2 + \sqrt{1 - \varepsilon^2}$$
$$\lambda_2 = -2 - \sqrt{1 - \varepsilon^2} \tag{6.3.41}$$

and the corresponding eigenvectors

$$\boldsymbol{u}_1 = \begin{bmatrix} 1 + \sqrt{1 - \varepsilon^2} \\ 1 \end{bmatrix}$$
$$\boldsymbol{u}_2 = \begin{bmatrix} 1 - \sqrt{1 - \varepsilon^2} \\ 1 \end{bmatrix}. \tag{6.3.42}$$

The solution of (6.3.40) can then be written

$$\begin{bmatrix} c_1(t) \\ c_3(t) \end{bmatrix} = \alpha_1 e^{-\lambda_1 t} \boldsymbol{u}_1 + \alpha_2 e^{-\lambda_2 t} \boldsymbol{u}_2 \qquad (t > 0) . \tag{6.3.43}$$

The initial condition is

$$c_1(0) = \langle y^2 \rangle_s$$
$$c_3(0) = \langle y^4 \rangle_s . \tag{6.3.44}$$

We can compute $\langle y^4 \rangle_s$ using the moment hierarchy (6.3.10) extended to M_0^4: we find

$$M_0^4 = -\frac{3}{2} \frac{\tilde{B}_0 M_0^2}{\tilde{A}_1} = \frac{1}{12} \tag{6.3.45}$$

then $\langle y^4 \rangle = 1/12$.

Hence, we obtain

$$\begin{bmatrix} 1/6 \\ 1/12 \end{bmatrix} = \alpha_1 \boldsymbol{u}_1 + \alpha_2 \boldsymbol{u}_2 \tag{6.3.46}$$

which have the solutions

$$\alpha_1 = \frac{1}{24}(1 + \sqrt{1 - \varepsilon^2})/\sqrt{1 - \varepsilon^2}$$
$$\alpha_2 = \frac{1}{24}(-1 + \sqrt{1 - \varepsilon^2})/\sqrt{1 - \varepsilon^2} . \tag{6.3.47}$$

The correlation function is, to 2nd order in ε (many terms cancel)

$$c_1(t) = \frac{1}{6} e^{-|\lambda_1| t} . \tag{6.3.48}$$

Notice that the eigenvalues λ_1 and λ_2 depend on ε^2. Any attempt to solve the

system (6.3.40) perturbatively would involve expanding $\exp(\lambda_1 t)$ and $\exp(\lambda_2 t)$ in powers of ε^2 and would yield terms like $t^N \exp(-2t)$ which would not be an accurate representation of the long-time behaviour of the autocorrelation function.

The x correlation function is

$$\langle x(t)x(0)\rangle_s = \varepsilon^2 c_1(t) \tag{6.3.49}$$

and the spectrum

$$S(\omega) = \varepsilon^2 \int_{-\infty}^{\infty} dt\, e^{-i\omega t} c_1(t)/2\pi \tag{6.3.50}$$

$$= \frac{\varepsilon^2}{12\pi} \frac{1}{\lambda_1^2 + \omega^2}\Big]. \tag{6.3.51}$$

6.3.3 Asymptotic Method for Stationary Distributions

For an arbitrary Fokker-Planck equation

$$\partial_t p = -\sum_i \partial_i A_i(\boldsymbol{x}) p + \tfrac{1}{2}\varepsilon^2 \sum_{i,j} \partial_i \partial_j B_{ij}(\boldsymbol{x}) p\,, \tag{6.3.52}$$

one can generate an asymptotic expansion for the stationary solution by setting

$$p_s(\boldsymbol{x}) = \exp[-\phi(\boldsymbol{x})/\varepsilon^2] \tag{6.3.53}$$

in terms of which we find

$$\Big[\sum_i A_i(\boldsymbol{x})\partial_i\phi + \tfrac{1}{2}\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\phi\partial_j\phi\Big] + \varepsilon^2\Big[-\sum_i \partial_i A_i(\boldsymbol{x}) + \sum_{i,j} \partial_i B_{ij}\partial_j\phi + \sum_{i,j} \partial_i\partial_i B_{ij}(\boldsymbol{x})\Big] = 0\,. \tag{6.3.54}$$

The first term, which is of order ε^0, is a Hamilton Jacobi equation. The main significance of the result is that an asymptotic expansion for $\phi(\boldsymbol{x})$ can be, in principle, developed:

$$\phi(\boldsymbol{x}) = \sum_n \varepsilon^n \phi_n(\boldsymbol{x}) \tag{6.3.55}$$

where $\phi_0(x)$ satisfies

$$\sum_i A_i(\boldsymbol{x})\partial_i\phi_0 + \tfrac{1}{2}\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\phi_0\partial_j\phi_0 = 0 \tag{6.3.56}$$

Graham and *Tel* [6.8, 9] have recently shown how equation (6.3.56) may be solved in the general case. Their main result is that solutions, though continuous, in general have infinitely many discontinuities in their derivatives, except in certain special cases, which are closely related to the situation in which the FPE satisfies potential conditions.

6.4 Adiabatic Elimination of Fast Variables

It is very often the case that a dynamical system can be described by stochastic equations which have widely different response times and in which the behaviour on a very short time scale is not of interest. The most natural example of this occurs in Brownian motion. Here it is normal to observe only the position of the Brownian particle, but the fundamental equations involve the momentum as well, which is normally unobservable. Thus, Langevin's equation (1.2.14) can be rewritten as an equation for position and velocity as

$$\frac{dx}{dt} = v \tag{6.4.1}$$

$$m\frac{dv}{dt} = -\beta v + \sqrt{2k\beta T}\,\xi(t)\,. \tag{6.4.2}$$

If we interpret the equations as Ito stochastic differential equations, the method of solution has already been given in Sect.4.4.6. However, it is simpler to integrate (6.4.2) first to give the solution

$$v(t) = v(0)\exp(-\beta t/m) + \frac{\sqrt{2kT\beta}}{m}\int_0^t \exp[-\beta(t-t')/m]\xi(t')dt'\,. \tag{6.4.3}$$

We now want to consider the situation in which the friction coefficient β is not small but the mass m is very small. Then for times t such that

$$t \gg m/\beta \equiv \tau\,, \tag{6.4.4}$$

the exponential in the first term will be negligible and the lower limit in the integral will be able to be extended to $-\infty$, without significant error. Hence,

$$v(t) \to \frac{\sqrt{2kT\beta}}{m}\int_{-\infty}^t \exp[-(t-t')/\tau]\xi(t')dt'\,. \tag{6.4.5}$$

Here τ will be called *relaxation time* since it determines the time scale of relaxation to (6.4.5).

Let us define

$$\eta(t,\tau) = \tau^{-1}\int_{-\infty}^t \exp[-(t-t')/\tau]dW(t') \tag{6.4.6}$$

which is, from Sect.4.4.4, a stationary Ornstein-Uhlenbeck process. The correlation function is

$$\langle\eta(t,\tau)\eta(t',\tau)\rangle = \frac{1}{2\tau}\exp(-|t-t'|/\tau) \tag{6.4.7}$$

$$\xrightarrow[\tau\to 0]{} \delta(t-t')\,. \tag{6.4.8}$$

We see that the limit $\tau \to 0$ corresponds to a white noise limit in which the correlation function becomes a delta function.

Thus, we can write (6.4.1) as

$$\frac{dx}{dt} = \sqrt{\frac{2kT}{\beta}}\,\eta(t, \tau) \tag{6.4.9}$$

and in the limit $\tau \to 0$, this should become

$$\frac{dx}{dt} = \sqrt{\frac{2kT}{\beta}}\,\xi(t)\,. \tag{6.4.10}$$

An alternative, and much more transparent way of looking at this is to say that in (6.4.2) the limit $m \to 0$ corresponds to setting the left-hand side equal to zero, so that

$$v(t) = \sqrt{\frac{2kT}{\beta}}\,\xi(t)\,. \tag{6.4.11}$$

The reasoning here is very suggestive but completely nonrigorous and gives no idea of any systematic approximation method, which should presumably be some asymptotic expansion in a small dimensionless parameter. Furthermore, there does not seem to be any way of implementing such an expansion directly on the stochastic differential equation—at least to the author's knowledge no one has successfully developed such a scheme.

The Fokker-Planck equation equivalent to (6.4.1, 2) for the distribution function $p(x, v, t)$ is

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(vp) + \frac{\partial}{\partial v}\left(\frac{vp}{\tau}\right) + \frac{kT}{\tau^2\beta}\frac{\partial^2 p}{\partial v^2}\,. \tag{6.4.12}$$

We define the position distribution function $\hat{p}(x, t)$ by

$$\hat{p}(x, t) = \int_{-\infty}^{\infty} dv\, p(x, v, t)\,. \tag{6.4.13}$$

Then we expect that, corresponding to the "reduced" Langevin equation (6.4.10) the FPE for $\hat{p}(x, t)$ is

$$\frac{\partial \hat{p}}{\partial t} = \frac{kT}{\beta}\frac{\partial^2 \hat{p}}{\partial x^2}\,. \tag{6.4.14}$$

We seek a way of deriving (6.4.14) from (6.4.12) in some perturbative manner, so that we obtain higher corrections in powers of the appropriate small parameter.

More generally, we can consider Brownian motion in a potential for which the Langevin equations are (Sect.5.3.6)

$$\begin{aligned} \frac{dx}{dt} &= v \\ m\frac{dv}{dt} &= -\beta v - V'(x) + \sqrt{2\beta kT}\,\xi(t)\,. \end{aligned} \tag{6.4.15}$$

The limit of large β should result in very rapid relaxation of the second equation to a quasistationary state in which $dv/dt \to 0$. Hence, we assume that for large enough β,

$$v = -\frac{1}{\beta}[V'(x) - \sqrt{2\beta kT}\,\xi(t)] \tag{6.4.16}$$

and substituting in (6.4.15) we get

$$\frac{dx}{dt} = -\beta^{-1}V'(x) + \sqrt{\frac{2kT}{\beta}}\,\xi(t) \tag{6.4.17}$$

corresponding to a FPE for $\hat{p}(x)$ known as the *Smoluchowski equation:*

$$\boxed{\frac{\partial \hat{p}}{\partial t} = \beta^{-1}\left\{\frac{\partial}{\partial x}\left[V'(x)\hat{p} + kT\frac{\partial \hat{p}}{\partial x}\right]\right\}.} \tag{6.4.18}$$

In this case we have eliminated the *fast variable* v, which is assumed to relax very rapidly to the value given by (6.4.16).

This procedure is the prototype of all *adiabatic elimination* procedures which have been used as the basis of Haken's *slaving principle* [6.1]. The basic physical assumption is that large β (or, in general, short relaxation times) force the variables governed by equations involving large β (e.g., v) to relax to a value given by assuming the slow variable (in this case x) to be constant. Such fast variables are then effectively *slaved* by the slow variables.

Surprisingly, the problem of a rigorous derivation of the Smoluchowski equation and an estimation of corrections to it, has only rather recently been solved. The first treatment was by *Brinkman* [6.2] who only estimated the order of magnitude of corrections to (6.4.18) but did not give all the correction terms to lowest order. The first correct solution was by *Stratonovich* [Ref. 6.3, Chap. 4, Sect. 11.1]. Independently, *Wilemski* [6.4] and *Titulaer* [6.5] have also given correct treatments.

In the following sections we will present a systematic and reasonably general theory of the problem of the derivation of the Smoluchowski equation and corrections to it, and will then proceed to more general adiabatic elimination problems. The procedure used is an adaptation of projection operator methods, which have been used in statistical physics, quantum optics, and related fields for many years. These methods can be formulated directly in the time domain, but we will find it more convenient to use a Laplace transform method, which was that originally used by Wilemski. The manner of presentation is similar to that of *Papanicolaou* [6.6], who has given a rigorous basis to its use in some problems. However, the demonstrations used here will be largely formal in character.

6.4.1 Abstract Formulation in Terms of Operators and Projectors

Let us consider the rescaled form of the Fokker-Planck equation (6.4.12) derived as (5.3.97) in Sect. 5.3.6, which we can write in the form

$$\frac{\partial p}{\partial t} = (\gamma L_1 + L_2)p\,, \tag{6.4.19}$$

where L_1 and L_2 are differential operators given by

$$L_1 = \frac{\partial}{\partial u}\left(u + \frac{\partial}{\partial u}\right) \tag{6.4.20}$$

$$L_2 = -\frac{\partial}{\partial y}u + U'(y)\frac{\partial}{\partial u}. \tag{6.4.21}$$

We would like to derive an equation for the distribution function in y,

$$\hat{p}(y, t) = \int du\, p(u, y, t) \tag{6.4.22}$$

which would be valid in the limit where γ becomes very large.

It is expected that an approximate solution to (6.4.19) would be obtained by multiplying $\hat{p}(y, t)$ by the stationary distribution of

$$\frac{\partial p}{\partial t} = L_1 p\,, \tag{6.4.23}$$

that is, by

$$(2\pi)^{-1/2}\exp(-\tfrac{1}{2}u^2)\,. \tag{6.4.24}$$

The reasoning is that for large γ, the velocity distribution is very rapidly thermalised or, more crudely, we can in (6.4.19) neglect L_2 compared to L_1 and the solution is a function of y multiplied by a solution of (6.4.23), which approaches a stationary solution in a time of order γ^{-1}, which will be very small.
We formalise this by defining a *projection operator* P by

$$(Pf)(u, y) = (2\pi)^{-1/2}\exp(-\tfrac{1}{2}u^2)\int du\, f(u, y)\,. \tag{6.4.25}$$

where $f(u, y)$ is an arbitrary function. The reader may easily check that

$$P^2 = P\,. \tag{6.4.26}$$

In terms of the vector space of all functions of u and y, P is an operator which projects any vector into the subspace of all vectors which can be written in the form of

$$g(u, y) = (2\pi)^{-1/2}\exp(-\tfrac{1}{2}u^2)\,\hat{g}(y)\,, \tag{6.4.27}$$

where $\hat{g}(y)$ is an arbitrary function of y. However, functions of the form (6.4.27) are all solutions of

$$L_1 g = 0\,, \tag{6.4.28}$$

that is, *the space into which P projects is the null space of* L_1. We may also note that in this case

$$\boxed{P = \lim_{t\to\infty} [\exp(L_1 t)]\,.} \tag{6.4.29}$$

To prove this, expand any function of u and y in eigenfunctions $P_\lambda(u)$ of L_1 as in Sect.5.2.5:

$$f(u, y) = \sum_\lambda A_\lambda(y)\, P_\lambda(u)\,, \tag{6.4.30}$$

where

$$A_\lambda(y) = \int du\, Q_\lambda(u) f(u, y)\,. \tag{6.4.31}$$

Then

$$\lim_{t\to\infty} [\exp(L_1 t) f(u, y)] = \sum_\lambda A_\lambda(y) \lim_{t\to\infty} \mathrm{e}^{-\lambda t} P_\lambda(u) \tag{6.4.32}$$

$$= P_0(u) \int du\, Q_0(u) f(u, y) \tag{6.4.33}$$

and noting that for this process (Ornstein-Uhlenbeck)

$$P_0(u) = (2\pi)^{-1/2} \exp(-\tfrac{1}{2}u^2) \tag{6.4.34}$$

$$Q_0(u) = 1\,. \tag{6.4.35}$$

We see that (6.4.29) follows.

In this case and in all other cases, we also have the essential relation

$$\boxed{PL_2P = 0\,,} \tag{6.4.36}$$

For, considering $PL_2Pf(u, y)$, we see from the definition of L_2, that $L_2Pf(u, y)$ is proportional to

$$u \exp(-\tfrac{1}{2}u^2) \propto P_1(u) \tag{6.4.37}$$

and

$$P\, P_1(u) = 0 \tag{6.4.38}$$

either by explicit substitution or by noting that $P_1(u)$ is not in the null space of L_1. Let us define

$$v = Pp \tag{6.4.39}$$

$$w = (1 - P)p \tag{6.4.40}$$

so that

$$p = v + w$$

and v is in the null space of L_1, while w is not in the null space of L_1.

We can now note that, from (6.4.29),

$$PL_1 = L_1P = 0 \tag{6.4.41}$$

and from the original equation we have

$$\begin{aligned}\frac{\partial v}{\partial t} &= P(\gamma L_1 + L_2)p \\ &= P(\gamma L_1 + L_2)[Pp + (1 - P)p] \\ &= PL_2(1 - P)p\,,\end{aligned}$$

[where we have used (6.4.41) and (6.4.36)] so that

$$\boxed{\frac{\partial v}{\partial t} = PL_2w\,.} \tag{6.4.42}$$

Similarly,

$$\begin{aligned}\frac{\partial w}{\partial t} &= (1- P)(\gamma L_1 + L_2)p = (1 - P)(\gamma L_1 + L_2)[Pp + (1 - P)p] \\ &= \gamma L_1(1 - P)p + (1 - P)L_2(1 - P)p + (1 - P)L_2Pp\end{aligned} \tag{6.4.43}$$

and using $PL_2P = 0$, we have

$$\boxed{\frac{\partial w}{\partial t} = \gamma L_1w + (1 - P)L_2w + L_2v\,.} \tag{6.4.44}$$

6.4.2 Solution Using Laplace Transform

The fundamental equations (6.4.42, 44) can be solved in a number of iterative ways. However, since they are linear equations, a solution in terms of the Laplace transform can be very appropriate and readily yields a perturbation expansion.

The Laplace transform of any function of time $f(t)$ is defined by

$$\tilde{f}(s) = \int_0^\infty e^{-st} f(t) \tag{6.4.45}$$

and it may quite readily be defined for operators and abstract vectors in exactly the same way. Our fundamental equations become, on using

$$\int_0^\infty e^{-st} \frac{df}{dt} = s\tilde{f}(s) - f(0)\,, \tag{6.4.46}$$

$$\begin{aligned} s\tilde{v}(s) &= PL_2\tilde{w}(s) + v(0)\,, \\ s\tilde{w}(s) &= [\gamma L_1 + (1 - P)L_2]\tilde{w}(s) + L_2\tilde{v}(s) + w(0)\,. \end{aligned} \tag{6.4.47}$$

These equations are linear operator equations, whose formal solution is straightforward. For simplicity, let us first assume

$$w(0) = 0 \tag{6.4.48}$$

which means that the initial distribution is assumed to be of the form

$$p(u, y, 0) = (2\pi)^{-1/2} \exp(-\tfrac{1}{2}u^2)\,\hat{p}(y, 0)\,, \tag{6.4.49}$$

that is, the initial thermalisation of the velocity u is assumed. Then we have formally

$$\tilde{w}(s) = [s - \gamma L_1 - (1 - P)L_2]^{-1} L_2\tilde{v}(s) \tag{6.4.50}$$

and hence,

$$\boxed{s\tilde{v}(s) = PL_2[s - \gamma L_1 - (1 - P)L_2]^{-1} L_2\tilde{v}(s) + v(0)\,.} \tag{6.4.51}$$

We have here, at least formally, the complete solution of the problem. For any *finite* s, we can take the large γ limit to find

$$s\tilde{v}(s) \simeq -\gamma^{-1} PL_2 L_1^{-1} L_2\tilde{v}(s) + v(0)\,. \tag{6.4.52}$$

Notice that L_1^{-1} does not always exist. However, we know that

$$PL_2\tilde{v}(s) = PL_2 P\,\tilde{p}(s) = 0 \tag{6.4.53}$$

from (6.4.36). Hence, $L_2\tilde{v}(s)$ contains no component in the null space of L_1 and thus $L_1^{-1}L_2\tilde{v}(s)$ exists.

In this case of Kramers' equation, let us now see what (6.4.52) looks like. It is equivalent to the differental equation

$$\boxed{\frac{\partial v}{\partial t} = -\gamma^{-1} PL_2 L_1^{-1} L_2 v\,.} \tag{6.4.54}$$

Now note that

$$L_2 v = \left[-\frac{\partial}{\partial y} u + U'(y) \frac{\partial}{\partial u} \right] P_0(u) \int du'\, p(u', y, t) \tag{6.4.55}$$

[with $P_0(u)$ defined by (6.4.34)].

For this problem is it useful to bring in the eigenfunctions of the Ornstein-Uhlenbeck process [Sect.5.2.6c] which for these parameters,

$$\tfrac{1}{2}D = k = 1 \tag{6.4.56}$$

take the form

$$P_n(u) = (2\pi)^{-1/2} \exp(-\tfrac{1}{2}u^2) Q_n(u) \tag{6.4.57}$$

with

$$Q_n(u) = (2^n n!)^{-1/2} \mathrm{H}_n(u/\sqrt{2}) . \tag{6.4.58}$$

Using

$$L_1 P_n(u) = -nP_n(u) \tag{6.4.59}$$

and the recursion formulae for Hermite polynomials

$$x\mathrm{H}_n(x) = \tfrac{1}{2}\mathrm{H}_{n+1}(x) + n\mathrm{H}_{n-1}(x) \tag{6.4.60}$$

$$\frac{d}{dx}[\mathrm{e}^{-x^2}\mathrm{H}_n(x)] = -\mathrm{e}^{-x^2}\mathrm{H}_{n+1}(x) . \tag{6.4.61}$$

we see that

$$L_2 v = -\left[U'(y) + \frac{\partial}{\partial y} \right] P_1(u)\hat{p}(y) \tag{6.4.62}$$

so that, using (6.4.59),

$$L_1^{-1} L_2 v = \left[U'(y) + \frac{\partial}{\partial y} \right] P_1(u)\hat{p}(y) . \tag{6.4.63}$$

We now apply L_2 once more and use the relations (6.4.60, 61) again.

We find

$$L_2 P_1(u) = -\left[\sqrt{2}\, P_2(u) + P_0(u) \right] \frac{\partial}{\partial y} - \sqrt{2}\, P_2(u) U'(y) \tag{6.4.64}$$

$$PL_2 L_1^{-1} L_2 v = -\frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y} \right] \hat{p}(y) P_0(u) \tag{6.4.65}$$

and the equation of motion (6.4.54) takes the form [after cancelling the factor $P_0(u)$]

$$\boxed{\frac{\partial \hat{p}}{\partial t} = \gamma^{-1} \frac{\partial}{\partial y}\left[U'(y)\hat{p} + \frac{\partial \hat{p}}{\partial y}\right]} \tag{6.4.66}$$

which is exactly the Smoluchowski equation, derived using the naive elimination given in Sect. 6.4.

6.4.3 Short-Time Behaviour

One should note that the limit $\gamma \to \infty$ in (6.4.52) implies that s is finite, or roughly speaking

$$\gamma \gg s\,. \tag{6.4.67}$$

Thus, the solution for the Lapace transform will only be valid provided

$$s \ll \gamma \tag{6.4.68}$$

or for the solution, this will mean that

$$t \gg \gamma^{-1}. \tag{6.4.69}$$

Let us define

$$s_1 = s\gamma^{-1} \tag{6.4.70}$$

so that (6.4.51) becomes

$$\gamma s_1 \tilde{v} = PL_2[s_1\gamma - L_1\gamma - (1 - P)L_2]^{-1}L_2\tilde{v} + v(0)\,. \tag{6.4.71}$$

The limit $\gamma \to \infty$ gives

$$\gamma s_1 \tilde{v} \simeq \gamma^{-1}PL_2(s_1 - L_1)^{-1}L_2\tilde{v} + v(0)\,. \tag{6.4.72}$$

Using the fact that $L_2\tilde{v}$ is proportional to $P_1(u)$ (6.4.62), we see that

$$\gamma s_1 \tilde{v} = \gamma^{-1}(s_1 + 1)^{-1}PL_2^2\tilde{v} + v(0)\,. \tag{6.4.73}$$

Changing back to the variable s again and rearranging, we find

$$s\tilde{v} = \gamma^{-1}\left(\frac{s}{\gamma} + 1\right)^{-1}PL_2^2\tilde{v} + v(0) \tag{6.4.74}$$

which is equivalent to

$$\frac{\partial v}{\partial t} = \int_0^t dt' \exp\left[\gamma(t' - t)\right]PL_2^2 v(t')dt\,. \tag{6.4.75}$$

Alternatively, we can rewrite (6.4.74) in the form

$$\frac{1}{\gamma}[s^2\tilde{v} - sv(0)] + [s\tilde{v} - v(0)] = \gamma^{-1}PL_2^2 v \tag{6.4.76}$$

which, using (6.4.46) and the same working out as in Sect. 6.4.2, is equivalent to the equation for $\hat{p}$:

$$\boxed{\frac{1}{\gamma}\frac{\partial^2 p}{\partial t^2} + \frac{\partial \hat{p}}{\partial t} = \gamma^{-1}\frac{\partial}{\partial y}\left[U'(y)\hat{p} + \frac{\partial \hat{p}}{\partial y}\right]} \tag{6.4.77}$$

in which the initial condition

$$\frac{\partial \hat{p}}{\partial t}(0) = 0$$

is implied because

$$\int_0^\infty e^{-st} f''(t) = s^2\tilde{f}(s) - sf(0) - f'(0)$$

and no constant term appears in the first bracket in (6.4.76). We may smiilarly rewrite (6.4.75) or integrate (6.4.77) to get

$$\boxed{\frac{\partial \hat{p}}{\partial t} = \frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right]\int_0^t dt' \exp[\gamma(t' - t)]\hat{p}(t').} \tag{6.4.78}$$

Equations (6.4.77, 78) demonstrate a non-Markov nature, seen explicitly in (6.4.78) which indicates that the prediction of $\hat{p}(t + \Delta t)$ requires the knowledge of $\hat{p}(t')$ for $0 \leqslant t' \leqslant t$. However, the kernel $\exp[\gamma(t' - t)]$ is significantly different from zero only for $|t' - t| \sim \gamma^{-1}$ and on a time scale much longer than this, (6.4.78) is approximated by the Smoluchowski equation (6.4.66). Formally, we achieve this by integrating by parts in (6.4.78):

$$\int_0^t dt' \exp[\gamma(t' - t)]\hat{p}(t') = \frac{\hat{p}(t) - e^{-\gamma t}\hat{p}(0)}{\gamma} - \gamma^{-1}\int_0^t \exp[\gamma(t' - t)]\frac{\partial \hat{p}}{\partial t}dt'. \tag{6.4.79}$$

Neglecting the last term as being of order γ^{-2}, we find the Smoluchowski equation replaced by

$$\boxed{\frac{\partial \hat{p}}{\partial t} = \frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right]\left[\hat{p}(t) - e^{-\gamma t}\hat{p}(0)\right]} \tag{6.4.80}$$

This equation is to lowest order in γ equivalent to (6.4.78) for *all times*, that is, very short ($<\gamma^{-1}$) and very long ($\gg\gamma^{-1}$) times. It shows the characteristic "*memory time*", γ^{-1}, which elapses before the equation approaches the Smoluchowski form.

To this order then, the process can be approximated by a Markov process, but to the same order, there are the alternative expressions (6.4.77, 78) which are not Markov processes. Clearly, to any higher order, we must have a non-Markov proces.

6.4.4 Boundary Conditions

Let us consider a barrier to be erected at $y = a$ with the particle confined to $y \leqslant a$. The behaviors for $u > 0$ and $u < 0$ are distinct.

From the stochastic differential equations

$$\begin{aligned} dy &= u\,dt \\ du &= -[U'(y) + \gamma u]dt + \sqrt{2\gamma}\,dW(t) \end{aligned} \tag{6.4.81}$$

We see that

$$\begin{aligned} &\text{for } u > 0, \quad y = a \text{ is an exit boundary} \\ &\text{for } u < 0, \quad y = a \text{ is an entrance boundary.} \end{aligned}$$

since a particle with $u > 0$ at $y = a$ must proceed to $y > a$ or be absorbed. Similarly, particles to the left of the boundary can never reach $y = a$ if $u < 0$. Conventionally, we describe $y = a$ as absorbing or reflecting as follows:

i) *absorbing barrier:* particle absorbed for $u > 0$, no particles with $u < 0$

$$\begin{aligned} \Rightarrow p(u, a, t) &= 0 \quad u > 0 \\ &= 0 \quad u < 0 \end{aligned} \tag{6.4.82}$$

The first condition is the usual absorbing boundary condition, as derived in Sect. 5.2.1. The second condition expresses the fact that any particle placed with $u < 0$ at $y = a$ *immediately* proceeds to $y < a$, and no further particles are introduced.

The absorbing barrier condition clearly implies

$$\hat{p}(a, t) = 0 \tag{6.4.83}$$

which is the usual absorbing barrier condition in a one variable Fokker-Planck equation.

ii) "*reflecting*" *barrier*: physically, a reflection at $y = a$ implies that the particle reaches $y = a$ with the velocity u, and is immediately reflected with a different but negative veloicty. If we *assume* that

$$u \to -u\,,$$

then we have a "periodic boundary condition" as in Sects. 5.2.1 and 5.3.2 This means that

$$p(u, a, t) = p(-u, a, t) \tag{6.4.84}$$

and that the normal component of the current leaving at (u, a) is equal to that

entering at $(-u, a)$. However, since for this equation

$$J = \left\{-up,\ [\gamma u + U'(y)]\, p + \gamma \frac{\partial p}{\partial u}\right\} \tag{6.4.85}$$

and the normal component is the component along the direction y, we see that (6.4.84, 85) are equivalent.

This boundary condition leads very naturally to that for the Smoluchowski equation.

(For (6.4.84) implies that only even-order eigenfunctions $P_n(u)$ can occur in the expansion of $p(u, a, t)$. Hence,

$$(1 - P)p(u, a, t) = w(u, a, t) \tag{6.4.86}$$

contains only *even* eigenfunctions, and the same is true of $\tilde{w}(u, a, s)$, the Laplace transform. But, from (6.4.50) we see that to lowest order in γ^{-1}

$$\tilde{w}(u, a, s) \simeq (-\gamma^{-1}L^{-1}L_2\tilde{v})(u, a, s) \tag{6.4.87}$$

and using (6.4.63)

$$= -P_1(u)\gamma^{-1}\left[U'(y)\hat{p}(y) + \frac{\partial \hat{p}(y)}{\partial y}\right]_{y=a}. \tag{6.4.88}$$

Since this is proportional to the *odd* eigenfunction $P_1(u)$, it vanishes. Hence we derive

$$\left[U'(y)\hat{p}(y) + \frac{\partial \hat{p}(y)}{\partial y}\right]_{y=a} = 0 \tag{6.4.89}$$

which is the correct reflecting barrier coundary condition for the Smoluchowski equation.

It is not difficult to show similarly that the same boundary conditions can be derived for the equations derived in Sect.6.4.3.

6.4.5 Systematic Perturbative Analysis

We return now to (6.4.47), but again with $w(0) = 0$ for simplicity.

Then we have

$$\tilde{w}(s) = [s - \gamma L_1 - (1 - P)L_2]^{-1}L_2\tilde{v}(s) \tag{6.4.90}$$

and

$$s\,\tilde{v}(s) = PL_2[s - \gamma L_1 - (1 - P)L_2]^{-1}L_2\tilde{v}(s) + v(0)\,. \tag{6.4.91}$$

We can straightforwardly expand the inverse in (6.5.91) in powers of γ. However, the order of magnitude of s in this expansion must be decided. From the previous sections we see that there is a possibility of defining

$$s_1 = s\gamma^{-1}, \qquad \tilde{v} \to \gamma^{-1}\tilde{v} \tag{6.4.92}$$

which gives rise to an expansion in which the initial time behaviour in times of order γ^{-1} is explicitly exhibited.

An alternative is to define

$$s_2 = s\gamma, \qquad \tilde{v} \to \gamma\tilde{v} \tag{6.4.93}$$

which we will find yields an expansion in which only the long-time behaviour is exhibited.

By substituting (6.4.93) we find that (6.4.91) takes the form

$$s_2\tilde{v} = -PL_2[L_1 + (1 - P)L_2\gamma^{-1} - s_2\gamma^{-2}]^{-1}L_2\tilde{v} + v(0) \tag{6.4.94}$$

and we may now write

$$s_2\tilde{v} = -PL_2L_1^{-1}L_2\tilde{v} + v(0) \qquad (\gamma \to \infty) \tag{6.4.95}$$

whereas, without this substitution, the limit $\gamma \to \infty$ yields simply

$$s\tilde{v} = v(0)\,. \tag{6.4.96}$$

This, is in fact, in agreement with (6.4.52), where we did not actually go to the limit $\gamma \to \infty$.

The substitution of $s_1 = s\gamma^{-1}$,

$$s_1\tilde{v} = \gamma^{-2}PL_2[s_1 - L_1 - (1 - P)L_2\gamma^{-1}]^{-1}L_2\tilde{v}(s) + v(0) \tag{6.4.97}$$

does not lead to a limit as $\gamma \to \infty$. However, it does yield an expansion which is similar to an ordinary perturbation problem, and, as we have seen, exhibits short-time behaviour.

a) Long-Time Perturbation Theory

We can expand (6.4.94) to order γ^{-2} in the form

$$s_2v = [A + B\gamma^{-1} + (C + Ds_2)\gamma^{-2}]\tilde{v} + v(0) \tag{6.4.98}$$

with

$$\begin{aligned} A &= -PL_2L_1^{-1}L_2 \\ B &= PL_2L_1^{-1}(1 - P)L_2L_1^{-1}L_2 \\ C &= -PL_2L_1^{-1}(1 - P)L_2L_1^{-1}(1 - P)L_2L_1^{-1}L_2 \\ D &= -PL_2L_1^{-2}L_2\,. \end{aligned} \tag{6.4.99}$$

Rearranging (6.4.98) we have

$$s_2(1 - \gamma^{-2}D)\tilde{v} = [A + B\gamma^{-1} + C\gamma^{-2}]\tilde{v} + v(0) \tag{6.4.100}$$

or, to order γ^{-2},

$$s_2\bar{v} = [A + B\gamma^{-1} + (C + DA)\gamma^{-2}]\bar{v} + (1 + \gamma^{-2}D)v(0)\,. \tag{6.4.101}$$

This gives the Laplace transform of an equation of motion for v, *in which the initial condition is not* $v(0)$, *but* $(1 + \gamma^{-2}D)v(0)$. Notice that this equation will be of first order in time, since s_2^n for $n > 1$ does not occur. This is not possible, of course, for an expansion to higher powers of γ.

b) Application to Brownian Motion

For Brownian motion we have already computed the operator A of (6.4.99); it is given by (6.4.65)

$$\boxed{A = -PL_2L_1^{-1}L_2 = \frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right].} \tag{6.4.102}$$

The other operators can be similarly calculated.

For example, from (6.4.63, 64),

$$L_2L_1^{-1}L_2 = -\left\{P_0(u)\frac{\partial}{\partial y} + \sqrt{2}\,P_2(u)\left[U'(y) + \frac{\partial}{\partial y}\right]\right\}\left\{U'(y) + \frac{\partial}{\partial y}\right\}.$$

Multiplying by $(1 - P)$ simply removes the term involving P_0, and multiplying by L_1^{-1} afterwards multiplies by $-\frac{1}{2}$. Hence,

$$L_1^{-1}(1 - P)L_2L_1^{-1}L_2 = \sqrt{2}\,P_2(u)\left[U'(y) + \frac{\partial}{\partial y}\right]^2.$$

We now use the Hermite polynomial recursion relations (6.4.60, 61) when multiplying by L_2: we derive

$$\begin{aligned} L_2P_2(u) &= \left[-\frac{\partial}{\partial y}u + U'(y)\frac{\partial}{\partial u}\right]P_2(u) \qquad (6.4.103)\\ &= -\sqrt{3}\,U'(y)P_3(u) - \frac{\partial}{\partial y}\left[\sqrt{3}\,P_3(u) + \sqrt{2}\,P_1(u)\right]. \end{aligned}$$

Finally, multiplying by P annihilates all terms since $P_0(u)$ does not occur. Hence,

$$\boxed{B = PL_2L_1^{-1}(1 - P)L_2L_1^{-1}L_2 = 0\,.} \tag{6.4.104}$$

The computation of C and DA follow similarly.

One finds that

$$C = \frac{\partial^2}{\partial y^2}\left[U'(y) + \frac{\partial}{\partial y}\right]\left[U'(y) + \frac{\partial}{\partial y}\right] \tag{6.4.105}$$

and

$$DA = -\frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right]\frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right] \tag{6.4.106}$$

so that

$$C + DA = \frac{\partial}{\partial y}U''(y)\left[U'(y) + \frac{\partial}{\partial y}\right] \tag{6.4.107}$$

and (6.4.101) is equivalent to the differential equation

$$\boxed{\frac{\partial \hat{p}}{\partial t} = \gamma^{-1}\frac{\partial}{\partial y}[1 + \gamma^{-2}U''(y)]\left[U'(y)\hat{p} + \frac{\partial \hat{p}}{\partial y}\right]} \tag{6.4.108}$$

with the initial condition

$$\lim_{t\to 0}\hat{p}(y, t) = \left\{1 - \gamma^{-2}\frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right]\right\}\hat{p}(y, 0)\,. \tag{6.4.109}$$

This alteration of the initial condition is a reflection of a "layer" phenomenon. Equation (6.4.108) is valid for $t \gg \gamma^{-1}$ and is known as the *corrected Smoluchowski equation.*

The exact solution would take account of the behaviour in the time up to $t \sim \gamma^{-1}$ in which terms like $\exp(-\gamma t)$ occur. Graphically, the situation is as in Fig. 6.1. the changed initial condition accounts for the effect of the initial layer near $t \sim \gamma^{-1}$.

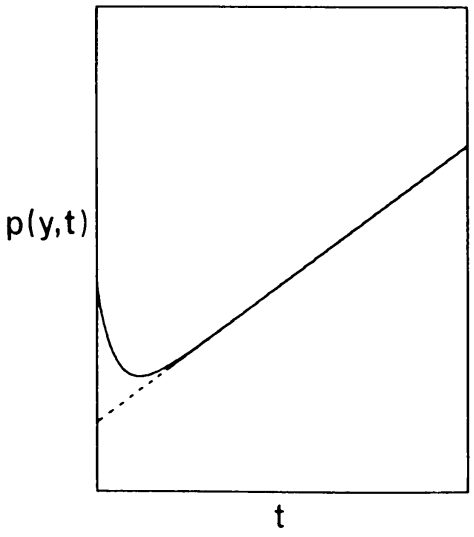


Fig. 6.1. Formation of a layer at a boundary. The exact solution (——) changes rapidly near the boundary on the left. The approximation (------) is good except near the boundary. The appropriate boundary condition for the approximation is thus given by the smaller value, where the dashed line meets the boundary

(c) Boundary Conditions

The higher order implementation of boundary conditions cannot be carried out by the methods of this section, since a rapidly varying layer occurs in the variable x near the boundary, and the assumption that the operator $\partial/\partial y$ is bounded becomes unjustified. Significant progress has been made by *Titulaer* and co-workers

[6.10–12]. Suppose the boundary is at $y = 0$, then we can substitute $z = \gamma y$, $\tau = \gamma t$ into the Kramers equation (6.4.19) to obtain

$$\frac{\partial P}{\partial \tau} = \left[\frac{\partial}{\partial u}\left(u + \frac{\partial}{\partial u}\right) - u\frac{\partial}{\partial z}\right] P + \frac{1}{\gamma} U'(z/\gamma)\frac{\partial P}{\partial u} \tag{6.4.110}$$

so that the zero order problem is the solution of that part of the equation independent of γ: to this order the potential is irrelevant. Only the stationary solution of (6.4.110) has so far been amenable to treatment. It can be shown that the stationary solution to the $\gamma \to \infty$ limit of (6.4.110) can be written in the form

$$P(u, z) = \psi_0'(u, z) + d_0^s \psi(u, z) + \sum_{n=1}^{\infty} d_n^s \psi_n(u, z) \tag{6.4.111}$$

where

$$\psi_0(u, z) = (2\pi)^{-1/2} \exp(-\tfrac{1}{2} u^2) \tag{6.4.112}$$

$$\psi_0'(u, z) = (2\pi)^{-1/2} (z - u) \exp(-\tfrac{1}{2} u^2) \tag{6.4.113}$$

and the $\psi_n(u, z)$ are certain complicated functions related to Hermite polynomials. The problem of determining the coefficients d_n^s is not straightforward, and the reader is referred to [6.12] for a treatment. It is found that the solution has an infinite derivative at $z = 0$, and for small z is of the form $a + bz^{1/2}$.

6.5 White Noise Process as a Limit of Nonwhite Process

The relationship between real noise and white noise has been mentioned previously in Sects.1.4.4, 4.1. We are interested in a limit of a differential equation

$$\frac{dx}{dt} = a(x) + b(x)\alpha_0(t)\,, \tag{6.5.1}$$

where $\alpha_0(t)$ is a stochastic source with some nonzero correlation time. We will show that if $\alpha_0(t)$ is a Markov process, then in the limit that it becomes a delta correlated process, the differential equation becomes a *Stratonovich stochastic differential equation* with the same coefficients, that is, it becomes

$$(S)\; dx = a(x)dt + b(x)dW(t) \tag{6.5.2}$$

which is equivalent to the Ito equation

$$dx = [a(x) + \tfrac{1}{2} b(x)b'(x)]dt + b(x)dW(t)\,. \tag{6.5.3}$$

To achieve the limit of a delta correlated Markov process, we must take the large γ limit of

$$\alpha_0(t) = \gamma\alpha(\gamma^2 t)\,, \tag{6.5.4}$$

where $\alpha(\tau)$ is a stationary stochastic process with

$$\langle \alpha(\tau) \rangle = 0 \tag{6.5.5}$$

$$\langle \alpha(\tau)\alpha(0) \rangle_s = g(\tau) . \tag{6.5.6}$$

Then,

$$\langle \alpha_0(t) \rangle = 0$$

$$\langle \alpha_0(t)\alpha_0(0) \rangle_s = \gamma^2 g(\gamma^2 t) . \tag{6.5.7}$$

In the limit of infinite γ, the correlation function becomes a delta function. More precisely, suppose

$$\int_{-\infty}^{\infty} g(\tau)d\tau = 1 \tag{6.5.8}$$

and

$$\int_{-\infty}^{\infty} |\tau| g(\tau)d\tau = \tau_c \tag{6.5.9}$$

defines the correlation time of the process $\alpha(t)$. [If $g(\tau)$ is exponential, then τ_c as defined in (6.5.9) requires that

$$g(\tau) \propto \exp(-\tau/\tau_c)$$

which agrees with the usage in Sects.1.4.4, 3.7.1.]

Then clearly, the correlation time of $\alpha_0(\tau)$ is τ_c/γ^2, which becomes zero as $\gamma \to \infty$; further,

$$\lim_{\gamma \to \infty} \langle \alpha_0(t)\alpha_0(0) \rangle_s = 0 \qquad (t \neq 0) . \tag{6.5.10}$$

But at all stages,

$$\int_{-\infty}^{\infty} \langle \alpha_0(t)\alpha_0(0) \rangle_s dt = \int_{-\infty}^{\infty} g(\tau)d\tau = 1 \tag{6.5.11}$$

so that we can write

$$\lim_{\gamma \to \infty} \langle \alpha_0(t)\alpha_0(0) \rangle_s = \delta(t) . \tag{6.5.12}$$

Therefore, the limit $\gamma \to \infty$ of $\alpha_0(t)$ does correspond to a normalised white noise limit. The higher-order correlation functions might be thought to be important too, but this turns out not to be the case.

We will give a demonstration in the case where $\alpha(\tau)$ is a Markov diffusion process whose Fokker-Planck equation is

$$\frac{\partial p(\alpha)}{\partial t} = -\frac{\partial}{\partial \alpha}[A(\alpha)p(\alpha)] + \frac{1}{2}\frac{\partial^2}{\partial \alpha^2}[B(\alpha)p(\alpha)] . \tag{6.5.13}$$

This means that the FPE for the pair (x, α) is

$$\frac{\partial p(x, \alpha)}{\partial t} = (\gamma^2 L_1 + \gamma L_2 + L_3)p(x, \alpha) \tag{6.5.14}$$

with

$$\begin{aligned} L_1 &= -\frac{\partial}{\partial \alpha} A(\alpha) + \frac{1}{2}\frac{\partial^2}{\partial \alpha^2} B(\alpha) \\ L_2 &= -\frac{\partial}{\partial x} b(x)\, \alpha \\ L_3 &= -\frac{\partial}{\partial x} a(x) . \end{aligned} \tag{6.5.15}$$

The asymptotic analysis now proceeds similarly to that used in Sects.6.4.1, 6.4.3, with a slight modification to take account of the operator L_3. Analogously to Sect. 6.4.1, we define a projector P on the space of functions of x and α by

$$(Pf)(x, \alpha) = p_s(\alpha) \int d\alpha f(x, \alpha) , \tag{6.5.16}$$

where $p_s(\alpha)$ is the solution of

$$L_1 p_s(\alpha) = 0 . \tag{6.5.17}$$

We assume that in the stationary distribution of α, the mean $\langle \alpha \rangle_s$ is zero. This means that the projection operator P satisfies the essential condition

$$PL_2P = 0 . \tag{6.5.18}$$

Since

$$\begin{aligned} (PL_2Pf)(x, \alpha) &= p_s(\alpha) \int d\alpha \left[-\frac{\partial}{\partial x} b(x)\alpha p_s(\alpha) \right] \int d\alpha' f(x, \alpha') \\ &= -p_s(\alpha)\langle \alpha \rangle_s \frac{\partial}{\partial x} b(x) \int d\alpha' f(x, \alpha') = 0 . \end{aligned} \tag{6.5.19}$$

Also, it is obvious that

$$PL_3 = L_3P \tag{6.5.20}$$

and, as before,

$$PL_1 = L_1P = 0 . \tag{6.5.21}$$

Defining, as before

$$v = Pp \tag{6.5.22}$$

$$w = (1 - P)p \tag{6.5.23}$$

and using the symbols $\tilde{v}$, $\tilde{w}$ for the corresponding Laplace transforms, we find

$$\begin{aligned} s\,\tilde{v}(s) &= P(\gamma^2 L_1 + \gamma L_2 + L_3)\tilde{p}(s) + v(0) \\ &= \gamma P L_2[P\tilde{p}(s) + (1 - P)\tilde{p}(s)] + L_3 P\,\tilde{p}(s) + v(0) \end{aligned} \tag{6.5.24}$$

so that

$$\boxed{s\,\tilde{v}(s) = \gamma P L_2 \tilde{w}(s) + L_3 \tilde{v}(s) + v(0)} \tag{6.5.25}$$

and similarly,

$$\boxed{s\,\tilde{w}(s) = [\gamma^2 L_1 + \gamma(1 - P)L_2 + L_3]\tilde{w}(s) + \gamma L_2 \tilde{v}(s) + w(0)} \tag{6.5.26}$$

which differ from (6.4.47) only by the existence of the $L_3\tilde{v}(s)$ term in (6.5.25) and $L_3\tilde{w}$ in (6.5.26). We again assume $w(0) = 0$, which means that $\alpha(t)$ is a *stationary* Markov process, so that

$$s\tilde{v}(s) = L_3\tilde{v}(s) - \gamma P L_2[-s + \gamma^2 L_1 + \gamma(1 - P)L_2 + L_3]^{-1}\gamma L_2 \tilde{v}(s) + v(0). \tag{6.5.27}$$

Now the limit $\gamma \to \infty$ gives

$$\boxed{s\tilde{v}(s) \simeq (L_3 - P L_2 L_1^{-1} L_2)\tilde{v}(s) + v(0)\,.} \tag{6.5.28}$$

We now compute $P L_2 L_1^{-1} L_2 \tilde{v}$. We write

$$\tilde{v}(s) = \tilde{p}(x) p_s(\alpha) \tag{6.5.29}$$

$$P L_2 L_1^{-1} L_2 \tilde{v} = p_s(\alpha) \int d\alpha' \left[-\frac{\partial}{\partial x} b(x)\alpha'\right] L_1^{-1} \left[-\frac{\partial}{\partial x} b(x)\alpha'\right] p_s(\alpha')\tilde{p}(x)\,. \tag{6.5.30}$$

We now need to evaluate

$$\int d\alpha\, \alpha L_1^{-1} \alpha p_s(\alpha) \equiv -D \tag{6.5.31}$$

and to do this we need a convenient expression for L_1^{-1}. Consider

$$\int_0^t \exp(L_1 t')dt' = L_1^{-1} \exp(L_1 t) - L_1^{-1} \tag{6.5.32}$$

and using (6.4.29)

$$\int_0^\infty \exp(L_1 t)\, dt = -L_1^{-1}(1 - P)\,. \tag{6.5.33}$$

Since, by assumption,

$$P\alpha p_s(\alpha) = p_s(\alpha) \int d\alpha' \alpha' p_s(\alpha') = p_s(\alpha')\langle\alpha\rangle_s = 0 \tag{6.5.34}$$

we have

$$D = \int d\alpha\, \alpha \int_0^\infty \exp(L_1 t)\alpha p_s(\alpha)\, dt \tag{6.5.35}$$

We note that

$$\exp(L_1 t)\alpha p_s(\alpha) \tag{6.5.36}$$

is the solution of the Fokker-Planck equation

$$\partial_t f = L_1 f \tag{6.5.37}$$

with initial condition

$$f(\alpha, 0) = \alpha p_s(\alpha)\,. \tag{6.5.38}$$

Hence,

$$\exp(L_1 t)\alpha p_s(\alpha) = \int d\alpha' p(\alpha, t\,|\,\alpha', 0)\alpha' p_s(\alpha') \tag{6.5.39}$$

and substituting in (6.5.35), we find

$$D = \int_0^\infty dt \int d\alpha\, d\alpha'\, \alpha\alpha' p(\alpha', t\,|\,\alpha, 0) p_s(\alpha)\,, \tag{6.5.40}$$

i.e.,

$$\boxed{D = \int_0^\infty dt\, \langle\alpha(t)\alpha(0)\rangle_s} \tag{6.5.41}$$

and from (6.5.8) and the symmetry of the correlation function,

$$D = 1/2\,. \tag{6.5.42}$$

Using (6.5.42) as the value of D, we find

$$-PL_2 L_1^{-1} L_2 \tilde{v} = \frac{1}{2} p_s(\alpha) \frac{\partial}{\partial x}\left[b(x) \frac{\partial}{\partial x} b(x)\hat{p}(x)\right] \tag{6.5.43}$$

so that the differential equation corresponding to (6.5.28) for

$$\hat{p}(x, t) = \int d\alpha\, p(x, \alpha) \tag{6.5.44}$$

is

$$\frac{\partial \hat{p}}{\partial t} = -\frac{\partial}{\partial x} a(x)\hat{p}(x) + \frac{1}{2}\frac{\partial}{\partial x} b(x)\frac{\partial}{\partial x} b(x)\hat{p}(x)\,. \tag{6.5.45}$$

This is, of course, the FPE in the Stratonovich form which corresponds to

$$(S)\; dx = a(x)dt + b(x)dW(t) \tag{6.5.46}$$

or which has the Ito form

$$dx = [a(x) + \tfrac{1}{2}b'(x)b(x)]dt + b(x)dW(t)\,, \tag{6.5.47}$$

as originally asserted.

6.5.1 Generality of the Result

A glance at the proof shows that all we needed was for $\alpha(t)$ to form a stationary Markov process with zero mean and with an evolution equation of the form

$$\frac{\partial p(\alpha)}{\partial t} = L_1 p(\alpha)\,, \tag{6.5.48}$$

where L_1 is a linear operator. This is possible for any kind of Markov process, in particular, for example, the random telegraph process in which $\alpha(t)$ takes on values $\pm a$. In the limit $\gamma \to \infty$, the result is still a Fokker-Planck equation. This is a reflection of the central limit theorem. For, the effective Gaussian white noise is made up of the sum of many individual components, as $\gamma \to \infty$, and the net result is still effectively Gaussian. In fact, *Papanicolaou* and *Kohler* [6.7] have rigorously shown that the result is valid even if $\alpha(t)$ is a non-Markov process, provided it is "strongly mixing" which, loosely speaking, means that all its correlation functions decay rapidly for large time differences.

6.5.2 More General Fluctuation Equations

Notice that in (6.5.1), instead of defining $\alpha_0(t)$ as simply $\gamma\alpha(t/\gamma^2)$, we can use the more general form

$$\alpha_0(t, x) = \gamma\psi[x, \alpha(t/\gamma^2)] \tag{6.5.49}$$

and now consider only $b(x) = 1$, since all x dependence can be included in ψ. We assume that

$$\int d\alpha\, \psi(x, \alpha)p_s(\alpha) = 0 \tag{6.5.50}$$

in analogy to the previous assumption $\langle\alpha\rangle_s = 0$.

Then D becomes x dependent, and we have to use

$$D(x) = \int_0^\infty dt \, \langle \psi[x, \alpha(t)] \psi[x, \alpha(0)] \rangle \tag{6.5.51}$$

and

$$E(x) = \int_0^\infty dt \, \left\langle \frac{\partial \psi}{\partial x}[x, \alpha(t)] \psi[x, \alpha(0)] \right\rangle \tag{6.5.52}$$

and the Fokker-Planck equation becomes

$$\frac{\partial \hat{p}}{\partial t} = -\frac{\partial}{\partial x}[\{a(x) + E(x)\}\hat{p}] + \frac{\partial^2}{\partial x^2}[D(x)\hat{p}] \, . \tag{6.5.53}$$

In this form we have agreement with the form derived by *Stratonovich* [Ref. 6.3, Eq.(4.4.39)].

6.5.3 Time Nonhomogeneous Systems

If instead of (6.5.1) we have

$$\frac{dx}{dt} = a(x, t) + b(x, t)\alpha_0(t) \, , \tag{6.5.54}$$

the Laplace transform method cannot be used simply. We can evade this difficulty by the following trick. Introduce the extra variable τ so that the equations become

$$dx = [a(x, \tau) + \gamma b(x, \tau)\alpha]dt \tag{6.5.55}$$

$$d\alpha = \gamma^2 A(\alpha)dt + \gamma\sqrt{B(\alpha)}\, dW(t) \tag{6.5.56}$$

$$d\tau = dt \, . \tag{6.5.57}$$

The final equation constrains t to be the same as τ, but the system now forms a homogeneous Markov process in the variables (x, α, τ). Indeed, any nonhomogeneous Markov process can be written as a homogeneous Markov process using this trick.
The Fokker-Planck equation is now

$$\frac{\partial p}{\partial t} = \gamma^2 L_1 + \gamma L_2 + L_3 \tag{6.5.58}$$

with

$$L_1 = -\frac{\partial}{\partial \alpha} A(\alpha) + \frac{1}{2}\frac{\partial^2}{\partial \alpha^2} B(\alpha) \tag{6.5.59}$$

$$L_2 = -\frac{\partial}{\partial x} b(x, \tau)\alpha \tag{6.5.60}$$

$$L_3 = -\frac{\partial}{\partial \tau} - \frac{\partial}{\partial x} a(x, \tau) . \tag{6.5.61}$$

Using the same procedure as before, we obtain

$$\frac{\partial \hat{p}}{\partial t} = \left[-\frac{\partial}{\partial \tau} - \frac{\partial}{\partial x} a(x, \tau) + \frac{1}{2} \frac{\partial}{\partial x} b(x, \tau) \frac{\partial}{\partial x} b(x, \tau) \right] \hat{p} \tag{6.5.62}$$

which yields

$$d\tau = dt \tag{6.5.63}$$

so that we have, after eliminating τ in terms of t,

$$\boxed{\frac{\partial \hat{p}}{\partial t} = \left[-\frac{\partial}{\partial x} a(x, t) + \frac{1}{2} \frac{\partial}{\partial x} b(x, t) \frac{\partial}{\partial x} b(x, t) \right] \hat{p}} \tag{6.5.64}$$

in exact analogy to (6.5.45).

6.5.4 Effect of Time Dependence in L_1

Suppose, in addition, that A and B depend on time as well, so that

$$L_1 = -\frac{\partial}{\partial \alpha} A(\alpha, \tau) + \frac{1}{2} \frac{\partial^2}{\partial \alpha^2} B(\alpha, \tau) . \tag{6.5.65}$$

In this case, we find P is a function of τ and hence does *not* commute with L_3. Thus,

$$PL_3 \neq L_3 P . \tag{6.5.66}$$

Nevertheless, we can take care of this. Defining $\tilde{v}(s)$ and $\tilde{w}(s)$ as before, we have

$$s\,\tilde{v}(s) = P(\gamma L_2 + L_3)\tilde{w}(s) + PL_3\tilde{v}(s) + v(0) \tag{6.5.67}$$

$$s\,\tilde{w}(s) = [\gamma^2 L_1 + \gamma(1 - P)L_2 + (1 - P)L_3]\tilde{w}(s) + \gamma L_2\tilde{v}(s) + (1 - P)L_3\tilde{v}(s) \tag{6.5.68}$$

so that

$$\begin{aligned} s\,\tilde{v}(s) = {} & PL_3\tilde{v}(s) + P(\gamma L_2 + L_3)[s - \gamma^2 L_1 - \gamma(1 - P)L_2 - (1 - P)L_3]^{-1} \\ & \times [\gamma L_2 + (1 - P)L_3]\tilde{v}(s) + v(0) . \end{aligned} \tag{6.5.69}$$

We see that because L_2 is multiplied by γ and L_3 is not, we get in the limit of large γ

$$s\,\tilde{v}(s) \sim (PL_3 - PL_3L_2^{-1}L_3)\tilde{v}(s) + v(0) \tag{6.5.70}$$

In this case we will not assume that we can normalise the autocorrelation function to a constant. The term $-PL_2L_1^{-1}L_2$ gives

$$\frac{\partial}{\partial x}b(x,\tau)\frac{\partial}{\partial x}b(x,\tau)\int_0^\infty dt\langle\alpha_\tau(t)\alpha_\tau(0)\rangle\,, \tag{6.5.71}$$

where by $\alpha_\tau(t)$ we mean the random variable whose FPE is

$$\frac{\partial p}{\partial t}=\left[\frac{\partial}{\partial\alpha}A(\alpha,\tau)+\frac{1}{2}\frac{\partial^2}{\partial\alpha^2}B(\alpha,\tau)\right]p\,. \tag{6.5.72}$$

Thus, the limit $\gamma\to\infty$ effectively makes the random motion of α infinitely faster than the motion due to the time dependence of α arising from the time dependence of A and B. Defining

$$D(\tau)=\int_0^\infty dt\langle\alpha_\tau(t)\alpha_\tau(0)\rangle\,, \tag{6.5.73}$$

we find, by eliminating τ as before,

$$\frac{\partial\hat{p}}{\partial t}=\left[-\frac{\partial}{\partial x}a(x,t)+D(t)\frac{\partial}{\partial x}b(x,t)\frac{\partial}{\partial x}b(x,t)\right]\hat{p}\,. \tag{6.5.74}$$

6.6 Adiabatic Elimination of Fast Variables: The General Case

We now want to consider the general case of two variables x and α which are coupled together in such a way that each affects the other. This is now a problem analogous to the derivation of the Smoluchowski equation with nonvanishing $V'(x)$, whereas the previous section was a generalization of the same equation with $V'(x)$ set equal to zero.

The most general problem of this kind would be so complex and unwieldy as to be incomprehensible. In order to introduce the concepts involved, we will first consider an example of a linear chemical system and then develop a generalised theory.

6.6.1 Example: Elimination of Short-Lived Chemical Intermediates

We consider the example of a chemically reacting system

$$X\underset{k}{\overset{\gamma}{\rightleftharpoons}}Y\underset{\gamma}{\overset{k}{\rightleftharpoons}}A \tag{6.6.1}$$

where X and Y are chemical species whose quantities vary, but A is by some means held fixed. The deterministic rate equations for this system are

$$\frac{dx}{dt}=-x+\gamma y \tag{6.6.2a}$$

$$\frac{dy}{dt} = -2\gamma y + x + a\,. \tag{6.6.2b}$$

Here x, y, a, are the concentrations of X, Y, A. The rate constants have been chosen so that $k = 1$, for simplicity.

The physical situation is often that Y is a very short-lived intermediate state which can decay to X or A, with time constant γ^{-1}. Thus, the limit of large γ in which the short-lived intermediate Y becomes even more short lived, and its concentration negligible, is of interest. This results in the situation where we solve (6.6.2b) with $dy/dt = 0$ so that

$$y = (x + a)/2\gamma\,, \tag{6.6.3}$$

and substitute this in (6.6.2a) to get

$$\frac{dx}{dt} = \frac{x}{2} - \frac{a}{2}\,. \tag{6.6.4}$$

The stochastic analogue of this procedure is complicated by the fact that the white noises to be added to (6.6.2) are correlated, and the stationary distribution of y depends on γ. More precisely, the stochastic differential equations corresponding to (6.3.2) are usually chosen to be (Sect.7.6.1)

$$\begin{aligned} dx &= (-x + \gamma y)dt + \varepsilon B_{11}dW_1(t) + \varepsilon B_{12}dW_2(t) \\ dy &= (-2\gamma y + x + a)dt + \varepsilon B_{21}dW_1(t) + \varepsilon B_{22}dW_2(t), \end{aligned} \tag{6.6.5}$$

where the matrix B satisfies

$$BB^{\mathrm{T}} = \begin{bmatrix} 2a & -2a \\ -2a & 4a \end{bmatrix}. \tag{6.6.6}$$

Here ε is a parameter, which is essentially the square root of the inverse volume of the reacting system and is usually small, though we shall not make use of this fact in what follows.

We wish to eliminate the variable y, whose mean value would be given by (6.6.3) and becomes vanishingly small in the limit. It is only possible to apply the ideas we have been developing if the variable being eliminated has a distribution function in the stationary state which is independent of γ. We will thus have to define a new variable as a function of y and x which possesses this desirable property.

The Fokker-Planck equation corresponding to (6.6.5) is

$$\begin{aligned} \frac{\partial p}{\partial t} = \Big[&\frac{\partial}{\partial x}(x - \gamma y) + \frac{\partial^2}{\partial x^2}\varepsilon^2 a - 2\frac{\partial^2}{\partial x \partial y}\varepsilon^2 a \\ &+ \frac{\partial}{\partial y}(2\gamma y - x - a) + 2\varepsilon^2 a \frac{\partial^2}{\partial y^2}\Big] p\,. \end{aligned} \tag{6.6.7}$$

It seems reasonable to define a new variable z by

$$z = 2\gamma y - x - a \tag{6.6.8}$$

which is proportional to the difference between y and its stationary value. Thus, we formally define a pair of new variables (x_1, z) by

$$\begin{aligned} x_1 &= x & x &= x_1 \\ z &= 2\gamma y - x - a & y &= (z + x_1 + a)/2\gamma \end{aligned} \tag{6.6.9}$$

so that we can transform the FPE using

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial}{\partial x_1} - \frac{\partial}{\partial z} \\ \frac{\partial}{\partial y} &= 2\gamma \frac{\partial}{\partial z} \end{aligned} \tag{6.6.10}$$

to obtain

$$\begin{aligned} \frac{\partial p}{\partial t} = \left[\frac{\partial}{\partial x_1}\left[\frac{x_1 - a}{2} - \frac{z}{2}\right] + \varepsilon^2 a \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_1 \partial z}(-2\varepsilon^2 a - 4\gamma\varepsilon^2 a)\right. \\ \left. + \frac{\partial}{\partial z}\left(2\gamma z - \frac{x - a}{2} + \frac{z}{z}\right) + \frac{\partial^2}{\partial z^2}(8\varepsilon^2\gamma^2 a + \varepsilon^2 a + 4\gamma\varepsilon^2 a)\right] p\,. \end{aligned} \tag{6.6.11}$$

The limit of $\gamma \to \infty$ does not yet give a Fokker-Planck operator in z, which is simply proportional to a fixed operator: we see that the drift and diffusion terms for z are proportional to γ and γ^2, respectively.

However, the substitution

$$\alpha = z\gamma^{-1/2} \tag{6.6.12}$$

changes this. In terms of α, the drift and diffusion coefficients become proportional to γ and we can see (now writing x instead of x_1)

$$\frac{\partial p}{\partial t} = [\gamma L_1 + \gamma^{1/2} L_2(\gamma) + L_3] p \tag{6.6.13}$$

in which

$$L_1 = 2\frac{\partial}{\partial \alpha}\alpha + 8\varepsilon^2 a \frac{\partial^2}{\partial \alpha^2} \tag{6.6.14}$$

$$\begin{aligned} L_2(\gamma) = \frac{\partial}{\partial x}\left[-4\varepsilon^2 a \frac{\partial}{\partial \alpha} - \tfrac{1}{2}\alpha\right] + 4\gamma^{-1/2}\varepsilon^2 a \frac{\partial^2}{\partial \alpha^2} - \gamma^{-1}\frac{\partial}{\partial \alpha}\left[\frac{x - a}{2}\right] \\ + \left\{-\gamma^{-1/2}\frac{\partial}{\partial \alpha}\alpha - \gamma^{-1} 2\varepsilon^2 a \frac{\partial^2}{\partial x \partial \alpha} + \gamma^{-3/2}\varepsilon^2 a \frac{\partial^2}{\partial \alpha^2}\right\} \end{aligned} \tag{6.6.15}$$

$$L_3 = \frac{\partial}{\partial x}\left(\frac{x - a}{2}\right) + \varepsilon^2 a \frac{\partial^2}{\partial x^2}\,. \tag{6.6.16}$$

Notice that $L_2(\gamma)$ has a large γ limit given by the first term of the first line of (6.6.15). The only important property of L_2 is that $PL_2P = 0$. Defining P as usual

$$Pf(x, \alpha) = p_s(\alpha) \int d\alpha' f(x, \alpha') \tag{6.6.17}$$

where $p_s(\alpha)$ is the stationary solution for L_1, we see that for any operator beginning with $\partial/\partial\alpha$ [such as the γ dependent part of $L_2(\gamma)$] we have

$$P\frac{\partial}{\partial\alpha}f = p_s(\alpha) \int d\alpha' \frac{\partial}{\partial\alpha'}f(x, \alpha') = 0\,, \tag{6.6.18}$$

provided we can drop boundary terms. Hence, the γ dependent part of $L_2(\gamma)$ satisfies $PL_2P = 0$. Further, it is clear from (6.6.14) that $\langle\alpha\rangle_s = 0$, so we find that the first, γ independent part, also satisfies this condition. Thus

$$PL_2(\gamma)P = 0\,. \tag{6.6.19}$$

Nevertheless, it is worth commenting that the γ dependent part of $L_2(\gamma)$ contains terms which look more appropriate to L_1, that is, terms not involving any x derivatives. However, by moving these terms into L_2, we arrange for L_1 to be independent of γ. Thus, P is independent of γ and the limits are clearer.

The procedure is now quite straightforward. Defining, as usual

$$\begin{aligned} P\tilde{p}(s) &= \tilde{v}(s) \\ (1 - P)\tilde{p}(s) &= \tilde{w}(s) \end{aligned} \tag{6.6.20}$$

and assuming, as usual, $w(0) = 0$, we find

$$s\,\tilde{v}(s) = P[\gamma L_1 + \gamma^{1/2}L_2(\gamma) + L_3]\,[\tilde{v}(s) + \tilde{w}(s)] + v(0) \tag{6.6.21}$$

and using

$$\begin{aligned} PL_1 &= L_1P = 0 \\ PL_2P &= 0 \\ PL_3 &= L_3P\,, \end{aligned} \tag{6.6.22}$$

we obtain

$$s\,\tilde{v}(s) = P\gamma^{1/2}L_2(\gamma)\tilde{w}(s) + L_3\tilde{v}(s) + v(0) \tag{6.6.23}$$

and similarly,

$$s\,\tilde{w}(s) = [\gamma L_1 + \gamma^{1/2}(1 - P)L_2(\gamma) + L_3]\tilde{w}(s) + \gamma^{1/2}L_2(\gamma)\tilde{v}(s) \tag{6.6.24}$$

so that

$$s\,\tilde{v}(s) = \{L_3 + \gamma P L_2(\gamma)[s - \gamma L_1 - \gamma^{1/2}(1 - P)L_2(\gamma) - L_3]^{-1}L_2\}\,\tilde{v}(s) + v(0)\,. \tag{6.6.25}$$

Now taking the large γ limit, we get

$$\boxed{s\tilde{v}(s) \simeq (L_3 - PL_2L_1^{-1}L_2)\tilde{v}(s) + v(0),} \tag{6.6.26}$$

where

$$L_2 = \lim_{\gamma\to\infty} L_2(\gamma) = \frac{\partial}{\partial x}\left(-4\varepsilon^2 a\frac{\partial}{\partial\alpha} - \tfrac{1}{2}\alpha\right). \tag{6.6.27}$$

Equation (6.6.26) is of exactly the same form as (6.5.28) and indeed the formal derivation from (6.6.7) is almost identical. The evaluation of $PL_2L_1^{-1}L_2$ is, however, slightly different because of the existence of terms involving $\partial/\partial\alpha$. Firstly, notice that since $P\,\partial/\partial\alpha = 0$, we can write

$$-PL_2L_1^{-1}L_2v = -p_s(\alpha)\int d\alpha'\left(-\tfrac{1}{2}\alpha'\frac{\partial}{\partial x}\right)L_1^{-1}\frac{\partial}{\partial x}\left(-4\varepsilon^2 a\frac{\partial}{\partial\alpha'} - \tfrac{1}{2}\alpha'\right)p_s(\alpha')\hat{p}(x) \tag{6.6.28}$$

and from the definition of $p_s(\alpha)$ as satisfying $L_1p_s(\alpha) = 0$, we see from (6.6.14) that

$$\frac{\partial}{\partial\alpha}p_s(\alpha) = -\alpha p_s(\alpha)/4\varepsilon^2 a \tag{6.6.29}$$

and hence that

$$-PL_2L_1^{-1}L_2v = \frac{1}{4}p_s(\alpha)\frac{\partial^2}{\partial x^2}\int d\alpha'\alpha' L_1^{-1}\alpha' p_s(\alpha')\hat{p}(x) \tag{6.6.30}$$

$$= -\frac{1}{4}p_s(\alpha)\frac{\partial^2\hat{p}(x)}{\partial x^2}\int_0^\infty dt\,\langle\alpha(t)\alpha(0)\rangle_s\,, \tag{6.6.31}$$

where we have used the reasoning given in Sect.6.5 to write the answer in terms of the correlation function. Here, L_1 is the generator of an Ornstein-Uhlenbeck process (Sect 3.8.4) with $k = 2$, $D = 16\varepsilon^2 a$, so that from (3.8.2),

$$-PL_2L_1^{-1}L_2v = -\frac{1}{4}p_s(\alpha)\frac{\partial^2\hat{p}(x)}{\partial x^2}\,4\varepsilon^2 a\int_0^\infty dt\,\mathrm{e}^{-2t} \tag{6.6.32}$$

$$= -\frac{1}{2}\varepsilon^2 a\frac{\partial^2\hat{p}(x)}{\partial x^2}\,p_s(\alpha)\,.$$

Hence, from (6.6.26), the effective Fokker-Planck equation is

$$\boxed{\frac{\partial\hat{p}}{\partial t} = \frac{\partial}{\partial x}\frac{x-a}{2}\hat{p}(x) + \frac{1}{2}\varepsilon^2 a\frac{\partial^2\hat{p}}{\partial x^2}\,.} \tag{6.6.33}$$

Comments

i) This is exactly the equation expected from the reaction

$$X \underset{k/2}{\overset{k/2}{\rightleftharpoons}} A \tag{6.6.34}$$

(with $k = 1$) (Sect.7.5.3). It is expected because general principles tell us that the stationary variance of the concentration fluctuations is given by

$$\mathrm{var}\{x(t)\}_s = \varepsilon^2 \langle x \rangle_s . \tag{6.6.35}$$

ii) Notice that the net effect of the adiabatic elimination is to *reduce* the coefficient of $\partial^2/\partial x^2$, a result of the correlation between the noise terms in x and y in the original equations.

iii) This result differs from the usual adiabatic elimination in that the noise term in the eliminated variable is important. There are cases where this is not so; they will be treated shortly.

6.6.2 Adiabatic Elimination in Haken's Model

Haken has introduced a simple model for demonstrating adiabatic elimination [Ref. 6.1, Sect.7.2]. The deterministic version of the model is a pair of coupled equations which may be written

$$\dot{x} = -\varepsilon x - a x \alpha \tag{6.6.36}$$

$$\dot{\alpha} = -\kappa \alpha + b x^2 . \tag{6.6.37}$$

One assumes that if κ is sufficiently large, we may, as before, replace α by the stationary solution (6.6.37) in terms of x to obtain

$$\alpha = \frac{b}{\kappa} x^2 \tag{6.6.38}$$

$$\dot{x} = -\varepsilon x - \frac{ab}{\kappa} x^3 . \tag{6.6.39}$$

The essential aim of the model is to obtain the cubic form on the right-hand side of (6.6.39).

In making the transition to a stochastic system, we find that there are various possibilities available. The usual condition for the validity of adiabatic elimination is

$$\varepsilon \ll \kappa . \tag{6.6.40}$$

In a stochastic version, all other parameters come into play as well, and the condition (6.6.40) is, in fact, able to be realised in at least three distinct ways with characteristically different answers.

Let us write stochastic versions of (6.6.36, 37):

$$\begin{aligned} dx &= -(\varepsilon x + axa)dt + C\,dW_1(t) \\ da &= (-\kappa a + bx^2)dt + D\,dW_2(t) \end{aligned} \tag{6.6.41}$$

and we assume here, for simplicity, that C and D are constants and $W_1(t)$ and $W_2(t)$ are independent of each other.

The Fokker-Planck equation is

$$\frac{\partial p}{\partial t} = \left[\frac{\partial}{\partial x}(\varepsilon x + axa) + \tfrac{1}{2}C^2\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial a}(\kappa a - bx^2) + \tfrac{1}{2}D^2\frac{\partial^2}{\partial a^2}\right]p\,. \tag{6.6.42}$$

We wish to eliminate a. It is convenient to define a new variable β by

$$\beta = a - \frac{b}{\kappa}x^2 \tag{6.6.43}$$

so that, for fixed x, the quantity β has zero mean. In terms of this variable, we can write a FPE:

$$\frac{\partial p}{\partial t} = (L_1^0 + L_2^0 + L_3^0)p \tag{6.6.44}$$

$$L_1^0 = \frac{\partial}{\partial \beta}\kappa\beta + \frac{D^2}{2}\frac{\partial^2}{\partial \beta^2} \tag{6.6.45}$$

$$\begin{aligned} L_2^0 = {} & \frac{\partial}{\partial \beta}a\beta x - \frac{2bx}{\kappa}\frac{\partial}{\partial \beta}\left(\varepsilon x + \frac{ab}{\kappa}x^3 + ax\beta\right) \\ & - C^2\left(\frac{bx}{\kappa}\frac{\partial^2}{\partial x \partial \beta} + \frac{\partial^2}{\partial x \partial \beta}\frac{bx}{\kappa}\right) + \frac{2b^2x^2C^2}{\kappa^2}\frac{\partial^2}{\partial \beta^2} \end{aligned} \tag{6.6.46}$$

$$L_3^0 = \left[\frac{\partial}{\partial x}\left(\varepsilon x + \frac{ab}{\kappa}x^3\right) + \frac{C^2}{2}\frac{\partial^2}{\partial x^2}\right] \tag{6.6.47}$$

In terms of these variables, the limit $\varepsilon \to 0$ is not interesting since we simply get the same system with $\varepsilon = 0$. No elimination is possible since L_1 is not multiplied by a large parameter.

In order for the limit $\varepsilon \to 0$ to have the meaning deterministically that (6.6.39) is a valid *limiting form*, there must exist an A such that

$$\frac{ab}{\kappa} = \varepsilon A, \qquad \text{as} \quad \varepsilon \to 0\,. \tag{6.6.48}$$

For this limit to be recognisable deterministically, it must not be swamped by noise so one must also have

$$\frac{C^2}{2} = \varepsilon B \qquad \text{as} \quad \varepsilon \to 0 \tag{6.6.49}$$

which means, as $\varepsilon \to 0$

$$L_3^0 \to \varepsilon \left[\frac{\partial}{\partial x}(x + Ax^3) + B \frac{\partial^2}{\partial x^2} \right]. \tag{6.6.50}$$

However, there are two distinct possibilities for L_2^0. In order for L_1^0 to be independent of ε, we must have κ independent of ε, which is reasonable. Thus, the limit (6.6.48) must be achieved by the product ab being proportional to ε. We consider various possibilities.

a) The Silent Slave: a Proportional to ε

We assume we can write

$$a = \varepsilon \tilde{a} \,. \tag{6.6.51}$$

We see that L_1^0 is independent of ε while L_2^0 and L_3^0 are proportional to ε. If we rescale time by

$$\tau = \varepsilon t \tag{6.6.52}$$

then

$$\frac{\partial p}{\partial \tau} = \left(\frac{1}{\varepsilon} L_1 + L_2 + L_3 \right) p \,, \tag{6.6.53}$$

where

$$\begin{aligned} L_1 &= L_1^0 \\ L_2 &= L_2^0/\varepsilon \\ L_3 &= L_3^0/\varepsilon \,. \end{aligned} \tag{6.6.54}$$

Clearly, the usual elimination procedure gives to lowest order

$$\boxed{\frac{\partial \hat{p}}{\partial t} = L_3 \hat{p} = \left[\frac{\partial}{\partial x}(x + Ax^3) + B \frac{\partial^2}{\partial x^2} \right] \hat{p}} \tag{6.6.55}$$

since L_2 does not become infinite as $\varepsilon \to 0$.

This corresponds exactly to eliminating α adiabatically, ignoring the fluctuations in α and simply setting the deterministic value in the x equation. I call it the 'silent slave', since (in Haken's terminology) α is slaved by x and makes no contribution

to the noise in the x equation. This is the usual form of slaving, as considered by Haken.

b) The Noisy Slave: a Proportional to $\varepsilon^{1/2}$

If we alternatively assume that both a and b are proportional to $\varepsilon^{1/2}$, we can write

$$a = \tilde{a}\varepsilon^{1/2} \qquad (6.6.56)$$
$$b = \tilde{b}\varepsilon^{1/2},$$

where

$$\tilde{a}\tilde{b} = \kappa A . \qquad (6.6.57)$$

L_1^0 stays constant, L_3^0 is proportional to ε and

$$L_2^0 = \varepsilon^{1/2}L_2 + \text{higher order terms in } \varepsilon , \qquad (6.6.58)$$

where

$$L_2 = \tilde{a}\beta\frac{\partial}{\partial x}x . \qquad (6.6.59)$$

Thus, the limiting equation is

$$\frac{\partial \hat{p}}{\partial \tau} = (L_3 - PL_2L_1^{-1}L_2)\hat{p} . \qquad (6.6.60)$$

The term $PL_2L_1^{-1}L_2$ can be worked out as previously; we find

$$-PL_2L_1^{-1}L_2 = \tilde{a}^2\frac{D^2}{2\kappa^2}\frac{\partial}{\partial x}x\frac{\partial}{\partial x}x \qquad (6.6.61)$$

so

$$\boxed{\frac{\partial \hat{p}}{\partial \tau} = \left\{\frac{\partial}{\partial x}\left[x\left[1 - \frac{\tilde{a}^2D^2}{2\kappa^2}\right] + Ax^3\right] + \frac{\partial^2}{\partial x^2}\left[B + \frac{\tilde{a}^2D^2x^2}{2\kappa^2}\right]\right\}\hat{p} .} \qquad (6.6.62)$$

I call this the "noisy slave", since the slave makes his presence felt in the final equation by adding noise (and affecting the drift, though this appears only in the Ito form as written; as a Stratonovich form, there would be no extra drift).

c) The General Case

Because we assume $ab \propto \varepsilon$, it can be seen that the second two terms in (6.6.46) are always proportional to ε^p, where $p > 1$, and hence are negligible (provided b is bounded). Thus, the only term of significance in L_2^0 is the first. Then it follows that if

$$a = \varepsilon^r\tilde{a} ,$$

we have the following possibilities:

$r > \frac{1}{2}$: no effect from L_2: limiting equation is (6.6.55),
$r = \frac{1}{2}$: limiting equation is (6.6.62)—a noisy slave,
$r < \frac{1}{2}$: the term $PL_2L_1^{-1}L_2$ becomes of order $\varepsilon^{2r-1} \to \infty$ and is dominant.

The equation is asymptotically (for $r < \frac{1}{2}$)

$$\frac{\partial \hat{p}}{\partial \tau} = \varepsilon^{2r-1}\tilde{a}^2 \frac{D^2}{2\kappa^2}\left(\frac{\partial}{\partial x} x \frac{\partial}{\partial x} x\right)\hat{p} . \tag{6.6.63}$$

These are quite distinct differences, all of which can be incorporated in the one formula, namely, in general

$$\boxed{\frac{\partial \hat{p}}{\partial \tau} = \left[\frac{\partial}{\partial x}(x + Ax^3) + \frac{\partial^2}{\partial x^2} B + \varepsilon^{2r-1}\tilde{a}^2 \frac{D^2}{2\kappa^2}\frac{\partial}{\partial k} x \frac{\partial}{\partial x} x\right]\hat{p} .} \tag{6.6.64}$$

In applying adiabatic elimination techniques, in general, one simply must take particular care to ensure that the correct dependence on small parameters of all constants in the system has been taken.

6.6.3 Adiabatic Elimination of Fast Variables: A Nonlinear Case

We want to consider the general case of two variables x and α which are coupled together in such a way that each affects the other, though the time scale of α is considerably faster than that of x.

Let us consider a system described by a pair of stochastic differential equations:

$$dx = [a(x) + b(x)\alpha]dt + c(x)dW_1(t) \tag{6.6.65}$$

$$d\alpha = \gamma^2[A(\alpha) - f(x)]dt + \gamma\sqrt{2B(\alpha)}\, dW_2(t) . \tag{6.6.66}$$

If we naively follow the reasoning of Sect.6.4, we immediately meet trouble. For in this limit, one would put

$$A(\alpha) - f(x) = -\sqrt{\frac{2B(\alpha)}{\gamma^2}}\frac{dW_2(t)}{dt} \tag{6.6.67}$$

on the assumption that for large α, (6.6.66) is always such that $d\alpha/dt = 0$. But then to solve (6.6.67) for α in terms x yields, in general, some complicated nonlinear function of x and $dW_2(t)/dt$ whose behaviour is inscrutable. However, if $B(\alpha)$ is zero, then we can define $u_0(x)$ to be

$$A[u_0(x)] = f(x) \tag{6.6.68}$$

and substitute in (6.6.65) to obtain

$$dx = [a(x) + b(x)u_0(x)]dt + c(x)dW_1(t)\,. \tag{6.6.69}$$

We shall devise a somewhat better procedure based on our previous methodology which can also take the effect of fluctuations in (6.6.66) into account.

The Fokker-Planck equation equivalent to (6.6.65, 66) is

$$\frac{\partial p}{\partial t} = (\gamma^2 L_1 + L_2 + L_3)p\,, \tag{6.6.70}$$

where

$$L_1 = \frac{\partial}{\partial \alpha}[f(x) - A(\alpha)] + \frac{\partial^2}{\partial \alpha^2} B(\alpha) \tag{6.6.71}$$

and L_2 and L_3 are chosen with hindsight in order for the requirement $PL_2P = 0$ to be satisfied. Firstly, we choose P, as usual, to be the projector into the null space of L_1. We write $p_x(\alpha)$ for the stationary solution i.e., the solution of

$$L_1 p_x(\alpha) = 0 \tag{6.6.72}$$

so that $p_x(\alpha)$ explicitly depends on x, because L_1 explicitly depends on x through the function $f(x)$. The projector P is defined by

$$(PF)\,(x, \alpha) = p_x(\alpha) \int d\alpha'\, F(x, \alpha') \tag{6.6.73}$$

for any function $F(x,\alpha)$.

We now define the function $u(x)$ as

$$u(x) = \int d\alpha\, \alpha p_x(\alpha) \equiv \langle \alpha \rangle_x\,. \tag{6.6.74}$$

Then we define

$$L_2 = -\frac{\partial}{\partial x}\left\{b(x)[\alpha - u(x)]\right\} \tag{6.6.75}$$

$$L_3 = -\frac{\partial}{\partial x}[a(x) + b(x)u(x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)]^2 \tag{6.6.76}$$

so that the term $\{\partial/\partial x\}\, b(x)u(x)$ cancels when these are added. Thus (6.6.70) is the correct FPE corresponding to (6.6.65, 66).

Now we have

$$PL_2PF = -p_x(\alpha) \int d\alpha' \frac{\partial}{\partial x}\{b(x)[\alpha' - u(x)]\} p_x(\alpha') \int d\alpha''\, F(x, \alpha'') \tag{6.6.77}$$

$$= 0$$

since $\int \alpha p_x(\alpha) d\alpha = u(x)$.

It is of course true that

$$PL_1 = L_1P = 0 \tag{6.6.78}$$

but

$$PL_3 \neq L_3P\,. \tag{6.6.79}$$

We now carry out the normal procedure. Writing, as usual,

$$\begin{aligned} P\,\tilde{p}(s) &= \tilde{v}(s) \\ (1 - P)\tilde{p}(s) &= \tilde{w}(s) \end{aligned} \tag{6.6.80}$$

and assuming $w(0) = 0$, we find

$$\begin{aligned} s\,\tilde{v}(s) &= P(L_2 + L_3)\tilde{w}(s) + PL_3\tilde{v}(s) + v(0) \\ s\,\tilde{w}(s) &= [\gamma^2 L_1 + (1 - P)L_2 + (1 - P)L_3]\tilde{w}(s) + L_2\tilde{v}(s) + (1 - P)L_3\tilde{v}(s) \end{aligned} \tag{6.6.81}$$

so that

$$\begin{aligned} s\,\tilde{v}(s) = {} & PL_3\tilde{v}(s) + P(L_2 + L_3)[s - \gamma^2 L_1 - (1 - P)L_2 - (1 - P)L_3]^{-1} \\ & \times [L_2 + (1 - P)L_3]\tilde{v}(s) + v(0)\,. \end{aligned} \tag{6.6.82}$$

To second order we have simply

$$\boxed{s\,\tilde{v}(s) \simeq \{(PL_3 - \gamma^{-2}P(L_2 + L_3)L_1^{-1}[L_2 + (1 - P)L_3]\}\,\tilde{v}(s) + v(0)\,.} \tag{6.6.83}$$

The term $PL_3\tilde{v}(s)$ is the most important term and yields the deterministic adiabatic elimination result. Writing

$$v(t) = p_\alpha(\alpha)\hat{p}(x)\,,$$

we find

$$PL_3v(t) = p_\alpha(\alpha)\int d\alpha' \left\{-\frac{\partial}{\partial x}[a(x) + b(x)u(x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)^2]\right\} p_\alpha(\alpha)\hat{p}(x) \tag{6.6.84}$$

and since

$$\int d\alpha\, p_\alpha(\alpha) = 1\,,$$

$$PL_3v(t) = p_\alpha(\alpha)\left\{-\frac{\partial}{\partial x}[a(x) + b(x)u(x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)^2]\right\}\hat{p}(x) \tag{6.6.85}$$

so the lowest-order differential equation is

$$\frac{\partial\hat{p}(x)}{\partial t} = -\frac{\partial}{\partial x}[a(x) + b(x)u(x)]\,\hat{p}(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\,c(x)^2\hat{p}(x) \tag{6.6.86}$$

which is equivalent to the stochastic differential equation

$$\boxed{dx = [a(x) + b(x)u(x)]dt + c(x)dW(t)\,.} \tag{6.6.87}$$

To this order, the equation of motion contains no fluctuating term whose origin is in the equation for α. However, the effect of completely neglecting fluctuations is given by (6.6.69) which is very similar to (6.6.87) but has $u_0(x)$ instead of $u(x)$. While it is expected that the average value of α in the stationary state would be similar to $u_0(x)$, it is not the same, and the similarity would only be close when the noise term $B(\alpha)$ was small.

Second-Order Corrections

It is possible to evaluate (6.6.83) to order γ^{-2}. At first glance, the occurrence of 2nd derivatives in L_3 would seem to indicate that to this order, 4th derivatives occur since L_3 occurs twice. However, we can show that the fourth-order terms vanish.

Consider the expression

$$P(L_2 + L_3)L_1^{-1}[L_2 + (1 - P)L_3]\tilde{v}(s)\,. \tag{6.6.88}$$

We know

$$\begin{aligned} &\text{i)}\ \ P\tilde{v}(s) = \tilde{v}(s) \\ &\text{ii)}\ (1 - P)L_2P\tilde{v}(s) = L_2P\tilde{v}(s)\,, \end{aligned} \tag{6.6.89}$$

where we have used $PL_2P = 0$.

Thus, (6.6.88) becomes

$$\begin{aligned} P(L_2 + L_3)L_1^{-1}(1 - P)(L_2 + L_3)P\tilde{v}(s) &= P\{PL_2 + [P, L_3] + L_3P\} \\ &\times (1 - P)L_1^{-1}(1 - P)\{L_2P + [L_3, P] + PL_3\}\tilde{v}(s) \end{aligned} \tag{6.6.90}$$

where the *commutator* $[A, B]$ is defined by

$$[A, B] = AB - BA\,. \tag{6.6.91}$$

We have noted that L_1^{-1} commutes with $(1 - P)$ and used $(1 - P)^2 = (1 - P)$ in (6.6.90) to insert another $(1 - P)$ before L_1^{-1}. We have also inserted another P in front of the whole expression, since $P^2 = P$. Using now

$$P(1 - P) = (1 - P)P = 0,$$

(6.6.90) becomes $P\{PL_2 + [P, L_3]\}L_1^{-1}(1 - P)\{L_2 + [L_3, P]\}\tilde{v}(s)\,.$ (6.6.92)

We will now compute $[P, L_3]$:

$$\begin{aligned} (PL_3f)(x, \alpha) = p_x(\alpha)\Big\{&-\frac{\partial}{\partial x}[a(x) + b(x)u(x)] \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)]^2\Big\}\int f(x, \alpha')\,d\alpha' \end{aligned} \tag{6.6.93}$$

and

$$(L_3P)f(x,\alpha) = \left\{-\frac{\partial}{\partial x}]a(x) + b(x)u(x)]\right. \\ \left. + \frac{1}{2}\frac{\partial^2}{\partial x^2}[a(x)]^2\right\} p_x(\alpha)\int d\alpha' f(x,\alpha') . \tag{6.6.94}$$

Subtracting these, and defining,

$$r_x(\alpha) = \frac{\partial p_x(\alpha)}{\partial x}\Big/ p_x(\alpha) \tag{6.6.95}$$

$$s_x(\alpha) = \frac{\partial^2 p_x(\alpha)}{\partial x^2}\Big/ p_x(\alpha) . \tag{6.6.96}$$

One finds

$$([P, L_3]f)(x,\alpha) = r_x(\alpha)[a(x) + b(x)u(x)]Pf(x,\alpha) \\ - \tfrac{1}{2}s_x(\alpha)c(x)^2Pf(x,\alpha) - r_x(\alpha)P\frac{\partial}{\partial x}[c(x)^2 f(x,\alpha)] . \tag{6.6.97}$$

The last term can be simplified even further since we are only interested in the case where $f(x,\alpha)$ is v, i.e.,

$$f(x,\alpha) = p_x(\alpha)\hat{p}(x) . \tag{6.6.98}$$

Then,

$$P\frac{\partial}{\partial x}c(x)^2 p_x(\alpha)\hat{p}(x) \tag{6.6.99}$$

$$= p_x(\alpha)\frac{\partial}{\partial x}c(x)^2\int d\alpha' p_x(\alpha')\hat{p}(x) \tag{6.6.100}$$

$$= p_x(\alpha)\frac{\partial}{\partial x}c(x)^2\hat{p}(x) . \tag{6.6.101}$$

We can further show that

$$P[P, L_3] = 0 . \tag{6.6.102}$$

For since

$$\int d\alpha\, p_x(\alpha) = 1 , \tag{6.6.103}$$

it follows that

$$\int d\alpha\, r_x(\alpha)p_x(\alpha) = \int d\alpha\, s_x(\alpha)p_x(\alpha) = 0 \tag{6.6.104}$$

which is sufficient to demonstrate (6.6.102). Hence, instead of (6.6.92) we may write (6.6.92) in the form

$$PL_2L_1^{-1}\{L_2 + [L_3, P]\}\tilde{v}(s) \tag{6.6.105}$$

and

$$[L_3, P]\tilde{v}(s) = -p_x(\alpha)\{r_x(\alpha)[a(x) + b(x)u(x)] - \tfrac{1}{2}s_x(\alpha)c(x)^2\}\hat{p}(x) + p_x(\alpha)r_x(\alpha)\frac{\partial}{\partial x}[c(x)^2\hat{p}(x)]\,. \tag{6.6.106}$$

6.6.4 An Example with Arbitrary Nonlinear Coupling

We consider the pair of equations

$$\begin{aligned} dx &= \gamma b(x)\alpha dt \\ d\alpha &= -\gamma^2 A(x, \alpha, \gamma)dt + \gamma\sqrt{2B(x, \alpha, \gamma)}\, dW(t) \end{aligned} \tag{6.6.107}$$

and assume the existence of the following limits and asymptotic expansions

$$\begin{aligned} A(x, \alpha, \gamma) &\sim \sum_{n=0}^{\infty} A_n(x, \alpha)\gamma^{-n} \\ B(x, \alpha, \gamma) &\sim \sum_{n=0}^{\infty} B_n(x, \alpha)\gamma^{-n} \end{aligned} \tag{6.6.108}$$

These expansions imply that there is an asymptotic stationary distribution of α at fixed x given by

$$p_s(\alpha, x) = \lim_{\gamma\to\infty} p_s(\alpha, x, \gamma) \tag{6.6.109}$$

$$p_s(\alpha, x) \propto B_0(x, \alpha)^{-1} \exp\{\int d\alpha[A_0(x, \alpha)/B_0(x, \alpha)]\}\,. \tag{6.6.110}$$

We assume that $A_0(x, \alpha)$ and $B_0(x, \alpha)$ are such that

$$\langle\alpha(x)\rangle_s = \int d\alpha\, \alpha p_s(\alpha, x) = 0 \tag{6.6.111}$$

so that we deduce from (6.6.108) that, for finite γ

$$\langle\alpha(x, \gamma)\rangle_s = \int d\alpha\, \alpha p_s(\alpha, x, \gamma) \sim \alpha_0(x)\gamma^{-1} \tag{6.6.112}$$

where $\alpha_0(x)$ can be determined from (6.6.108).

We define the new variables

$$\begin{aligned} \beta &= \alpha - \frac{1}{\gamma}\alpha_0(x) \\ x_1 &= x\,. \end{aligned} \tag{6.6.113}$$

In terms of which the Fokker-Planck operator becomes (the Jacobian is a constant, as usual) on changing x_1 back to x

$$L = -\frac{\partial}{\partial x}\alpha_0(x)b(x) - \gamma\beta\frac{\partial}{\partial x}b(x) + \frac{1}{\gamma}\alpha_0'(x)\alpha_0(x)b(x)\frac{\partial}{\partial\beta} + \frac{\partial}{\partial\beta}\beta\alpha_0'(x)b(x) + \gamma^2\left[\frac{\partial}{\partial\beta}A\left(\frac{\alpha_0(x)}{\gamma} + \beta, x, \gamma\right) + \frac{\partial^2}{\partial\beta^2}B\left(\frac{\alpha_0(x)}{\gamma} + \beta, x, \gamma\right)\right] \tag{6.6.114}$$

and, by using the asymptotic expansions (6.6.108), we can write this as

$$L = L_3 + \gamma L_2(\gamma) + \gamma^2 L_1 \tag{6.6.115}$$

with

$$L_3 = -\frac{\partial}{\partial x}\alpha_0(x)b(x) \tag{6.6.116}$$

$$L_1 = \frac{\partial}{\partial\beta}A_0(\beta, x) + \frac{\partial^2}{\partial\beta^2}B_0(\beta, x) \tag{6.6.117}$$

$$L_2(\gamma) = L_2 + O(\gamma^{-1}) \tag{6.6.118}$$

$$L_2 = -\beta\frac{\partial}{\partial x}b(x) - \frac{\partial}{\partial\beta}\left[\frac{\partial A_0(\beta, x)}{\partial\beta}\alpha_0(x) + A_1(\beta,x)\right] - \frac{\partial^2}{\partial\beta^2}\left[\frac{\partial B_0(\beta, x)}{\partial\beta}\alpha_0(x) + B_1(\beta, x)\right]. \tag{6.6.119}$$

We note that L_3 and L_1 do not commute, but, as in Sect. 6.5.4, this does not affect the limiting result,

$$\frac{\partial\hat{p}}{\partial t} = (L_3 - PL_2L_1^{-1}L_2)\hat{p}\ . \tag{6.6.120}$$

The evaluation of the $PL_2L_1^{-1}L_2$ term is straightforward, but messy. We note that the terms involving $\partial/\partial\beta$ vanish after being operated on by P. From the explicit form of $p_s(\alpha, x)$ one can define $G(\beta, x)$ by

$$\frac{\partial}{\partial\beta}\left\{\left[\frac{\partial A_0(\beta, x)}{\alpha\beta}\alpha_0(x) + A_1(\beta, x)\right]p_s(\beta, x)\right\} + \left\{\frac{\partial^2}{\partial\beta^2}\left[\frac{\partial B_0(\beta, x)}{\partial\beta}\alpha_0(x) + B_1(\beta, x)\right]\right\}p_s(\beta, x) = G(\beta, x)p_s(\beta, x) \tag{6.6.121}$$

and one finds that

$$PL_2L_1^{-1}L_2\hat{p} = \left[\frac{\partial}{\partial x}b(x)D(x)\frac{\partial}{\partial x}b(x) + \frac{\partial}{\partial x}b(x)E(x)\right]\hat{p} \tag{6.6.122}$$

with

$$D(x) = \int_0^\infty dt\langle\beta(t), \beta(0)\,|\,x\rangle \tag{6.6.123}$$

$$E(x) = \int_0^\infty dt\langle\beta(t), G(\beta, x)\,|\,x\rangle$$

where $\langle \ldots | x\rangle$ indicates an average over $p_s(\beta, x)$. This is a rather strong adiabatic elimination result, in which an arbitrary nonlinear elimination can be handled and a finite resulting noise dealt with. The calculation is simpler than that in the previous section, since the terms involving L_3 are of lower order here than those involving L_2.

7. Master Equations and Jump Processes

It is very often the case that in systems involving numbers of particles, or individual objects (animals, bacteria, etc) that a description in terms of a jump process can be very plausibly made. In such cases we find, as first mentioned in Sect. 1.1, that in an appropriate limit *macroscopic deterministic* laws of motion arise, about which the random nature of the process generates a fluctuating part. However the deterministic motion and the fluctuations arise directly out of the same description in terms of individual jumps, or transitions. In this respect, a description in terms of a jump process (and its corresponding master equation) is very satisfactory.

In contrast, we could model such a system approximately in terms of stochastic differential equations, in which the deterministic motion and the fluctuations have a completely independent origin. In such a model this independent description of fluctuations and deterministic motion is an embarrassment, and fluctuation dissipation arguments are necessary to obtain some information about the fluctuations. In this respect the master equation approach is a much more complete description.

However the existence of the macroscopic deterministic laws is a very significant result, and we will show in this chapter that there is often a limit in which the solution of a master equation can be approximated asymptotically (in terms of a large parameter Ω describing the system size) by a deterministic part (which is the solution of a deterministic differential equation), plus a fluctuating part, describable by a stochastic differential equation, whose coefficients are given by the original master equation. Such asymptotic expansions have already been noted in Sect. 3.8.3, when we dealt with the Poisson process, a very simple jump process, and are dealt with in detail in Sect. 7.2.

The result of these expansions is the development of rather simple rules for writing Fokker-Planck equations equivalent (in an asymptotic approximation) to master equations, and in fact it is often in practice quite simple to write down the appropriate approximate Fokker-Planck equation without ever formulating the master equation itself. There are several different ways of formulating the first-order approximate Fokker-Planck equation, all of which are equivalent. However, there is as yet only one way of systematically expanding in powers of Ω^{-1}, and that is the system size expansion of van Kampen.

The chapter concludes with an outline of the *Poisson representation*, a method devised by the author and co-workers, which, for a class of master equations, sets up a Fokker-Planck equation *exactly* equivalent to the master equation. In this special case, the system size expansion arises as a small noise expansion of the Poisson representation Fokker-Planck equation.

7.1 Birth-Death Master Equations—One Variable

The one dimensional prototype of all birth-death systems consists of a population of individuals X in which the number that can occur is called x, which is a non-negative integer. We are led to consider the conditional probability $P(x, t \mid x', t')$ and its corresponding master equation. The concept of birth and death is usually that *only a finite number of X are created (born) or destroyed (die) in a given event.* The simplest case is when the X are born or die one at a time, with a time independent probability so that the transition probabilities $W(x \mid x', t)$ can be written

$$W(x \mid x', t) = t^+(x')\delta_{x,x'+1} + t^-(x')\delta_{x,x'-1} . \tag{7.1.1}$$

Thus there are two processes,

$$x \to x + 1: \qquad t^+(x) = \text{transition probability per unit time.} \tag{7.1.2}$$

$$x \to x - 1: \qquad t^-(x) = \text{transition probability per unit time.} \tag{7.1.3}$$

The general master equation (3.5.5) then takes the form

$$\begin{aligned} \partial_t P(x, t \mid x', t') = {} & t^+(x-1)P(x-1, t \mid x', t') + t^-(x+1)P(x+1, t \mid x', t') \\ & - [t^+(x) + t^-(x)]P(x, t \mid x', t') . \end{aligned} \tag{7.1.4}$$

There are no general methods of solving this equation, except in the time-independent situation.

7.1.1 Stationary Solutions

We can write the equation for the stationary solution $P_s(x)$ as

$$0 = J(x + 1) - J(x) \tag{7.1.5}$$

with

$$J(x) = t^-(x)P_s(x) - t^+(x - 1)P_s(x - 1) . \tag{7.1.6}$$

We now take note of the fact that x is a non-negative integer; we cannot have a negative number of individuals. This requires

$$\text{(i) } t^-(0) = 0: \qquad \text{no probability of an individual dying if there are none present;} \tag{7.1.7}$$

$$\text{(ii) } P(x, t \mid x', t') = 0 \qquad \text{for } x < 0 \text{ or } x' < 0 . \tag{7.1.8}$$

This means that

$$J(0) = t^-(0)P_s(0) - t^+(-1)P_s(-1) = 0 . \tag{7.1.9}$$

We now sum (7.1.5) so

$$0 = \sum_{z=0}^{x-1} [J(z+1) - J(z)] = J(x) - J(0) . \tag{7.1.10}$$

Hence,

$$J(x) = 0 \tag{7.1.11}$$

and thus

$$P_s(x) = \frac{t^+(x-1)}{t^-(x)} P_s(x-1) \tag{7.1.12}$$

so that

$$\boxed{P_s(x) = P_s(0) \prod_{z=1}^{x} \frac{t^+(z-1)}{t^-(z)} .} \tag{7.1.13}$$

a) Detailed Balance Interpretation

The condition $J(x) = 0$ can be viewed as a *detailed balance requirement*, in which x is an even variable. For, it is clear that it is a form of the detailed balance condition (5.3.74), which takes the form here of

$$P(x, \tau | x', 0) P_s(x') = P(x', \tau | x, 0) P_s(x) , \tag{7.1.14}$$

Setting $x' = x \pm 1$ and taking the limit $\tau \to 0$, and noting that by definition (3.4.1),

$$W(x | x', t) = \lim_{\tau \to 0} P(x, t + \tau | x', t)/\tau , \tag{7.1.15}$$

the necessity of this condition is easily proved.

b) Rate Equations

We notice that the mean of x satisfies

$$\partial_t \langle x(t) \rangle = \partial_t \sum_{x=0}^{\infty} x P(x, t | x', t') \tag{7.1.16}$$

$$= \sum_{x=0}^{\infty} x[t^+(x-1)P(x-1, t | x', t') - t^+(x)P(x, t | x', t')]$$

$$+ \sum_{x=0}^{\infty} x[t^-(x+1)P(x+1, t | x', t') - t^-(x)P(x, t | x', t')] \tag{7.1.17}$$

$$= \sum_{x=0}^{\infty} [(x+1)t^+(x) - xt^+(x) + (x-1)t^-(x)$$

$$- xt^-(x)]P(x, t | x', t') , \tag{7.1.18}$$

i.e.,

$$\boxed{\frac{d}{dt} \langle x(t) \rangle = \langle t^+[x(t)] \rangle - \langle t^-[x(t)] \rangle .} \tag{7.1.19}$$

The corresponding deterministic equation is that which would be obtained by neglecting fluctuations, i.e.,

$$\frac{dx}{dt} = t^+(x) - t^-(x) . \tag{7.1.20}$$

Notice that a stationary state occurs deterministically when

$$t^+(x) = t^-(x) . \tag{7.1.21}$$

Corresponding to this, notice that the maximum value of $P_s(x)$ occurs when

$$P_s(x)/P_s(x-1) \simeq 1 , \tag{7.1.22}$$

which from (7.1.12) corresponds to

$$t^+(x-1) = t^-(x) . \tag{7.1.23}$$

Since the variable x takes on only integral values, for sufficiently large x (7.1.21) and (7.1.23) are essentially the same.

Thus, the *modal value of* x, which corresponds to (7.1.23), is the stationary stochastic analogue of the deterministic steady state which corresponds to (7.1.21).

7.1.2 Example: Chemical Reaction $X \rightleftharpoons A$

We treat the case of a reaction $X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} A$ in which it is assumed that A is a fixed concentration. Thus, we assume

$$t^+(x) = k_2 a \tag{7.1.24}$$

$$t^-(x) = k_1 x \tag{7.1.25}$$

so that the Master equation takes the simple form [in which we abbreviate $P(x, t \mid x', t')$ to $P(x, t)$]

$$\partial_t P(x,t) = k_2 a P(x-1,t) + k_1(x+1)P(x+1,t) - (k_1 x + k_2 a)P(x,t). \tag{7.1.26}$$

a) Generating Function

To solve the equation, we introduce the generating function (c.f. Sects.1.4.1, 3.8.2)

$$G(s,t) = \sum_{x=0}^{\infty} s^x P(x,t) \tag{7.1.27}$$

so that

$$\boxed{\partial_t G(s,t) = k_2 a(s-1)G(s,t) - k_1(s-1)\partial_s G(s,t) .} \tag{7.1.28}$$

If we substitute

$$\phi(s, t) = G(s, t) \exp(-k_2 a s/k_1), \tag{7.1.29}$$

(7.1.28) becomes

$$\partial_t \phi(s, t) = -k_1(s-1)\partial_s \phi(s, t). \tag{7.1.30}$$

The further substitution $s - 1 = e^z$,

$$\phi(s, t) = \psi(z, t)$$

gives

$$\partial_t \psi(z, t) + k_1 \partial_z \psi(z, t) = 0 \tag{7.1.31}$$

whose solution is an arbitrary function of $k_1 t - z$. For convenience, write this as

$$\begin{aligned}\psi(z, t) &= F[\exp(-k_1 t + z)]\, e^{-k_2 a/k_1} \\ &= F[(s-1)e^{-k_1 t}]\, e^{-k_2 a/k_1}\end{aligned}$$

so

$$\boxed{G(s, t) = F[(s-1)e^{-k_1 t}] \exp[(s-1)k_2 a/k_1].} \tag{7.1.32}$$

Normalisation requires $G(1, t) = 1$, and hence

$$F(0) = 1. \tag{7.1.33}$$

b) Conditional Probability

The initial condition determines F; this is (setting $t' = 0$)

$$P(x, 0 \mid N, 0) = \delta_{x,N} \tag{7.1.34}$$

which means

$$G(s, 0) = s^N = F(s-1) \exp[(s-1)k_2 a/k_1] \tag{7.1.35}$$

so that

$$G(s, t) = \exp\left[\frac{k_2 a}{k_1}(s-1)(1 - e^{-k_1 t})\right][1 + (s-1)e^{-k_1 t}]^N. \tag{7.1.36}$$

This can now be expanded in a power series in s giving

$$\boxed{\begin{aligned}P(x, t \mid N, 0) = \exp\left[-\frac{k_2 a}{k_1}(1 - e^{-k_1 t})\right] \sum_{r=0}^{x} \frac{N!}{(N-r)!\,r!\,(x-r)!} \left(\frac{k_2 a}{k_1}\right)^{x-r} \\ \times (1 - e^{-k_1 t})^{N+x-2r} e^{-k_1 t r}.\end{aligned}} \tag{7.1.37}$$

This very complicated answer is a complete solution to the problem but is of very little practical use. It is better to work either directly from (7.1.36), the generating function, or from the equations for mean values.

From the generating function we can compute

$$\langle x(t)\rangle = \partial_s G(s=1,t) = \frac{k_2 a}{k_1}(1 - e^{-k_1 t}) + N e^{-k_1 t} \tag{7.1.38}$$

$$\langle x(t)[x(t)-1]\rangle = \partial_s^2 G(s=1,t) = \langle x(t)\rangle^2 - N e^{-2k_1 t} \tag{7.1.39}$$

$$\mathrm{var}\{x(t)\} = \left(N e^{-k_1 t} + \frac{k_2 a}{k_1}\right)(1 - e^{-k_1 t}) . \tag{7.1.40}$$

c) Moment Equations

From the differential equation (7.1.28) we have

$$\partial_t[\partial_s^n G(s,t)] = \{n[k_2 a \partial_s^{n-1} - k_1 \partial_s^n] + (s-1)[k_2 a \partial_s^n - k_1 \partial_s^{n+1}]\} G(s,t) . \tag{7.1.41}$$

Setting $s = 1$ and using

$$\partial_s^n G(s,t)|_{s=1} = \langle x(t)^n\rangle_f , \tag{7.1.42}$$

we find

$$\frac{d}{dt}\langle x(t)^n\rangle_f = n[k_2 a\langle x(t)^{n-1}\rangle_f - k_1\langle x(t)^n\rangle_f] \tag{7.1.43}$$

and these equations form a closed hierarchy. Naturally, the mean and variance solutions correspond to (7.1.38, 40).

d) Autocorrelation Function and Stationary Distribution

As $t \to \infty$ for any F, we find from (7.1.32, 33)

$$G(s, t \to \infty) = \exp[(s-1)k_2 a/k_1] \tag{7.1.44}$$

corresponding to the *Poissonian solution:*

$$P_s(x) = \exp(-k_2 a/k_1)(k_2 a/k_1)^x/x!. \tag{7.1.45}$$

Since the equation of time evolution for $\langle x(t)\rangle$ is *linear*, we can apply the methods of Sect.3.7.4, namely, the regression theorem, which states that the stationary autocorrelation function has the same time dependence as the mean, and its value at $t = 0$ is the stationary variance. Hence,

$$\langle x(t)\rangle_s = k_2 a/k_1 \tag{7.1.46}$$

$$\mathrm{var}\{x(t)\}_s = k_2 a/k_1 \tag{7.1.47}$$

$$\langle x(t), x(0)\rangle_s = e^{-k_1 t} k_2 a/k_1 . \tag{7.1.48}$$

The Poissonian stationary solution also follows from (7.1.13) by direct substitution.

e) Poissonian Time-Dependent Solutions

A very interesting property of this equation is the existence of *Poissonian time-dependent solutions*. For if we choose

$$P(x, 0) = \frac{e^{-\alpha_0}\alpha_0^x}{x!}, \tag{7.1.49}$$

then

$$G(s, 0) = \exp[(s - 1)\alpha_0] \tag{7.1.50}$$

and from (7.1.32) we find

$$G(s, t) = \exp\{(s - 1)[\alpha_0 e^{-k_1 t} + (k_2 a/k_1)(1 - e^{-k_1 t})]\} \tag{7.1.51}$$

corresponding to

$$P(x, t) = \frac{e^{-\alpha(t)}\alpha(t)^x}{x!} \tag{7.1.52}$$

with

$$\alpha(t) = \alpha_0 e^{-k_1 t} + (k_2 a/k_1)(1 - e^{-k_1 t}) . \tag{7.1.53}$$

Here $\alpha(t)$ is seen to be the solution of the deterministic equation

$$\frac{dx}{dt} = k_2 a - k_1 x \tag{7.1.54}$$

with the initial condition $x(0) = \alpha_0$. (7.1.55)

This result can be generalised to many variables and forms the rationale for the *Poisson representation* which will be developed in Sect. 7.7. The existence of Poissonian propagating solutions is a consequence of the linearity of the system.

7.1.3 A Chemical Bistable System

We consider the system

$$A + 2X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3X \tag{7.1.56}$$

$$A \underset{k_4}{\overset{k_3}{\rightleftharpoons}} X \tag{7.1.57}$$

which has been studied by many authors [7.1]. The concentration of A is held fixed so that we have

$$t^+(x) = k_1 A x(x-1) + k_3 A$$
$$t^-(x) = k_2 x(x-1)(x-2) + k_4 x . \tag{7.1.58}$$

The corresponding deterministic equation is, of course,

$$\frac{dx}{dt} = t^+(x) - t^-(x)$$
$$\simeq -k_2 x^3 + k_1 A x^2 - k_4 x + k_3 A , \tag{7.1.59}$$

where it is assumed that $x \gg 1$ so that we set $x(x-1)(x-2) \simeq x^3$, etc. The solution of this equation, with the initial condition $x(0) = x_0$, is given by

$$\left(\frac{x-x_1}{x_0-x_1}\right)^{x_3-x_2}\left(\frac{x-x_2}{x_0-x_2}\right)^{x_1-x_3}\left(\frac{x-x_3}{x_0-x_3}\right)^{x_2-x_1}$$
$$= \exp\left[-k_2(x_1-x_2)(x_2-x_3)(x_3-x_1)t\right] . \tag{7.1.60}$$

Here, x_1, x_2, x_3 are roots of

$$k_2 x^3 - k_1 A x^2 + k_4 x - x_3 A = 0 \tag{7.1.61}$$

with $x_3 \geqslant x_2 \geqslant x_1$.

Clearly these roots are the stationary values of the solutions $x(t)$ of (7.1.59). From (7.1.59) we see that

$$\begin{aligned} x < x_1 \quad &\Longrightarrow \frac{dx}{dt} > 0 \\ x_2 > x > x_1 \quad &\Longrightarrow \frac{dx}{dt} < 0 \\ x_3 > x > x_2 \quad &\Longrightarrow \frac{dx}{dt} > 0 \\ x > x_3 \quad &\Longrightarrow \frac{dx}{dt} < 0 . \end{aligned} \tag{7.1.62}$$

Thus, in the region $x < x_2$, $x(t)$ will be attracted to x_1 and in the region $x > x_2$, $x(t)$ will be attracted to x_3. The solution $x(t) = x_2$ will be unstable to small perturbations. This yields a system with two deterministically stable stationary states.

a) Stochastic Stationary Solution

From (7.1.13)

$$P_s(x) = P_s(0) \prod_{z=1}^{x} \left\{\frac{B[(z-1)(z-2)+P]}{z(z-1)(z-2)+Rz}\right\} , \tag{7.1.63}$$

where

$$B = k_1 A/k_2$$
$$R = k_4/k_2 \qquad (7.1.64)$$
$$P = k_3/k_1 .$$

Notice that if $P = R$, the solution (7.1.63) is Poissonian with mean B. In this case, we have a stationary state in which reactions (7.1.56, 57) are simultaneously in balance. This is *chemical equilibrium*, in which, as we will show later, there is always a Poissonian solution (Sects. 7.5.1 and 7.7b). The maxima of (7.1.63) occur, according to (7.1.21), when

$$B = x[(x-1)(x-2) + R]/[P + x(x-1)] . \qquad (7.1.65)$$

The function $x = x(B)$, found by inverting (7.1.65), gives the maxima (or minima) corresponding to that value of B for a given P and R.

There are the two asymptotic forms:

$$x(B) \sim B \quad \text{large } B$$
$$x(B) \sim PB/R \quad \text{small } B \qquad (7.1.66)$$

If $R > 9P$, we can show that the slope of $x(B)$ becomes negative for some range of $x > 0$ and thus we get three solutions for a given B, as shown in Fig. 7.1. The transition from one straight line to the other gives the kink that can be seen.

Notice also that for the choice of parameters shown, the bimodal shape is significant over a very small range of B. This range is very much narrower than the range over which $P(x)$ is two peaked, since the ratio of the heights of the peaks can be very high.

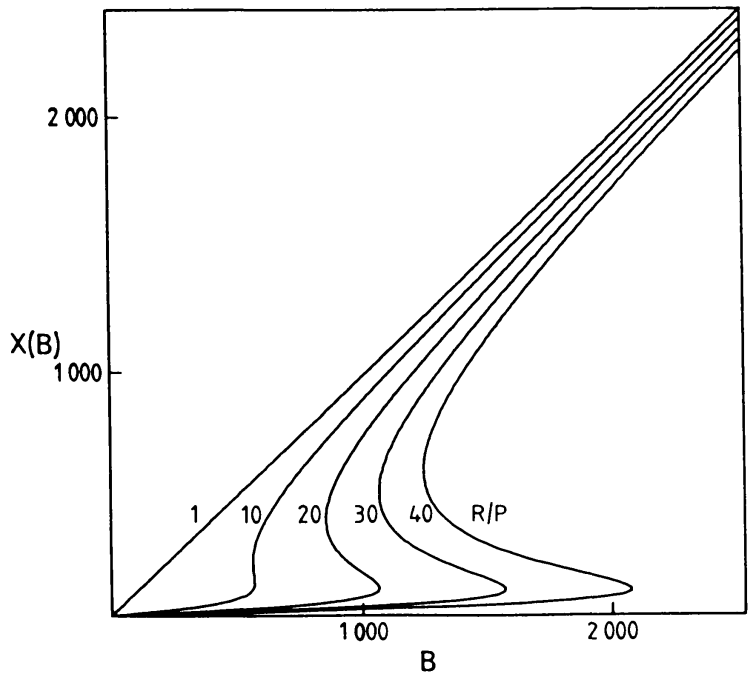


Fig. 7.1. Plot of $x(B)$ against B, as given by the solution of (7.1.65) for various values of R/P, and $P = 10{,}000$

A more precise result can be given. Suppose the volume V of the system becomes very large and the concentration y of X given by

$$y = x/V\,,$$

is constant. Clearly the transition probabilities must scale like V, since the rate of production of X will scale like $x = yV$.

Hence,

$$\begin{aligned} k_1 A &\sim 1/V \\ k_3 A &\sim V \\ k_2 &\sim 1/V^2 \\ k_4 &\sim 1 \end{aligned} \tag{7.1.67}$$

which means that

$$\begin{aligned} B &\sim V \\ R &\sim V^2 \\ P &\sim V^2\,. \end{aligned} \tag{7.1.68}$$

We then write

$$\begin{aligned} B &= \tilde{B}V \\ R &= \tilde{R}V^2 \\ P &= \tilde{P}V^2 \end{aligned}$$

so that (7.1.65) becomes

$$\tilde{B} \simeq y(y^2 + \tilde{R})/(y^2 + \tilde{P})\,.$$

And if y_1 and y_2 are two values of y,

$$\begin{aligned} \log\,[P_s(y_2)/P_s(y_1)] \simeq \sum_{z=y_1V}^{y_2V} \{&\log \tilde{B}V + \log\,(z^2 + \tilde{P}V^2) \\ &- \log\,[z(z^2 + \tilde{R}V^2)]\} \end{aligned} \tag{7.1.69}$$

and we now approximate by an integral

$$\sim V\int_{y_1}^{y_2} dy\left[\log\left(\frac{\tilde{B}(y^2 + \tilde{P})}{y(y^2 + \tilde{R})}\right)\right].$$

Hence,

$$\frac{P_s(y_2)}{P_s(y_1)} \simeq \exp\left[V\int_{y_1}^{y_2} \log\left(\frac{\tilde{B}(y^2 + \tilde{P})}{y(y^2 + \tilde{R})}\right)\right] \tag{7.1.70}$$

and as $V \to \infty$, depending on the sign of the integral, this ratio becomes either zero or infinity. Thus, in a large volume limit, the two peaks, unless precisely equal in height, become increasingly unequal and only one survives.

The variance of the distribution can be obtained by a simple trick. Notice from (7.1.63) that $P_s(x)$ can be written

$$P_s(x) = B^x G(x)\,, \tag{7.1.71}$$

where $G(x)$ is a function defined through (7.1.63). Then,

$$\left.\begin{aligned} &\langle x^k\rangle = \left[\sum_{x=0}^{\infty} x^k B^x G(x)\right]\left[\sum_{x=0}^{\infty} B^x G(x)\right]^{-1} \\ &\text{and} \\ &B\frac{\partial}{\partial B}\langle x^k\rangle = \langle x^{k+1}\rangle - \langle x\rangle\langle x^k\rangle \end{aligned}\right\} \tag{7.1.72}$$

so that

$$\mathrm{var}\{x\} = B\frac{\partial}{\partial B}\langle x\rangle\,. \tag{7.1.73}$$

From this we note that as $V \to \infty$,

$$\mathrm{var}\{y\} \sim \frac{1}{V} \to 0\,. \tag{7.1.74}$$

So a deterministic limit is approached. Further, notice that if $\langle x\rangle$ is proportional to B, the variance is equal to the mean, in fact, we find on the two branches (7.1.66),

$$\begin{aligned} \langle x(B)\rangle &= \mathrm{var}\{x(B)\} = B \qquad \text{large } B \\ \langle x(B)\rangle &= \mathrm{var}\{x(B)\} = PB/R \quad \text{small } B \end{aligned}$$

which means that the distributions are roughly Poissonian on these limiting branches.

The stochastic mean is not, in fact, given exactly by the peak values but approximates it very well. Of course, for any B there is one well defined $\langle x(B)\rangle$, not three values. Numerical computations show that the mean closely follows the lower branch and then suddenly makes a transition at B_c to the upper branch. This will be the value at which the two maxima have equal height and can, in principle, be determined from (7.1.70).

b) Time-Dependent Behaviour

This is impossible to deduce exactly. Almost all approximation methods depend on the large volume limit, whose properties in a stationary situation have just been noted and which will be dealt with systematically in the next section.

7.2 Approximation of Master Equations by Fokker-Planck Equations

The existence of the parameter V in terms of which well-defined scaling laws are valid leads to the concept of *system size expansions*, first put on a systematic basis by *van Kampen* [7.2]. There is a confused history attached to this, which arises out of repeated attempts to find a limiting form of the Master equation in which a Fokker-Planck equation arises. However, the fundamental result is that a diffusion process can always be approximated by a jump process, not the reverse.

7.2.1 Jump Process Approximation of a Diffusion Process

The prototype result is that found for the random walk in Sect.3.8.2, that in the limit of infinitely small jump size, the Master equation becomes a Fokker-Planck equation. Clearly the jumps must become more probable and smaller, and this can be summarised by a *scaling assumption:* that there is a parameter δ, such that the average step size and the variance of the step size are proportional to δ, and such that the jump probabilities increase as δ becomes small.

We assume that the jump probabilities can be written

$$W_\delta(x' \mid x) = \Phi\left(\frac{x' - x - A(x)\delta}{\sqrt{\delta}}, x\right)\delta^{-3/2}, \tag{7.2.1}$$

where

$$\int dy\, \Phi(y, x) = Q \tag{7.2.2}$$

and

$$\int dy\, y\, \Phi(y, x) = 0\,. \tag{7.2.3}$$

This means that

$$\begin{aligned} \alpha_0(x) &\equiv \int dx' W_\delta(x' \mid x) = Q/\delta \\ \alpha_1(x) &\equiv \int dx'(x' - x) W_\delta(x' \mid x) = A(x)Q \\ \alpha_2(x) &\equiv \int dx'(x' - x)^2 W_\delta(x' \mid x) = \int dy\, y^2 \Phi(y, x)\,. \end{aligned} \tag{7.2.4}$$

We further assume that $\Phi(y, x)$ vanishes sufficiently rapidly as $y \to \infty$, so that

$$\lim_{\delta \to 0} W_\delta(x' \mid x) = \lim_{y \to \infty}\left[\left(\frac{y}{x' - x}\right)^3 \Phi(y, x)\right] = 0 \qquad \text{for} \quad x' \neq x\,. \tag{7.2.5}$$

The conditions (7.2.4, 5) are very similar to those in Sect.3.4, namely, (3.4.1, 4, 5) and by taking a twice differentiable function $f(z)$, one can carry out much the same procedure as that used in Sect. 3.4 to show that

$$\lim_{\delta \to 0} \left\langle \frac{\partial f(z)}{\partial t} \right\rangle = \left\langle \alpha_1(z) \frac{\partial f}{\partial z} + \frac{1}{2} \alpha_2(z) \frac{\partial^2 f}{\partial z^2} \right\rangle \tag{7.2.6}$$

implying that in the limit $\delta \to 0$, the Master equation

$$\frac{\partial P(x)}{\partial t} = \int dx' [W(x|x')P(x') - W(x'|x)P(x)] \tag{7.2.7}$$

becomes the FPE

$$\frac{\partial P(x)}{\partial t} = -\frac{\partial}{\partial x} \alpha_1(x)P(x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \alpha_2(x)P(x) . \tag{7.2.8}$$

Thus, given (7.2.8), one can always construct a Master equation depending on a parameter δ which approximates it as closely as desired. Such a Master equation will have transition probabilities which satisfy the criteria of (7.2.4). If they do *not* satisfy these criteria, then this approximation is not possible. Some examples are appropriate.

a) Random Walk (Sect.3.8.2)
Let $x = nl$, then

$$W(x|x') = d(\delta_{x',x-l} + \delta_{x',x+l}) . \tag{7.2.9}$$

Then

$$\begin{aligned} \alpha_0(x) &= 2d \\ \alpha_1(x) &= 0 \\ \alpha_2(x) &= 2l^2 d ; \end{aligned} \tag{7.2.10}$$

let

$$\delta = l^2 \tag{7.2.11}$$

and

$$D = l^2 d . \tag{7.2.12}$$

Then all requirements are met, so the limiting equation is

$$\frac{\partial P}{\partial t} = D \frac{\partial^2 P}{\partial x^2} \tag{7.2.13}$$

as found in Sect.3.8.2.

b) Poisson Process (Sect.3.8.3)
Here, letting $x = nl$,

$$W(x|x') = d\delta_{x,x'+l} \tag{7.2.14}$$

and

$$\begin{aligned}\alpha_0(x) &= d \\ \alpha_1(x) &= ld \\ \alpha_2(x) &= l^2 d\,.\end{aligned} \tag{7.2.15}$$

There is no way of parametrising l and d in terms of δ such that $l \to 0$ and both $\alpha_1(x)$ and $\alpha_2(x)$ are finite. In this case, there is no Fokker-Planck limit.

c) General Approximation of Diffusion Process by a Birth-Death Master Equation
Suppose we have a Master equation such that

$$W_\delta(x'|x) = \left(\frac{A(x)}{2\delta} + \frac{B(x)}{2\delta^2}\right)\delta_{x',x+\delta} + \left(-\frac{A(x)}{2\delta} + \frac{B(x)}{2\delta^2}\right)\delta_{x',x-\delta} \tag{7.2.16}$$

so that for sufficiently small δ, $W_\delta(x'|x)$ is positive and we assume that this is uniformly possible over the range of x of interest. The process then takes place on a range of x composed of integral multiples of δ. This is *not* of the form of (7.2.1) but, nevertheless, in the limit $\delta \to 0$ gives a FPE. For

$$\alpha_0(x) = B(x)/\delta^2 \tag{7.2.17a}$$

$$\alpha_1(x) = A(x) \tag{7.2.17b}$$

$$\alpha_2(x) = B(x) \tag{7.2.17c}$$

and

$$\lim_{\delta\to 0} W_\delta(x'|x) = 0 \qquad \text{for} \qquad x' \neq x\,. \tag{7.2.17d}$$

Here, however, $\alpha_0(x)$ diverges like $1/\delta^2$, rather than like $1/\delta$ as in (7.2.4) and the picture of a jump taking place according to a smooth distribution is no longer valid. The proof carries through, however, since the behaviour of $\alpha_0(x)$ is irrelevant and the limiting FPE is

$$\frac{\partial P(x)}{\partial t} = -\frac{\partial}{\partial x}A(x)P(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}B(x)P(x)\,. \tag{7.2.18}$$

In this form, we see that we have a possible tool for simulating a diffusion process by an approximating birth-death process. The method fails if $B(x) = 0$ anywhere in the range of x, since this leads to negative $W_\delta(x'|x)$. Notice that the stationary solution of the Master equation in this case is

$$\begin{aligned}P_s(x) &= P_s(0)\prod_{z=\delta}^{x}\left[\frac{\delta A(z-\delta) + B(z-\delta)}{-\delta A(z) + B(z)}\right] \\ &= P_s(0)\left[\frac{-\delta A(0) + B(0)}{\delta A(x) + B(x)}\right]\prod_{z=0}^{x}\left[\frac{1 + \delta A(z)/B(z)}{1 - \delta A(z)/B(z)}\right]\end{aligned} \tag{7.2.19}$$

so that, for small enough δ

$$\log P_s(x) \to \text{const} - \log B(x) + \sum_{z=0}^{x} 2\delta\, A(z)/B(z)\,, \tag{7.2.20}$$

i.e.,

$$P_s(x) \to \frac{\mathcal{N}}{B(x)} \exp\left[2 \int_0^x dz\; A(z)/B(z)\right] \tag{7.2.21}$$

as required. The limit is clearly uniform in any finite interval of x provided $A(x)/B(x)$ is bounded there.

7.2.2 The Kramers-Moyal Expansion

A simple but nonrigorous derivation was given by *Kramers* [7.3] and considerably improved by *Moyal* [7.4]. It was implicitly used by *Einstein* [7.5] as explained in Sect. 1.2.1.

In the Master equation (7.2.7), we substitute x' by defining

$y = x - x'$ in the first term, and

$y = x' - x$ in the second term.

Defining

$$t(y, x) = W(x + y \mid x)\,, \tag{7.2.22}$$

the master equation becomes

$$\frac{\partial P(x)}{\partial t} = \int dy\, [t(y, x - y)P(x - y) - t(y, x)P(x)]\,. \tag{7.2.23}$$

We now expand in power series,

$$= \int dy \sum_{n=1}^{\infty} \frac{(-y)^n}{n!} \frac{\partial^n}{\partial x^n} [t(y, x)P(x)] \tag{7.2.24}$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} [\alpha_n(x)P(x)]\,, \tag{7.2.25}$$

where

$$\alpha_n(x) = \int dx'(x' - x)^n W(x' \mid x) = \int dy\; y^n\, t(y, x)\,. \tag{7.2.26}$$

By terminating the series (7.2.25) at the second term, we obtain the Fokker-Planck equation (7.2.8).

In introducing the system size expansion, van Kampen criticised this "proof", because there is no consideration of what small parameter is being considered.

Nevertheless, this procedure enjoyed wide popularity—mainly because of the convenience and simplicity of the result. However, the demonstration in Sect.7.2.1 shows that there are limits to its validity. Indeed, if we assume that $W(x'|x)$ has the form(7.2.1), we find that

$$\alpha_n(x) = \delta^{n/2-1} \int dy\, y^n \Phi(y, x)\,. \tag{7.2.27}$$

So that as $\delta \to 0$, terms higher than the second in the expansion (7.2.25) (*the Kramers-Moyal expansion*) do vanish. And indeed in his presentation, *Moyal* [7.4] did require conditions equivalent to (7.2.4, 5).

7.2.3 Van Kampen's System Size Expansion [7.2]

Birth-death master equations provide good examples of cases where the Kramers-Moyal expansion fails, the simplest being the Poisson process mentioned in Sect. 7.2.1.

In all of these, the size of the jump is ± 1 or some small integer, whereas typical sizes of the variable may be large, e.g., the number of molecules or the position of the random walker on a long lattice.

In such cases, we can introduce a system size parameter Ω such that the transition probabilities can be written in terms of the intensive variables x/Ω etc. For example, in the reaction of Sect.7.1.3, Ω was the volume V and x/Ω the concentration. Let us use van Kampen's notation:

a = extensive variable (number of molecules, etc $\propto \Omega$)

$x = a/\Omega$ intensive variable (concentration of molecules).

The limit of interest is large Ω at fixed x. This corresponds to the approach to a macroscopic system. We can rewrite the transition probability as

$$W(a|a') = W(a'; \Delta a)$$
$$\Delta a = a - a'. \tag{7.2.28}$$

The essential point is that the size of the jump is expressed in terms of the extensive quantity Δa, but the dependence on a' is better expressed in terms of the intensive variable x.

Thus, we assume that we can write

$$W(a'; \Delta a) = \Omega \psi\left(\frac{a'}{\Omega}; \Delta a\right). \tag{7.2.29}$$

If this is the case, we can now make an expansion. We choose a new variable z so that

$$a = \Omega\phi(t) + \Omega^{1/2} z\,, \tag{7.2.30}$$

where $\phi(t)$ is a function to be determined. It will now be the case that the $\alpha_n(a)$ are proportional to Ω: we will write

$$\alpha_n(a) = \Omega\tilde{\alpha}_n(x)\,. \tag{7.2.31}$$

We now take the Kramers-Moyal expansion (7.2.25) and change the variable to get

$$\frac{\partial P(z,t)}{\partial t} - \Omega^{1/2}\phi'(t)\frac{\partial P(z,t)}{\partial z} = \sum_{n=1}^{\infty}\frac{\Omega^{1-n/2}}{n!} \times \left(-\frac{\partial}{\partial z}\right)^n \tilde{\alpha}_n[\phi(t) + \Omega^{-1/2}z]P(z,t). \tag{7.2.32}$$

The terms of order $\Omega^{1/2}$ on either side will cancel if $\phi(t)$ obeys

$$\phi'(t) = \tilde{\alpha}_1[\phi(t)] \tag{7.2.33}$$

which is the *deterministic equation* expected. We expand $\tilde{\alpha}_n[\phi(t) + \Omega^{-1/2}z]$ in powers of $\Omega^{-1/2}$, rearrange and find

$$\frac{\partial P(z,t)}{\partial t} = \sum_{m=2}^{\infty}\frac{\Omega^{-(m-2)/2}}{m!}\sum_{n=1}^{m}\frac{m!}{n!(m-n)!}\tilde{\alpha}_n^{(m-n)}[\phi(t)]\left(-\frac{\partial}{\partial z}\right)^n z^{m-n}P(z,t)\,. \tag{7.2.34}$$

Taking the large Ω limit, only the $m = 2$ term survives giving

$$\boxed{\frac{\partial P(z,t)}{\partial t} = -\tilde{\alpha}_1'[\phi(t)]\frac{\partial}{\partial z}z\,P(z,t) + \frac{1}{2}\tilde{\alpha}_2[\phi(t)]\frac{\partial^2}{\partial z^2}P(z,t)\,.} \tag{7.2.35}$$

a) Comparison with Kramers-Moyal Result

The Kramers-Moyal Fokker-Planck equation, obtained by terminating (7.2.25) after two terms, is

$$\frac{\partial P(a)}{\partial t} = -\frac{\partial}{\partial a}[\alpha_1(a)P(a)] + \frac{1}{2}\frac{\partial^2}{\partial a^2}[\alpha_2(a)P(a)] \tag{7.2.36}$$

and changing variables to $x = a/\Omega$, we get

$$\frac{\partial P(x)}{\partial t} = -\frac{\partial}{\partial x}[\tilde{\alpha}_1(x)P(x)] + \frac{1}{2\Omega}\frac{\partial^2}{\partial x^2}[\tilde{\alpha}_2(x)P(x)]\,. \tag{7.2.37}$$

We can now use the small noise theory of Sect. 6.3, with

$$\varepsilon^2 = \frac{1}{\Omega} \tag{7.2.38}$$

and we find that substituting

$$z = \Omega^{1/2}[x - \phi(t)]\,, \tag{7.2.39}$$

the lowest-order FPE for z is exactly the same as the lowest-order term in van

Kampen's method (7.2.35). This means that if we are only interested in the lowest order, we may use the Kramers-Moyal Fokker-Planck equation which may be easier to handle than van Kampen's method. The results will differ, but to lowest order in $\Omega^{-1/2}$ will agree, and each will only be valid to this order.

Thus, if a FPE has been obtained from a Master equation, its validity depends on the kind of limiting process used to derive it. If it has been derived in a limit $\delta \to 0$ of the kind used in Sect.7.2.1, then it can be taken seriously and the full nonlinear dependence of $\alpha_1(a)$ and $\alpha_2(a)$ on a can be exploited.

On the other hand, if it arises as the result of an Ω expansion like that in Sect. 7.2.3, only the small noise approximation has any validity. There is no point in considering anything more than the linearisation, (7.2.35), about the deterministic solution. The solution of this equation is given in terms of the corresponding stochastic differential equation

$$dz = \tilde{\alpha}_1'[\phi(t)]z\,dt + \sqrt{\tilde{\alpha}_2[\phi(t)]}\,dW(t)\,. \tag{7.2.40}$$

by the results of Sect. 4.4.7 (4.4.69), or Sect. 4.4.9 (4.4.83).

b) Example: Chemical Reaction $X \rightleftharpoons A$

From Sect. 7.1.2, we have

$$W(x|x') = \delta_{x,x'+1}k_2a + \delta_{x,x'-1}k_1x'\,. \tag{7.2.41}$$

The assumption is

$$\begin{aligned} a &= a_0V \\ x &= x_0V\,, \end{aligned} \tag{7.2.42}$$

where V is the volume of the system. This means that we assume the total amounts of A and X to be proportional to V (a reasonable assumption) and that the rates of production and decay of X are proportional to a and x, respectively.

Thus,

$$W(x_0'; \Delta x) = V[k_2a_0\delta_{\Delta x,1} + k_1x_0'\delta_{\Delta x,-1}] \tag{7.2.43}$$

which is in the form of (7.2.29), with $\Omega \to V$, $a \to x$, etc.
Thus,

$$\begin{aligned} \alpha_1(x) &= \sum(x' - x)W(x'|x) = k_2a - k_1x = V(k_2a_0 - k_1x_0) \\ \alpha_2(x) &= \sum(x' - x)^2W(x'|x) = k_2a + k_1x = V(k_2a_0 + k_1x_0)\,. \end{aligned} \tag{7.2.44}$$

The deterministic equation is

$$\phi'(t) = [k_2a_0 - k_1\phi(t)] \tag{7.2.45}$$

whose solutions is

$$\phi(t) = \phi(0)e^{-k_1t} + \frac{k_2a_0}{k_1}(1 - e^{-k_1t})\,. \tag{7.2.46}$$

The Fokker-Planck equation is

$$\frac{\partial P(z)}{\partial t} = k_1 \frac{\partial}{\partial z} zP(z) + \frac{1}{2} \frac{\partial^2}{\partial z^2} [k_2 a_0 + k_1 \phi(t)] P(z) . \tag{7.2.47}$$

From (4.4.84, 85) we can compute that

$$\langle z(t) \rangle = z(0) e^{-k_1 t} . \tag{7.2.48}$$

Usually, one would assume $z(0) = 0$, since the initial condition can be fully dealt with by the initial condition on ϕ. Assuming $z(0)$ is zero, we find

$$\text{var}\{z(t)\} = \left[\frac{k_2 a_0}{k_1} + \phi(0) e^{-k_1 t} \right] (1 - e^{-k_1 t}) \tag{7.2.49}$$

so that

$$\langle x(t) \rangle = V\phi(t) = V\phi(0) e^{-k_1 t} + \frac{k_2 a}{k_1} (1 - e^{-k_1 t}) \tag{7.2.50}$$

$$\text{var}\{x(t)\} = V \text{ var}\{z(t)\} = \left[\frac{k_2 a}{k_1} + V\phi(0) e^{-k_1 t} \right] (1 - e^{-k_1 t}) . \tag{7.2.51}$$

With the identification $V\phi(0) = N$, these are exactly the same as the exact solutions (7.1.38–40) in Sect. 7.1.2. The stationary solution of (7.2.47) is

$$P_s(z) = \mathcal{N} \exp\left(- \frac{k_1 z^2}{2 k_2 a_0} \right) \tag{7.2.52}$$

which is Gaussian approximation to the exact Poissonian.

The stationary solution of the Kramers-Moyal equation is

$$\begin{aligned} P_s(x) &= \frac{\mathcal{N}}{\alpha_2(x)} \exp\left[\int_0^x 2 \frac{\alpha_1(x')}{\alpha_2(x')} dx' \right] \\ &= \mathcal{N} (k_2 a + k_1 x)^{-1 + 4k_2 a / k_1} e^{-2x} . \end{aligned} \tag{7.2.53}$$

In fact, one can explicitly check the limit by setting

$$x = V(k_2 a_0 / k_1) + \delta \tag{7.2.54}$$

so that

$$(7.2.53) \simeq \mathcal{N} (2Vk_2 a_0 + k_1 \delta)^{-1 + 4Vk_2 a_0 / k_1} e^{-2Vk_2 a_0 / k_1 - 2\delta} . \tag{7.2.55}$$

Then,

$$\log P_s(x) \simeq \text{const} - \frac{k_2}{2 k_2 a_0 V} (\delta - \delta^2) . \tag{7.2.56}$$

Using the exact Poissonian solution, making the same substitution and using Stirling's formula

$$\log x! \sim (x + \tfrac{1}{2}) \log x - x + \text{const}\,, \tag{7.2.57}$$

one finds the same result as (7.2.56), but the exact results are different, in the sense that even the ratio of the logarithms is different.

The term linear in δ is, in fact, of lower order in V: because using (7.2.39), we find $\delta = z\sqrt{V}$ and

$$\log P_s(z) \sim \text{const} - \frac{k_1}{2k_2 a_0}\left(\frac{z}{\sqrt{V}} - z^2\right) \tag{7.2.58}$$

so that in the large V limit, we have a simple Gaussian with zero mean.

c) Moment Hierarchy

From the expansion (7.2.34), we can develop equations for the moments

$$\langle z^k \rangle = \int dz\, P(z, t) z^k \tag{7.2.59}$$

by direct substitution and integration by parts:

$$\frac{d}{dt}\langle z^k \rangle = \sum_{m=2}^{\infty} \frac{\Omega^{-(m-2)/2}}{m!} \sum_{n=1}^{m,k} \frac{m!k!}{n!(m-n)!(k-n)!} \tilde{\alpha}_n^{(m-n)}[\phi(t)] \langle z^{m+k-2n} \rangle\,. \tag{7.2.60}$$

One can develop a hierarchy by expanding $\langle z^k \rangle$ in inverse powers of $\Omega^{1/2}$:

$$\langle z^k \rangle = \sum_{r=0}^{\infty} M_r^k \Omega^{-r/2}\,. \tag{7.2.61}$$

From such a hierarchy one can compute stationary moments and autocorrelation functions using the same techniques as those used in handling the moment hierarchy for the small noise expansion of the Fokker-Planck equation in Sect.6.3.1. *Van Kampen* [7.2] has carried this out.

7.2.4 Kurtz's Theorem

Kurtz [7.6] has demonstrated that in a certain sense, the Kramers-Moyal expansion can give rise to a slightly stronger result than van Kampen's expansion. For the restricted class of birth-death processes with polynomial transition probabilities, he has shown the following. We consider the stochastic process obeying a birth-death master equation

$$\partial_t P(a, t) = \sum_{a'} W(a|a')P(a', t) - \sum_{a'} W(a'|a)P(a, t) \tag{7.2.62}$$

in which the scaling condition (7.2.29) is satisfied.

Then the process $b(t)$, satisfying the stochastic differential equation

$$db(t) = \alpha_1(b)dt + \sqrt{\alpha_2(b)}\, dW(t) \tag{7.2.63}$$

exists, and to each sample path $a(t)$ of (7.2.62) a sample path of $b(t)$ of (7.2.63) exists such that

$$|b(t) - a(t)| \sim \log V \tag{7.2.64}$$

for all *finite* t.

This result implies the lowest order result of van Kampen. For, we make the the substitution of the form (7.2.30)

$$a(t) = V\phi(t) + V^{1/2}z(t) \tag{7.2.65}$$

$$b(t) = V\phi(t) + V^{1/2}y(t)\,. \tag{7.2.66}$$

Then the characteristic function of $z(t)$ is

$$\begin{aligned}\langle \exp[isz(t)]\rangle &= \langle \exp[isV^{-1/2}a(t) - isV^{1/2}\phi(t)]\rangle \\ &= \exp[-isV^{1/2}\phi(t)]\langle \exp[isV^{-1/2}b(t)]\rangle + O(V^{-1/2}\log V) \\ &= \langle \exp[isy(t)]\rangle + O(V^{-1/2}\log V)\,. \end{aligned} \tag{7.2.67}$$

Using now the asymptotic expansion for the FPE we know the distribution function of $y(t)$ approaches that of the FPE (7.2.35) to $O(V^{-1/2})$ and the result follows with, however, a slightly weaker convergence because of the log V term involved. Thus, in terms of quantities which can be calculated and measured, means, variances, etc, Kurtz's apparently stronger result is equivalent to van Kampen's system size expansion.

7.2.5 Critical Fluctuations

The existence of a system size expansion as outlined in Sect.7.2.3 depends on the fact that $\tilde{\alpha}_1'(a)$ does not vanish. It is possible, however, for situations to arise where

$$\tilde{\alpha}_1'(\phi_s) = 0\,, \tag{7.2.68}$$

where ϕ_s is a stationary solution of the deterministic equation. This occurs, for example, when we consider the reaction of Sect.7.1.3 for which (using the notation of that section)

$$\tilde{\alpha}_1(y) = (\tilde{B}y^2 + \tilde{p} - y^3 - y\tilde{R})\tilde{k}_2\,,$$

where $\tilde{k}_2 = V^2k_2$.

Two situations can occur, corresponding to A and B in Fig. 7.2. The situation A corresponds to an unstable stationary state—any perturbation to the left will eventually lead to C, but B is stable. Clearly the deterministic equation takes the form

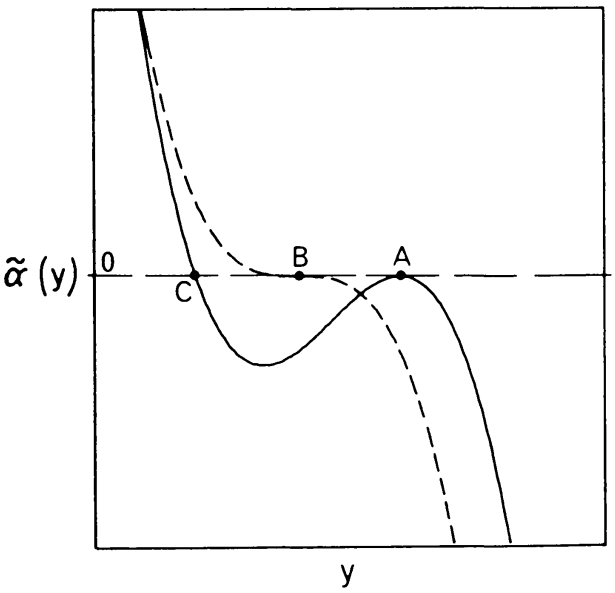


Fig. 7.2. Graph showing different kinds of behaviour of $\alpha_1(y)$ which lead to $\alpha_1'(y) = 0$

$$\dot{y} = -\tilde{k}_2(y - \phi_s)^3 \tag{7.2.69}$$

and we have a Master equation analogue of the cubic process of Sect.6.2.4a.

Van Kampen [7.13] has shown that in this case we should write

$$a = \Omega\,\phi(t) + \Omega^{\mu}u \tag{7.2.70}$$

in which case (7.2.32) becomes

$$\frac{\partial P(z,t)}{\partial t} - \Omega^{1-\mu}\phi'(t)\frac{\partial P(z,t)}{\partial z} = \sum_{n=1}^{\infty}\frac{\Omega^{1-\mu n}}{n!}\left(-\frac{\partial}{\partial z}\right)^n \tilde{\alpha}_n[\phi(t) + \Omega^{\mu-1}z]P(z,t). \tag{7.2.71}$$

Suppose now that the first $q-1$ derivatives of $\tilde{\alpha}_1(\phi_s)$ vanish. Then if we choose ϕ_s for $\phi(t)$, (7.2.71) becomes to lowest order

$$\frac{\partial P(z,t)}{\partial t} = -\frac{1}{q!}\tilde{\alpha}_1^{(q)}(\phi_s)\Omega^{(1-q)(1-\mu)}\frac{\partial}{\partial z}(z^q p) + \tfrac{1}{2}\alpha_2(\phi_s)\Omega^{1-2\mu}\frac{\partial^2 p}{\partial z^2} + \text{higher terms.} \tag{7.2.72}$$

To make sure z remains of order unity, set

$$(1-q)(1-\mu) = (1-2\mu), \text{ i.e., } \mu = \frac{q}{q+1} \tag{7.2.73}$$

so the result is

$$\frac{\partial P}{\partial t} = \Omega^{-(q-1)/(q+1)}\left(-\frac{1}{q!}\tilde{\alpha}_1^{(q)}\frac{\partial}{\partial z}z^q p + \tfrac{1}{2}\tilde{\alpha}_2\frac{\partial^2 P}{\partial z^2}\right) \tag{7.2.74}$$

(where $\tilde{\alpha}_1^{(q)}$ and $\tilde{\alpha}_2$ are evaluated at ϕ_s.) The fluctuations now vary on a slower time scale τ given by

$$\tau = t\Omega^{-(q-1)/(q+1)} \tag{7.2.75}$$

and the equation for the average is

$$\frac{d\langle x\rangle}{d\tau} = \frac{1}{q!}\tilde{\alpha}_1^{(q)}\langle x^q\rangle \tag{7.2.76}$$

which is no longer that associated with the linearised deterministic equation. Of course, stability depends on the sign of $\tilde{\alpha}_1^{(q)}$ and whether q is odd or even. The simplest stable case occurs for $q = 3$ which occurs at the critical point B of Fig.7.2, and in this case we have the cubic process of Sect.6.3.4a. The long-time scale is

$$\tau = \Omega^{-1/2}t\,. \tag{7.2.77}$$

We see that for large Ω, the system's time dependence is given as a function of $\tau = \Omega^{-1/2}t$. Only for times $t \gtrsim \Omega^{1/2}$ does τ become a significant size, and thus it is only for very long times t that any significant time development of the system takes place. Thus, the motion of the system becomes *very slow* at large Ω.

The condition (7.2.68) is normally controllable by some external parameter, (say, for example, the temperature), and the point in the parameter space where (7.2.68) is satisfied is called a *critical point*. This property of very slow time development at a critical point is known as *critical slowing down*.

7.3 Boundary Conditions for Birth-Death Processes

For birth-death processes, we have a rather simple way of implementing boundary conditions. For a process confined within an interval $[a, b]$, it is clear that reflecting and absorbing boundary conditions are obtained by forbidding the exit from the interval or the return to it, respectively. Namely,

	Reflecting	Absorbing
Boundary at a	$t^-(a) = 0$	$t^+(a-1) = 0$
Boundary at b	$t^+(b) = 0$	$t^-(b+1) = 0$.

(7.3.1)

It is sometimes useful, however, to insert boundaries in a process and, rather than set certain transition probabilities equal to zero, impose boundary conditions similar to those used for Fokker-Planck equations (Sect.5.2.1) so that the resulting solution in the interval $[a, b]$ is a solution of the Master equation with the appropriate vanishing transition probabilities. This may be desired in order to preserve the particular analytic form of the transition probabilities, which may have a certain convenience.

a) Forward Master Equation

We can write the forward Master equation as

$$\begin{aligned}\partial_t P(x, t\,|\,x', t') = {} & t^+(x-1)P(x-1, t\,|\,x', t') + t^-(x+1)P(x+1, t\,|\,x', t') \\ & - [t^+(x) + t^-(x)]P(x, t\,|\,x', t')\,.\end{aligned} \tag{7.3.2}$$

Suppose we want a reflecting barrier at $x = a$. Then this could be obtained by requiring

$$t^-(a) = 0$$

and

$$P(a-1, t|x', t) = 0\,. \tag{7.3.3}$$

The only equation affected by this requirement is that for $\partial_t P(a, t|x', t')$ for which the same equation can be obtained by not setting $t^-(a) = 0$ but instead introducing a *fictitious* $P(a-1, t|x', t')$ such that

$$t^+(a-1)P(a-1, t|x', t') = t^-(a)P(a, t|x', t')\,. \tag{7.3.4}$$

This can be viewed as the analogue of the zero current requirement for a reflecting barrier in a Fokker-Planck equation.

If we want an absorbing barrier at $x = a$, we can set

$$t^+(a-1) = 0\,. \tag{7.3.5}$$

After reaching the point $a-1$, the process never returns and its behaviour is now of no interest. The only equation affected by this is that for $\partial_t P(a, t|x', t')$ and the same equation can be again obtained by introducing a *fictitious* $P(a-1, t|x', t')$ such that

$$P(a-1, t|x', t') = 0\,. \tag{7.3.6}$$

Summarising, we have the alternative formulation of imposed boundary conditions which yield the same effect in $[a, b]$ as (7.3.1):

Foward Master Equation on interval $[a, b]$		
	Reflecting	Absorbing
Boundary at a	$t^-(a)P(a) = t^+(a-1)P(a-1)$	$P(a-1) = 0$
Boundary at b	$t^+(b)P(b) = t^-(b+1)P(b+1)$	$P(b+1) = 0$

(7.3.7)

b) Backward Master Equation

The backward Master equation is (see Sect. 3.6)

$$\begin{aligned}\partial_{t'}P(x, t|x', t') = {} & t^+(x')[P(x, t|x'+1, t') - P(x, t|x', t')] \\ & + t^-(x')[P(x, t|x'-1, t') - P(x, t|x', t')]\,.\end{aligned} \tag{7.3.8}$$

In the case of a reflecting barrier, at $x = a$, it is clear that $t^-(a) = 0$ is equivalent to constructing a fictitious $P(x, t|a-1, t')$ such that

$$P(x, t|a-1, t') = P(x, t|a, t') \tag{7.3.9}$$

In the absorbing barrier case, none of the equations for $P(x, t|x', t')$ with $x,x' \in [a, b]$ involve $t^+(a-1)$. However, because $t^+(a-1) = 0$, the equations in which $x' < a - 1$ will clearly preserve the condition

$$P(x, t|x', t') = 0, \qquad x \in [a, b], \qquad x' \leqslant a - 1 \tag{7.3.10}$$

and the effect of this on the equation with $x' = a$ will be to impose

$$P(x, t|a - 1, t') = 0 \tag{7.3.11}$$

which is therefore the required boundary condition. Summarising:

Backward Master Equation on interval $[a, b]$		
	Reflecting	Absorbing
Boundary at a	$P(\cdot\|a-1) = P(\cdot\|a)$	$P(\cdot\|a-1) = 0$
Boundary at b	$P(\cdot\|b+1) = P(\cdot\|b)$	$P(\cdot\|b+1) = 0$

(7.3.12)

7.4 Mean First Passage Times

The method for calculating these in the simple one-step case parallels that of the Fokker-Planck equation (Sect.5.2.7) very closely. We assume the system is confined to the range

$$a \leqslant x \leqslant b \tag{7.4.1}$$

and is abosrbed or reflected at either end, as the case may be. For definiteness we take a system with

reflecting barrier at $x = a$;

absorbing state at $x = b + 1$.

The argument is essentially the same as that in Sect.5.2.7 and we find that $T(x)$, the mean time for a particle initially at x to be absorbed, satisfies the equation related to the *backward Master equation* (7.3.8):

$$t^+(x)[T(x+1) - T(x)] + t^-(x)[T(x-1) - T(x)] = -1 \tag{7.4.2}$$

with the boundary condition corresponding to (5.2.159) and arising from (7.3.12):

$$T(a - 1) = T(a) \tag{7.4.3a}$$

$$T(b + 1) = 0\,. \tag{7.4.3b}$$

Define

$$U(x) = T(x+1) - T(x) \tag{7.4.4}$$

so (7.4.2) becomes

$$t^+(x)U(x) - t^-(x)U(x-1) = 1\,. \tag{7.4.5}$$

Define

$$\phi(x) = \prod_{z=a+1}^{x} \frac{t^-(z)}{t^+(z)} \qquad \text{and} \tag{7.4.6}$$

$$S(x) = U(x)/\phi(x) \tag{7.4.7}$$

then (7.4.5) is equivalent to

$$t^+(x)\phi(x)[S(x) - S(x-1)] = -1 \tag{7.4.8}$$

with a solution

$$S(x) = -\sum_{z=a}^{x} 1/[t^+(z)\phi(z)]\,. \tag{7.4.9}$$

This satisfies the boundary condition (7.4.3a) which implies that

$$U(a-1) = S(a-1) = 0\,. \tag{7.4.10}$$

Hence,

$$T(x+1) - T(x) = -\phi(x)\sum_{z=a}^{x} 1/[t^+(z)\phi(z)] \tag{7.4.11}$$

and

$$\boxed{T(x) = \sum_{y=x}^{b} \phi(y) \sum_{z=a}^{y} 1/[t^+(z)\phi(z)]} \qquad \begin{array}{l} a \text{ reflecting} \\ b \text{ absorbing} \\ b > a \end{array} \tag{7.4.12}$$

which also satisfies the boundary condition $T(b+1) = 0$, (7.4.3b).

Similarly, if a is absorbing and b reflecting

$$\boxed{T(x) = \sum_{y=a}^{x} \phi(y) \sum_{z=y}^{b} 1/[t^+(z)\phi(z)]} \qquad \begin{array}{l} a \text{ absorbing} \\ b \text{ reflecting} \\ b > a \end{array} \tag{7.4.13}$$

and a formula corresponding to (5.2.158) for both a and b absorbing can be similarly deduced.

7.4.1 Probability of Absorption

The mean time to absorption is always finite when a and b are finite. If, however, b is at infinity and is reflecting, the mean time may diverge. This does not itself mean that there is a finite probability of not being absorbed. The precise result [Ref. 7.7, Sect.4.7] is the following.

If the process takes place on the interval (a, ∞) and a is absorbing, then the probability of absorption into state $a-1$ from state x is given as follows. Define the function $M(x)$ by

$$M(x) = \sum_{y=a}^{\infty}\left[\prod_{z=a}^{y}\frac{t^+(y)}{t^-(y)}\right]. \tag{7.4.14}$$

Then if $M(x) < \infty$, the probability of absorption at $a-1$, from state x, is

$$\frac{M(x)}{1+M(x)} \tag{7.4.15}$$

and if $M(x) = \infty$, this probability is one. If this probability is 1, then the mean time to absorption is (7.4.13).

7.4.2 Comparison with Fokker-Planck Equation

The formulae (7.4.12, 13) are really very similar to the corresponding formulae (7.4.1, 2) for a diffusion process. In fact, using the model of Sect. 7.2.1c it is not difficult to show that in the limit $\delta \to 0$ the two become the same.

If we wish to deal with the kind of problem related to escape over a potential barrier (Sect.5.2.7c) which turn up in the context of this kind of master equation, for example, in the bistable reaction discussed in Sect.7.1.3, very similar approximations can be made. In this example, let us consider the mean first passage time from the stable stationary state x_1 to the other stable stationary state x_3.

Then the point $x = 0$ is a reflecting barrier, so the interval under consideration is $(0, x_3)$ with initial point x_1. Notice that

$$\phi(x) = \prod_{z=1}^{x}\frac{t^-(z)}{t^+(z)} = \frac{P_s(0)}{P_s(x)}\frac{t^+(0)}{t^+(x)} \tag{7.4.16}$$

so that

$$T(x_1 \to x_3) = \sum_{y=x_1}^{x_3}[P_s(y)t^+(y)]^{-1}\sum_{z=0}^{y}P_s(z)\,. \tag{7.4.17}$$

If we assume that $P_s(y)^{-1}$ has a *sharp* maximum at the unstable point x_2, we can set $y = x_2$ in all other factors in (7.4.17) to obtain

$$T(x_1 \to x_3) \sim \frac{n_1}{t^+(x_2)}\sum_{y=x_1}^{x_3}[P_s(y)]^{-1}\,, \tag{7.4.18}$$

$$n_1 = \sum_{z=0}^{x_2} P_s(z) \tag{7.4.19}$$

and is the total probability of being in the lower peak of the stationary distribution. The result is a discrete analogue of those obtained in Sect. 5.2.7c.

7.5 Birth-Death Systems with Many Variables

There is a very wide class of systems whose time development can be considered as the result of individual encounters between members of some population. These include, for example,

—chemical reactions, which arise by transformations of molecules on collision;
—population systems, which die, give birth, mate and consume each other;
—systems of epidemics, in which diseases are transmitted from individual to individual by contact.

All of these can usually be modelled by what I call *combinatorial kinetics*, in which the transition probability for a certain transformation consequent on that encounter is proportional to the number of possible encounters of that type.

For example, in a chemical reaction $X \rightleftharpoons 2Y$, the reaction $X \to 2Y$ occurs by spontaneous decay, a degenerate kind of encounter, involving only one individual. The number of encounters of this kind is the number of X; hence, we say

$$t(x \to x - 1, y \to y + 2) = k_1 x \,. \tag{7.5.1}$$

For the reverse reaction, one can assemble pairs of molecules of Y in $y(y-1)/2$ different ways. Hence

$$t(x \to x + 1, y \to y - 2) = k_2 y(y - 1) \,. \tag{7.5.2}$$

In general, we can consider encounters of many kinds between molecules, species, etc., of many kinds. Using the language of chemical reactions, we have the general formulation as follows.

Consider an n-component reacting system involving s different reactions:

$$\sum_a N_a^A X_a \underset{k_A^-}{\overset{k_A^+}{\rightleftharpoons}} \sum_a M_a^A X_a \qquad (A = 1, 2, \ldots, s) \,. \tag{7.5.3}$$

The coefficient N_a^A of X_a is the number of molecules of X_a involved on the left and M_a^A is the number involved on the right. We introduce a vector notation so that if x_a is the number of molecules of X_a, then

$$\begin{aligned} \boldsymbol{x} &= (x_1, x_2, \ldots, x_n) \\ \boldsymbol{N}^A &= (N_1^A, N_2^A, \ldots, N_n^A) \\ \boldsymbol{M}^A &= (M_1^A, M_2^A, \ldots, M_n^A) \end{aligned} \tag{7.5.4}$$

and we also define

$$\boldsymbol{r}^A = \boldsymbol{M}^A - \boldsymbol{N}^A . \tag{7.5.5}$$

Clearly, as reaction A proceeds one step in the forward direction,

$$\boldsymbol{x} \to \boldsymbol{x} + \boldsymbol{r}^A \tag{7.5.6}$$

and in the backward direction,

$$\boldsymbol{x} \to \boldsymbol{x} - \boldsymbol{r}^A . \tag{7.5.7}$$

The rate constants are defined by

$$\begin{aligned} t_A^+(\boldsymbol{x}) &= k_A^+ \prod_a \frac{x_a!}{(x_a - N_a^A)!} \\ t_A^-(\boldsymbol{x}) &= k_A^- \prod_a \frac{x_a!}{(x_a - M_a^A)!} \end{aligned} \tag{7.5.8}$$

which are proportional, respectively, to the number of ways of choosing the combination $\boldsymbol{N}^A$ or $\boldsymbol{M}^A$ from $\boldsymbol{x}$ molecules. The Master equation is thus

$$\begin{aligned} \partial_t P(\boldsymbol{x}, t) = \sum_A \{ & [t_A^-(\boldsymbol{x} + \boldsymbol{r}^A) P(\boldsymbol{x} + \boldsymbol{r}^A, t) - t_A^+(\boldsymbol{x}) P(\boldsymbol{x}, t)] \\ & + [t_A^+(\boldsymbol{x} - \boldsymbol{r}^A) P(\boldsymbol{x} - \boldsymbol{r}^A, t) - t_A^-(\boldsymbol{x}) P(\boldsymbol{x}, t)] \} . \end{aligned} \tag{7.5.9}$$

This form is, of course, a completely general way of writing a time-homogeneous Master equation for an integer variable $\boldsymbol{x}$ in which steps of size $\boldsymbol{r}^A$ can occur. It is only by making the special choice (7.5.8) for the transition probabilities per unit time that the general *combinatorial Master equation* arises. Another name is the *chemical Master equation*, since such equations are particularly adapted to chemical reactions.

7.5.1 Stationary Solutions when Detailed Balance Holds

In general, there is no explicit way of writing the stationary solution in a practical form. However, if detailed balance is satisfied, the stationary solution is easily derived. The variable $\boldsymbol{x}$, being simply a vector of numbers, can only be an even variable, hence, detailed balance must take the form (from Sect. 5.3.5)

$$t_A^-(\boldsymbol{x} + \boldsymbol{r}^A) P_s(\boldsymbol{x} + \boldsymbol{r}^A) = t_A^+(\boldsymbol{x}) P_s(\boldsymbol{x}) \tag{7.5.10}$$

for all A. The requirement that this holds for all A puts quite stringent requirements on the $t_A^\pm$. This arises from the fact that (7.5.10) provides a way of calculating $P_s(\boldsymbol{x}_0 + n\boldsymbol{r}^A)$ for any n and any initial $\boldsymbol{x}_0$. Using this method for all available A, we can generate $P_s(\boldsymbol{x})$ on the space of all $\boldsymbol{x}$ which can be written

$$\boldsymbol{x} = \boldsymbol{x}_0 + \sum n_A \boldsymbol{r}^A ; \qquad (n_A \text{ integral}) \tag{7.5.11}$$

but the solutions obtained may be ambiguous since, for example, from (7.5.10) we may write

$$P_s[(\boldsymbol{x}+\boldsymbol{r}^A)+\boldsymbol{r}^B] = \frac{P_s(\boldsymbol{x})\, t_A^+(\boldsymbol{x})\, t_B^+(\boldsymbol{x}+\boldsymbol{r}^A)}{t_A^-(\boldsymbol{x}+\boldsymbol{r}^A)\, t_B^-(\boldsymbol{x}+\boldsymbol{r}^A+\boldsymbol{r}^B)}$$

but

$$P_s[(\boldsymbol{x}+\boldsymbol{r}^B)+\boldsymbol{r}^A] = \frac{P_s(\boldsymbol{x})\, t_B^+(\boldsymbol{x})\, t_A^+(\boldsymbol{x}+\boldsymbol{r}^B)}{t_B^-(\boldsymbol{x}+\boldsymbol{r}^B)\, t_A^-(\boldsymbol{x}+\boldsymbol{r}^A+\boldsymbol{r}^B)} \qquad (7.5.12)$$

Using the combinatorial forms (7.5.8) and substituting in (7.5.12), we find that this condition is automatically satisfied.

The condition becomes nontrivial when the same two points can be connected to each other in two essentially different ways, i.e., if, for example,

$$\begin{aligned} \boldsymbol{N}^A &\neq \boldsymbol{N}^B \\ \boldsymbol{M}^A &\neq \boldsymbol{M}^B \end{aligned} \qquad (7.5.13)$$

but $\boldsymbol{r}^A = \boldsymbol{r}^B \equiv \boldsymbol{r}$.

In this case, uniqueness of $P_s(\boldsymbol{x}+\boldsymbol{r}^A)$ in (7.5.10) requires

$$\frac{t_A^+(\boldsymbol{x})}{t_A^-(\boldsymbol{x}+\boldsymbol{r})} = \frac{t_B^+(\boldsymbol{x})}{t_B^-(\boldsymbol{x}+\boldsymbol{r})} \qquad (7.5.14)$$

and this means

$$\frac{k_A^+}{k_A^-} = \frac{k_B^+}{k_B^-}. \qquad (7.5.15)$$

If there are two chains $A, B, C, \ldots$, and $A', B', C', \ldots$, of reactions such that

$$\boldsymbol{r}^A + \boldsymbol{r}^B + \boldsymbol{r}^C + \ldots = \boldsymbol{r}^{A'} + \boldsymbol{r}^{B'} + \boldsymbol{r}^{C'} + \ldots . \qquad (7.5.16)$$

Direct substitution shows that

$$P_s(\boldsymbol{x}+\boldsymbol{r}^A+\boldsymbol{r}^B+\boldsymbol{r}^C+\ldots) = P_s(\boldsymbol{x}+\boldsymbol{r}^{A'}+\boldsymbol{r}^{B'}+\boldsymbol{r}^{C'}+\ldots) \qquad (7.5.17)$$

only if

$$\boxed{\frac{k_A^+}{k_A^-}\frac{k_B^+}{k_B^-}\frac{k_C^+}{k_C^-}\ldots = \frac{k_{A'}^+}{k_{A'}^-}\frac{k_{B'}^+}{k_{B'}^-}\frac{k_{C'}^+}{k_{C'}^-}\ldots} \qquad (7.5.18)$$

which is, therefore, the condition for detailed balance in a Master equation with combinatorial kinetics.

A solution for $P_s(\boldsymbol{x})$ in this case is a multivariate Poisson

$$P_s(\boldsymbol{x}) = \prod_a \frac{\alpha_a^{x_a}\, \mathrm{e}^{-\alpha_a}}{x_a!} \qquad (7.5.19)$$

which we check by substituting into (7.5.10) which gives

$$\prod_a \frac{\alpha_a^{(x_a+r_a^A)}e^{-\alpha_a}}{(x_a+r_a^A)!}\frac{k_A^-(x_a+r_a^A)!}{(x_a+r_a^A-M_a^A)!} = \prod_a \frac{\alpha_a^{x_a}e^{-\alpha_a}}{x_a!}\frac{k_A^+ x_a!}{(x_a-N_a^A)!}. \tag{7.5.20}$$

Using the fact that

$$r_a^A = M_a^A - N_a^A,$$

we find that

$$k_A^+ \prod_a \alpha_a{}^{N_a^A} = k_A^- \prod_a \alpha_a{}^{M_a^A}. \tag{7.5.21}$$

However, the most general solution will have this form only subject to conservation laws of various kinds. For example, in the reaction

$$X \rightleftharpoons 2Y, \tag{7.5.22}$$

the quantity $2x + y$ is conserved. Thus, the stationary distribution is

$$\frac{e^{-\alpha_1}\alpha_1^x}{x!}\frac{e^{-\alpha_2}\alpha_2^y}{y!}\phi(2x+y) \tag{7.5.23}$$

where ϕ is an arbitrary function. Choosing $\phi(2x + y) = 1$ gives the Poissonian solution. Another choice is

$$\phi(2x+y) = \delta(2x+y, N) \tag{7.5.24}$$

which corresponds to fixing the total of $2x + y$ at N.

As a degenerate form of this, one sometimes considers a reaction written as

$$A \rightleftharpoons 2Y \tag{7.5.25}$$

in which, however, A is considered a fixed, deterministic number and the possible reactions are

$$\begin{aligned} &A \to 2Y: \quad t^+(y) = k_1 a \\ &2Y \to A \quad t^-(y) = k_2 y(y-1). \end{aligned} \tag{7.5.26}$$

In this case, the conservation law is now simply that y is always even, or always odd. The stationary solution is of the form

$$P_s(y) = \frac{\alpha^y}{y!}\psi(y, \alpha) \tag{7.5.27}$$

where $\psi(y, \alpha)$ is a function which depends on y only through the evenness or oddness of y.

7.5.2 Stationary Solutions Without Detailed Balance (Kirchoff's Solution)

There is a method which, in principle, determines stationary solutions in general, though it does not seem to have found great practical use. The interested reader is referred to *Haken* [Ref. 7.8, Sect.4.8] and *Schnakenberg* [7.4] for a detailed treatment. In general, however, approximation methods have more to offer.

7.5.3 System Size Expansion and Related Expansions

In general we find that in chemical Master equations a system size expansion does exist. The rate of production or absorption is expected to be proportional to Ω, the size of the system. This means that as $\Omega \to \infty$, we expect

$$\boldsymbol{x} \sim \Omega \boldsymbol{\rho} \,, \tag{7.5.28}$$

where $\boldsymbol{\rho}$ is the set of chemical concentrations. Thus, we must have $t_A^{\pm}(\boldsymbol{x})$ proportional to Ω as $\Omega \to \infty$, so that this requires

$$\boxed{\begin{aligned} k_A^+ &\sim \kappa_A^+ \, \Omega^{-\sum_a N_a^A + 1} \\ k_A^- &\sim \kappa_A^- \, \Omega^{-\sum_a M_a^A + 1} \,. \end{aligned}} \tag{7.5.29}$$

Under these circumstances, a multivariate form of van Kampen's system size expansion can be developed. This is so complicated that it will not be explicitly derived here, but as in the single variable case, we have a Kramers-Moyal expansion whose first two terms give a diffusion process whose asymptotic form is the same as that arising from a system size expansion.

The Kramers-Moyal expansion from (7.5.9) can be derived in exactly the same way as in Sect.7.2.2, in fact, rather more easily, since (7.5.9) is already in the appropriate form. Thus, we have

$$\partial_t P(\boldsymbol{x}, t) = \sum_{A,n} \left\{ \frac{(\boldsymbol{r}^A \cdot \nabla)^n}{n!} [t_A^-(\boldsymbol{x}) P(\boldsymbol{x}, t)] + \frac{(-\boldsymbol{r}^A \cdot \nabla)^n}{n!} [t_A^+(\boldsymbol{x}) P(\boldsymbol{x}, t)] \right\} \tag{7.5.30}$$

and we now truncate this to second order to obtain

$$\boxed{\begin{aligned} &\partial_t P(\boldsymbol{x}, t) = -\sum_a \partial_a [A_a(\boldsymbol{x}) P(\boldsymbol{x}, t)] + \tfrac{1}{2} \sum_{a,b} \partial_a \partial_b [B_{ab}(\boldsymbol{x}) P(\boldsymbol{x}, t)] \,, && (7.5.31) \\ &\left.\begin{aligned} A_a(\boldsymbol{x}) &= \sum_A r_a^A [t_A^+(\boldsymbol{x}) - t_A^-(\boldsymbol{x})] \\ B_{ab}(\boldsymbol{x}) &= \sum_A r_a^A r_b^A [t_A^+(\boldsymbol{x}) + t_A^-(\boldsymbol{x})] \,. \end{aligned}\right\} && (7.5.32) \end{aligned}}$$

In this form we have the *chemical Fokker-Planck equation* corresponding to the Master equation. However, we note that this is really only valid as an approxima-

tion, whose large volume asymptotic expansion is identical to order $1/\Omega$ with that of the corresponding Master equation.

If one is satisfied with this degree of approximation, it is often simpler to use the Fokker-Planck equation than the Master equation.

7.6 Some Examples

7.6.1 $X + A \rightleftharpoons 2X$

Here,

$$\begin{aligned} t^+(x) &= k_1 a x \\ t^-(x) &= k_2 x(x-1)\ . \end{aligned} \tag{7.6.1}$$

Hence,

$$\begin{aligned} A(x) &= k_1 a x - k_2 x(x-1) \sim k_1 a x - k_2 x^2 \quad \text{to order } 1/\Omega \\ B(x) &= k_1 a x + k_2 x(x-1) \sim k_1 a x + k_2 x^2 \quad \text{to order } 1/\Omega\ . \end{aligned} \tag{7.6.2}$$

7.6.2 $X \underset{k}{\overset{\gamma}{\rightleftharpoons}} Y \underset{\gamma}{\overset{k}{\rightleftharpoons}} A$

Here we have

$$\left.\begin{aligned} t_1^+(\boldsymbol{x}) &= \gamma y \\ t_1^-(\boldsymbol{x}) &= k x \end{aligned}\right\} \qquad \boldsymbol{r}^1 = (1, -1) \tag{7.6.3}$$

$$\left.\begin{aligned} t_2^+(\boldsymbol{x}) &= k a \\ t_2^-(\boldsymbol{x}) &= \gamma y \end{aligned}\right\} \qquad \boldsymbol{r}^2 = (0, 1)\ . \tag{7.6.4}$$

Hence,

$$\boldsymbol{A}(\boldsymbol{x}) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} (\gamma y - k x) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} (k a - \gamma y) \tag{7.6.5}$$

$$= \begin{bmatrix} \gamma y - k x \\ k x + k a - 2\gamma y \end{bmatrix} \tag{7.6.6}$$

$$\underset{\sim}{\boldsymbol{B}}(\boldsymbol{x}) = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} (1, -1) \right\} (\gamma y + k x) + \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} (0, 1) \right\} (k a + \gamma y) \tag{7.6.7}$$

$$= \begin{bmatrix} \gamma y + k x & -\gamma y - k x \\ -\gamma y - k x & 2\gamma y + k x + k a \end{bmatrix}. \tag{7.6.8}$$

If we now use the linearised form about the stationary state,

$$\gamma y = kx = ka \tag{7.6.9}$$

$$\underset{\sim}{B}_s = \begin{bmatrix} 2ka & -2ka \\ -2ka & 4ka \end{bmatrix}. \tag{7.6.10}$$

7.6.3 Prey-Predator System

The prey-predator system of Sect. 1.3 provides a good example of the kind of system in which we are interested. As a chemical reaction we can write it as

$$\begin{aligned} &\text{i)} \quad X + A \to 2X \qquad \boldsymbol{r}^1 = (1, 0) \\ &\text{ii)} \quad X + Y \to 2Y \qquad \boldsymbol{r}^2 = (-1, 1) \\ &\text{iii)} \quad Y \to B \qquad \boldsymbol{r}^3 = (0, -1)\,. \end{aligned} \tag{7.6.11}$$

The reactions are all irreversible (though reversibility may be introduced) so we have

$$t_A^-(\boldsymbol{x}) = 0 \qquad (A = 1, 2, 3)$$

but

$$\begin{aligned} t_1^+(\boldsymbol{x}) &= k_1 a \frac{x!}{(x-1)!} \frac{y!}{y!} = k_1 a x \\ t_2^+(\boldsymbol{x}) &= k_2 \frac{x!}{(x-1)!} \frac{y!}{(y-1)!} = k_2 xy \\ t_3^+(\boldsymbol{x}) &= k_3 \frac{x!}{x!} \frac{y!}{(y-1)!} = k_3 y\,. \end{aligned} \tag{7.6.12}$$

The Master equation can now be explicitly written out using (7.5.9): one obtains

$$\begin{aligned} \partial_t P(x, y) = {} & k_1 a(x-1) P(x-1, y) + k_2(x+1)(y-1) P(x+1, y-1) \\ & + k_3(y+1)\, P(x, y+1) - (k_1 a x + k_2 xy + k_3 y)\, P(x, y). \end{aligned} \tag{7.6.13}$$

There are no exact solutions of this equation, so approximation methods must be used.

Kramers-Moyal. From (7.5.32)

$$\boldsymbol{A}(\boldsymbol{x}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} k_1 a x + \begin{bmatrix} -1 \\ 1 \end{bmatrix} k_2 xy + \begin{bmatrix} 0 \\ -1 \end{bmatrix} k_3 y \tag{7.6.14}$$

$$= \begin{bmatrix} k_1 a x - k_2 xy \\ k_2 xy - k_3 y \end{bmatrix}, \tag{7.6.15}$$

$$\underset{\sim}{B}(x) = \begin{bmatrix}1\\0\end{bmatrix}(1, 0)k_1 ax + \begin{bmatrix}-1\\1\end{bmatrix}(-1, 1)k_2 xy + \begin{bmatrix}0\\-1\end{bmatrix}(0, -1)k_3 y \tag{7.6.16}$$

$$= \begin{bmatrix}k_1 ax + k_2 xy & -k_2 xy\\ -k_2 xy & k_2 xy + k_3 y\end{bmatrix}. \tag{7.6.17}$$

The deterministic equations are

$$\frac{d}{dt}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}k_1 ax - k_2 xy\\ k_2 xy - k_3 y\end{bmatrix}. \tag{7.6.18}$$

Stationary State at

$$\begin{bmatrix}x_s\\y_s\end{bmatrix} = \begin{bmatrix}k_3/k_2\\ k_1 a/k_2\end{bmatrix}. \tag{7.6.19}$$

To determine the stability of this state, we check the stability of the iinearised deterministic equation

$$\frac{d}{dt}\begin{bmatrix}\delta x\\ \delta y\end{bmatrix} = \frac{\partial A(x_s)}{\partial x_s}\delta x + \frac{\partial A(x_s)}{\partial y_s}\delta y$$

$$= \begin{bmatrix}k_1 a - k_2 y_s\\ k_2 y_s\end{bmatrix}\delta x + \begin{bmatrix}-k_2 x_s\\ k_2 x_s - k_3\end{bmatrix}\delta y \tag{7.6.20}$$

$$= \begin{bmatrix}0 & -k_3\\ k_1 a & 0\end{bmatrix}\begin{bmatrix}\delta x\\ \delta y\end{bmatrix}. \tag{7.6.21}$$

The eigenvalues of the matrix are

$$\lambda = \pm\, \mathrm{i}(k_1 k_3 a)^{1/2} \tag{7.6.22}$$

which indicates a periodic motion of any small deviation from the stationary state. We thus have neutral stability, since the disturbance neither grows nor decays.

This is related to the existence of a conserved quantity

$$V = k_2(x + y) - k_3 \log x - k_1 a \log y \tag{7.6.23}$$

which can readily be checked to satisfy $dV/dt = 0$. Thus, the system conserves V and this means that there are different circular trajectories of constant V.

Writing again

$$x = x_s + \delta x \tag{7.6.24}$$

$$y = y_s + \delta y$$

and expanding to second order, we see that

$$V = \frac{k_2^2}{2}\left(\frac{\delta x^2}{k_3^2} + \frac{\delta y^2}{(k_1 a)^2}\right) \tag{7.6.25}$$

so that the orbits are initially elliptical (this can also be deduced from the linearised analysis).

As the orbits become larger, they become less elliptic and eventually either x or y may become zero.

If x is the first to become zero (all the prey have been eaten), one sees that y inevitably proceeds to zero as well. If y becomes zero (all predators have starved to death), the prey grow unchecked with exponential growth.

Stochastic Behaviour. Because of the conservation of the quantity V, the orbits have neutral stability which means that when the fluctuations are included, the system will tend to change the size of the orbit with time. We can see this directly from the equivalent stochastic differential equations

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} k_1 ax - k_2 xy \\ k_2 xy - k_3 y \end{bmatrix} dt + \underset{\sim}{C}(x, y) \begin{bmatrix} dW_1(t) \\ dW_2(t) \end{bmatrix}, \tag{7.6.26}$$

where

$$\underset{\sim}{C}(x, y)\underset{\sim}{C}(x, y)^{\mathrm{T}} = \underset{\sim}{B}(x) . \tag{7.6.27}$$

Then using Ito's formula

$$dV(x, y) = \frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy + \frac{1}{2}\left(\frac{\partial^2 V}{\partial x^2} dx^2 + 2\frac{\partial^2 V}{\partial x \partial y} dx\, dy + \frac{\partial^2 V}{\partial y^2} dy\right) \tag{7.6.28}$$

so that

$$\begin{aligned} \langle dV(x, y)\rangle = &\left\langle \frac{\partial V}{\partial x}(k_1 ax - k_2 xy) + \frac{\partial V}{\partial y}(k_2 xy - k_3 y)\right\rangle dt \\ &+ \left\langle B_{11}\frac{k_3}{2x^2} + B_{22}\frac{k_1}{2y^2}\right\rangle dt . \end{aligned} \tag{7.6.29}$$

The first average vanishes since V is deterministically conserved and we find

$$\langle dV(x, y)\rangle = \frac{1}{2}\left\langle \frac{k_3 k_1 a}{x} + \frac{k_3 k_2 y}{x} + \frac{k_1 k_2 ax}{y} + \frac{k_1 k_3 a}{y}\right\rangle . \tag{7.6.30}$$

All of these terms are of order Ω^{-1} and are positive when x and y are positive. Thus, in the mean, $V(x, y)$ increases steadily. Of course, eventually one or other of the axes is hit and similar effects occur to the deterministic case. We see that when x or y vanish, $V = \infty$.

Direct implementation of the system size expansion is very cumbersome in this case, and moment equations prove more useful. These can be derived directly from the Master equation or from the Fokker-Planck equation. The results differ slightly

from each other, by terms of order inverse volume. For simplicity, we use the FPE so that

$$\frac{d}{dt}\begin{bmatrix}\langle x\rangle\\ \langle y\rangle\end{bmatrix}=\begin{bmatrix}k_1a\langle x\rangle - k_2\langle xy\rangle\\ k_2\langle xy\rangle - k_3\langle y\rangle\end{bmatrix} \tag{7.6.31}$$

$$\frac{d}{dt}\begin{bmatrix}\langle x^2\rangle\\ \langle xy\rangle\\ \langle y^2\rangle\end{bmatrix}=\frac{1}{dt}\begin{bmatrix}\langle 2x\,dx + dx^2\rangle\\ \langle x\,dy + y\,dx + dx\,dy\rangle\\ \langle 2y\,dy + dy^2\rangle\end{bmatrix} \tag{7.6.32}$$

$$=\begin{bmatrix}2k_1a\langle x^2\rangle - 2k_2\langle x^2y\rangle + k_1a\langle x\rangle + k_2\langle xy\rangle\\ k_2\langle x^2y - y^2x\rangle + (k_1a - k_3 - k_2)\langle xy\rangle\\ 2k_2\langle xy^2\rangle - 2k_3\langle y^2\rangle + k_2\langle xy\rangle + k_3\langle y\rangle\end{bmatrix}. \tag{7.6.33}$$

Knowing a system size expansion is valid means that we know all correlations and variances are of order $1/\Omega$ compared with the means.

We therefore write

$$\begin{aligned} x &= \langle x\rangle + \delta x \\ y &= \langle y\rangle + \delta y \end{aligned} \tag{7.6.34}$$

and keep terms only of lowest order. Noting that terms arising from $\langle dx^2\rangle$, $\langle dxdy\rangle$ and $\langle dy^2\rangle$ are one order in Ω smaller than the others, we get

$$\frac{d}{dt}\begin{bmatrix}\langle x\rangle\\ \langle y\rangle\end{bmatrix}=\begin{bmatrix}k_1a\langle x\rangle - k_2\langle x\rangle\langle y\rangle\\ k_2\langle x\rangle\langle y\rangle - k_3\langle x\rangle\langle y\rangle\end{bmatrix}+\begin{bmatrix}-k_2\langle\delta x\delta y\rangle\\ k_2\langle\delta x\delta y\rangle\end{bmatrix} \tag{7.6.35}$$

$$\begin{aligned}\frac{d}{dt}\begin{bmatrix}\langle\delta x^2\rangle\\ \langle\delta x\delta y\rangle\\ \langle\delta y^2\rangle\end{bmatrix}&=\begin{bmatrix}k_1a\langle x\rangle + k_2\langle x\rangle\langle y\rangle\\ -k_2\langle x\rangle\langle y\rangle\\ k_2\langle x\rangle\langle y\rangle + k_3\langle y\rangle\end{bmatrix}\\ &+\begin{bmatrix}2k_1a - 2k_2\langle y\rangle, & -2k_2\langle x\rangle, & 0\\ k_2\langle y\rangle, & k_1a - k_3 + k_2(\langle x\rangle - \langle y\rangle), & -k_2\langle x\rangle\\ 0, & 2k_2\langle y\rangle, & 2k_2\langle x\rangle - 2k_3\end{bmatrix}\begin{bmatrix}\langle\delta x^2\rangle\\ \langle\delta x\delta y\rangle\\ \langle\delta y^2\rangle\end{bmatrix}\end{aligned} \tag{7.6.36}$$

We note that the means, to lowest order, obey the deterministic equations, but to next order, the term containing $\langle\delta x\delta y\rangle$ will contribute. Thus, let us choose a simplified case in which

$$k_1a = k_3 = 1, \quad k_2 = \alpha \tag{7.6.37}$$

which can always be done by rescaling variables. Also abbreviating

$$\langle x\rangle \to x,\ \langle y\rangle \to y,\ \langle\delta x^2\rangle \to f,\ \langle\delta x\delta y\rangle \to g,\ \langle\delta y^2\rangle \to h\,, \tag{7.6.38}$$

we obtain

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - \alpha xy \\ \alpha xy - y \end{bmatrix} + \begin{bmatrix} -\alpha g \\ \alpha g \end{bmatrix}$$

$$\frac{d}{dt}\begin{bmatrix} f \\ g \\ h \end{bmatrix} = \begin{bmatrix} x + \alpha xy \\ -\alpha xy \\ \alpha xy + y \end{bmatrix} + \begin{bmatrix} 2 - 2\alpha y & -2\alpha x & 0 \\ \alpha y & \alpha(x - y) & -\alpha x \\ 0 & 2\alpha y & 2\alpha x - 2 \end{bmatrix}\begin{bmatrix} f \\ g \\ h \end{bmatrix}. \tag{7.6.39}$$

We can attempt to solve these equations in a stationary state. Bearing in mind that f, g, h, are a factor Ω^{-1} smaller than x and y, this requires α to be of order Ω^{-1} [this also follows from the scaling requirements (7.5.29)]. Hence α is small. To lowest order one has

$$x_s = y_s = 1/\alpha\,. \tag{7.6.40}$$

But the equations for f, g, h in the stationary state then become

$$\begin{bmatrix} 2g \\ h - f \\ -2g \end{bmatrix} = \begin{bmatrix} 2/\alpha \\ -1/\alpha \\ 2/\alpha \end{bmatrix} \tag{7.6.41}$$

which are inconsistent. Thus this method does not yield a stationary state. Alternatively one can solve all of (7.6.39) in a stationary state.

After some manipulation one finds

$$\begin{aligned} x_s &= y_s \\ g_s &= \alpha^{-1}(x_s - \alpha x_s^2) \end{aligned} \tag{7.6.42}$$

so that

$$\begin{aligned} f_s &= x_s(-2\alpha x_s^2 + x_s(2 - \alpha) - 1)\,/\,(2 - 2\alpha x_s) \\ h_s &= x_s(-2\alpha x_s^2 + x_s(2 + \alpha) + 1)/(2 - 2\alpha x_s) \end{aligned} \tag{7.6.43}$$

and the equation for g_s gives

$$-\alpha x_s^2 + \alpha x_s(f_s - h_s) = 0 \tag{7.6.44}$$

giving a solution for x_s, y_s, etc.

$$\left.\begin{aligned} x_s &= y_s = \tfrac{1}{2} \\ f_s &= \tfrac{1}{2}\alpha/(\alpha - 2) \\ g_s &= \tfrac{1}{4}(2 - \alpha)/\alpha \\ h_s &= -1\,/\,(\alpha - 2) \end{aligned}\right\} \tag{7.6.45}$$

and for small α in which the method is valid, this leads to a negative value for f_s, which is by definition, postive. Thus there is no stationary solution.

By again approximating $x_s = y_s = 1/\alpha$, the differential equations for f, g, and h can easily be solved. We find, on assuming that initially the system has zero variances and correlations,

$$\begin{aligned} f(t) &= \frac{1}{2\alpha}(\cos 2t - 1) + \frac{2t}{\alpha} \\ g(t) &= -\frac{1}{2\alpha}\sin 2t \\ h(t) &= -\frac{1}{2\alpha}(\cos 2t - 1) + \frac{2t}{\alpha}. \end{aligned} \tag{7.6.46}$$

Notice that $f(t)$ and $h(t)$ are, in fact, always positive and increase steadily. The solution is valid only for a short time since the increasing value of $g(t)$ will eventually generate a time-dependent mean.

7.6.4 Generating Function Equations

In the case of combinatorial kinetics, a relatively simple differential equation can be derived for the generating function:

$$G(\boldsymbol{s}, t) = \sum_{x} \left(\prod_{a} s_a^{x_a}\right) P(\boldsymbol{x}, t). \tag{7.6.47}$$

For we note that

$$\partial_t G(\boldsymbol{s}, t) = \partial_t^+ G(\boldsymbol{s}, t) + \partial_t^- G(\boldsymbol{s}, t) \tag{7.6.48}$$

where the two terms correspond to the t^+ and t^- parts of the master equation. Thus

$$\begin{aligned} \partial_t^+ G(\boldsymbol{s}, t) = \sum_{A, x} k_A^+ \Bigg\{ &\prod_a \left[\frac{(x_a - r_a^A)!}{(x_a - r_a^A - N_a^A)!} s_a^{x_a}\right] P(\boldsymbol{x} - \boldsymbol{r}^A, t) \\ &- \prod_a \left[\frac{x_a!}{(x_a - N_a^A)!} s_a^{x_a}\right] P(\boldsymbol{x}, t)\Bigg\}. \end{aligned} \tag{7.6.49}$$

Changing the summation variable to $\boldsymbol{x} - \boldsymbol{r}^A$ and renaming this as $\boldsymbol{x}$, in the first term we find

$$\partial_t^+ G(\boldsymbol{s}, t) = \sum_{A, x} k_A^+ \left[\prod_a \frac{x_a!}{(x_a - N_a^A)!} s_a^{x_a + r_a^A} - \prod_a \frac{x_a!}{(x_a - N_a^A)!} s_a^{x_a}\right] P(\boldsymbol{x}, t). \tag{7.6.50}$$

Note that

$$\prod_a \frac{s_a^{x_a} x_a!}{(x_a - N_a^A)!} = \prod_a \left(\partial_a^{N_a^A} s_a^{x_a}\right) s_a^{N_a^A} \tag{7.6.51}$$

and that

$$\prod_a \frac{s_a^{x_a + r_a^A x_a!}}{(x_a - N_a^A)!} = \prod_a \left(\partial_a^{N_a^A} s_a^{x_a}\right) s_a^{M_a^A} \tag{7.6.52}$$

so that

$$\partial_t^+ G(\boldsymbol{s}, t) = \sum_A k_A^+ \left(\prod_a s_a^{M_a^A} - \prod_a s_a^{N_a^A}\right) \partial_a^{N_a^A} G(\boldsymbol{s}, t) . \tag{7.6.53}$$

Similarly, we derive a formula for $\partial_t^- G(\boldsymbol{s}, t)$ and put these together to get

$$\boxed{\partial_t G(\boldsymbol{s}, t) = \sum_A \left(\prod_a s_a^{M_a^A} - \prod_a s_a^{N_a^A}\right) \left(k_A^+ \prod_a \partial_a^{N_a^A} - k_A^- \prod_a \partial_a^{M_a^A}\right) G(\boldsymbol{s}, t)} \tag{7.6.54}$$

which is the general formula for a generating function differential equation. We now give a few examples.

a) An Exactly Soluble Model

Reactions: (A, B, C held fixed)

$$A + X \xrightarrow{k_2} 2X + D \qquad \begin{aligned} &N^1 = 1,\ M^1 = 2;\ r^1 = 1 \\ &k_1^+ = k_2 A \\ &k_1^- = 0 \end{aligned} \tag{7.6.55}$$

$$B + X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} C \qquad \begin{aligned} &N^2 = 1,\ M^2 = 0;\ r^2 = -1 \\ &k_2^+ = k_1 B \\ &k_2^- = k_3 C . \end{aligned} \tag{7.6.56}$$

Hence, from (7.6.54), the generating function equation is

$$\partial_t G = (s^2 - s)(k_2 A \partial_s G) + (1 - s)(k_1 B \partial_s G - k_3 C G) . \tag{7.6.57}$$

Solve by characteristics.

Set

$$k_1 B = \beta, \qquad k_2 A = \alpha, \qquad k_3 C = \gamma . \tag{7.6.58}$$

The characteristics are

$$\frac{dt}{1} = -\frac{ds}{(1 - s)(\beta - \alpha s)} = \frac{dG}{\gamma(1 - s)G} \tag{7.6.59}$$

. . .

$$\left(\frac{1-s}{\beta-\alpha s}\right)e^{(\alpha-\beta)t}=u \tag{7.6.60}$$

$$(\beta-\alpha s)^{\gamma/\alpha}G=v\,. \tag{7.6.61}$$

The general solution can be written $v=F(u)$, i.e.,

$$G=(\beta-\alpha s)^{-\gamma/\alpha}F\left\{e^{(\alpha-\beta)t}\left(\frac{1-s}{\beta-\alpha s}\right)\right\}\,. \tag{7.6.62}$$

From this we can find various time-dependent solutions. The conditional probability $P(x,t\,|\,y,0)$ comes from the initial condition

$$G_y(s,0)=s^y \tag{7.6.63}$$

$$\Rightarrow F(z)=(1-\beta z)^y(1-\alpha z)^{-\gamma/\alpha-y}(\beta-\alpha)^{\gamma/\alpha} \tag{7.6.64}$$

$$\boxed{\begin{aligned}\Rightarrow G_y(s,t)&=\lambda^{\gamma/\alpha}[\beta(1-e^{-\lambda t})-s(\alpha-\beta e^{-\lambda t})]^y\\&\quad\times[(\beta-\alpha e^{-\lambda t})-\alpha s(1-e^{-\lambda t})]^{-\gamma/\alpha-y}\end{aligned}} \tag{7.6.65}$$

(with $\lambda=\beta-\alpha$).

As $t\to\infty$, a stationary state exists only if $\beta>\alpha$ and is

$$G_y(s,\infty)=(\beta-\alpha s)^{-\gamma/\alpha}(\beta-\alpha)^{\gamma/\alpha} \tag{7.6.66}$$

$$\Rightarrow P_s(x)=\frac{\Gamma(x+\gamma/\alpha)(\alpha/\beta)^x}{\Gamma(\gamma/\alpha)x!}(\beta-\alpha)^{\gamma/\alpha}\,. \tag{7.6.67}$$

We can also derive moment equations from the generating function equations by noting

$$\begin{aligned}\partial_s G(s,t)\,|_{s=1}&=\langle x(t)\rangle\\ \partial_s^2 G(s,t)\,|_{s=1}&=\langle x(t)[x(t)-1]\rangle\,.\end{aligned} \tag{7.6.68}$$

Proceeding this way we have

$$\frac{d}{dt}\langle x(t)\rangle=(k_2A-k_1B)\langle x(t)\rangle+k_3C \tag{7.6.69}$$

and

$$\begin{aligned}\frac{d}{dt}\langle x(t)[x(t)-1]\rangle&=2(k_2A-k_1B)\langle x(t)[x(t)-1]\rangle\\&\quad+2k_2A\langle x(t)\rangle+2k_3C\langle x(t)\rangle\,.\end{aligned} \tag{7.6.70}$$

These equations have a stable stationary solution provided

$$k_2A < k_1B, \qquad \text{i.e.,} \quad \alpha < \beta .$$

In this case, the stationary mean and variance are

$$\langle x\rangle_s = k_3C/(k_1B - k_2A) \tag{7.6.71}$$

$$\text{var}\{x\}_s = k_1k_3BC/(k_2A - k_1B)^2 . \tag{7.6.72}$$

This model is a simple representation of the processes taking place in a nuclear reactor. Here X is a neutron. The first reaction represents the fission by absorption of a neutron by A to produce residue(s) D plus two neutrons. The second represents absorption of neutrons and production by means other than fission.

As k_2A approaches k_1B, we approach a critical situation where neutrons are absorbed and created in almost equal numbers. For $k_2A > k_1B$, an explosive chain reaction occurs. Notice that $\langle x_s\rangle$ and $\text{var}\{x\}_s$ both become very large as a critical point is approached and, in fact,

$$\frac{\text{var}\{x_s\}}{\langle x\rangle_s} = \frac{k_1B}{k_1B - k_2A} \to \infty . \tag{7.6.73}$$

Thus, there are *very* large fluctuations in $\langle x_s\rangle$ near the critical point.

Note also that the system has linear equations for the mean and is Markovian, so the methods of Sect. 3.7.4 (the regression theorem) show that

$$\langle x(t), x(0)\rangle_s = \exp[(k_2A - k_1B)t]\text{var}\{x\}_s \tag{7.6.74}$$

so that the fluctuations become vanishingly slow as the critical point is approached, i.e., the time correlation function decays very slowly with time.

b) Chemical Reaction $X_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} X_2$

One reaction

$$N = \begin{bmatrix}1\\0\end{bmatrix}, \qquad M = \begin{bmatrix}0\\1\end{bmatrix}, \qquad r = \begin{bmatrix}-1\\1\end{bmatrix}$$

$$k_1^+ = k_1, \qquad k_1^- = k_2 \tag{7.6.75}$$

$$\partial_t G(s_1, s_2, t) = (s_2 - s_1)(k_1\partial_{s_1} - k_2\partial_{s_2})G(s_1, s_2, t)$$

can be solved by characteristics. The generating function is an arbitrary function of solutions of

$$\frac{dt}{1} = -\frac{ds_1}{k_1(s_2 - s_1)} = \frac{ds_2}{k_2(s_2 - s_1)} . \tag{7.6.76}$$

Two integrals are solutions of

$$k_2ds_1 + k_1ds_2 = 0 \Longrightarrow k_2s_1 + k_1s_2 = v , \tag{7.6.77}$$

$$(k_1 + k_2)dt = \frac{d(s_2 - s_1)}{s_2 - s_1} \Longrightarrow (s_2 - s_1)e^{-(k_1+k_2)t} = u \tag{7.6.78}$$

$$\therefore\ G(s_1, s_2, t) = F[k_2 s_1 + k_1 s_2, (s_2 - s_1)e^{-(k_1+k_2)t}]\,. \tag{7.6.79}$$

The initial condition (Poissonian)

$$G(s_1, s_2, 0) = \exp[\alpha(s_1 - 1) + \beta(s_2 - 1)] \tag{7.6.80}$$

gives the Poissonian solution:

$$G(s_1, s_2, t) = \exp\left\{\frac{k_2\beta - k_1\alpha}{k_1 + k_2}(s_2 - s_1)e^{-(k_1+k_2)t} + \frac{\alpha + \beta}{k_1 + k_2}[k_1(s_2 - 1) + k_2(s_1 - 1)]\right\}. \tag{7.6.81}$$

In this case, the stationary solution is not unique because $x + y$ is a conserved quantity. From (7.6.79) we see that the general stationary solution is of the form

$$G(s_1, s_2, \infty) = F(k_2 s_1 + k_1 s_2, 0)\,. \tag{7.6.82}$$

Thus,

$$k_1^n \frac{\partial^n G}{\partial s_1^n} = k_2^n \frac{\partial^n G}{\partial s_2^n} \tag{7.6.83}$$

which implies that, setting $s_1 = s_2 = 1$,

$$k_1^n \langle x_1^n \rangle_f = k_2^n \langle x_2^n \rangle_f\,. \tag{7.6.84}$$

7.7 The Poisson Representation [7.10]

This is a particularly elegant technique which generates Fokker-Planck equations which are equivalent to chemical Master equations of the form (7.5.9).

We *assume* that we can expand $P(\boldsymbol{x}, t)$ as a superposition of multivariate uncorrelated Poissons:

$$P(\boldsymbol{x}, t) = \int d\boldsymbol{\alpha} \prod_a \frac{e^{-\alpha_a}\alpha_a^{x_a}}{x_a!} f(\boldsymbol{\alpha}, t)\,. \tag{7.7.1}$$

This means that the generating function $G(\boldsymbol{s}, t)$ can be written

$$G(\boldsymbol{s}, t) = \int d\boldsymbol{\alpha}\ \exp\left[\sum_a (s_a - 1)\alpha_a\right] f(\boldsymbol{\alpha}, t)\,. \tag{7.7.2}$$

We substitute this in the generating function equation (7.6.54) to get

$$\partial_t G(\boldsymbol{s}, t) = \sum_A \int d\boldsymbol{\alpha} \left\{ \left[\prod_a \left(\frac{\partial}{\partial \alpha_a} + 1 \right)^{M_a^A} - \prod_a \left(\frac{\partial}{\partial \alpha_a} + 1 \right)^{N_a^A} \right] \times \left(k_A^+ \prod_a \alpha_a^{N_a^A} - k_A^- \prod_a \alpha_a^{M_a^A} \right) \exp \left[\sum_a (s_a - 1)\alpha_a \right] \right\} f(\boldsymbol{\alpha}, t) . \tag{7.7.3}$$

We now integrate by parts, drop surface terms and finally equate coefficients of the exponential to obtain

$$\boxed{\frac{\partial f(\boldsymbol{\alpha}, t)}{\partial t} = \sum_A \left[\prod_a \left(1 - \frac{\partial}{\partial \alpha_a} \right)^{M_a^A} - \prod_a \left(1 - \frac{\partial}{\partial \alpha_a} \right)^{N_a^A} \right] \times \left[k_A^+ \prod_a \alpha_a^{N_a^A} - k_A^- \prod_a \alpha_a^{M_a^A} \right] f(\boldsymbol{\alpha}, t) .} \tag{7.7.4}$$

a) Fokker-Planck Equations for Bimolecular Reaction Systems

This equation is of the Fokker-Planck form if we have, as is usual in real chemical reactions,

$$\sum_a M_a^A \leqslant 2 \qquad \sum_a N_a^A \leqslant 2 \tag{7.7.5}$$

which indicates that only pairs of molecules at the most participate in reactions. The FPE can then be simplified as follows. Define the currents

$$J_A(\boldsymbol{\alpha}) = k_A^+ \prod_a \alpha_a^{N_a^A} - k_A^- \prod_a \alpha_a^{M_a^A}, \tag{7.7.6}$$

the drifts

$$A_a[\boldsymbol{J}(\boldsymbol{\alpha})] = \sum_A r_a^A J_A(\boldsymbol{\alpha}) , \tag{7.7.7}$$

and the diffusion matrix elements by

$$B_{ab}[\boldsymbol{J}(\boldsymbol{\alpha})] = \sum_A J_A(\boldsymbol{\alpha})(M_a^A M_b^A - N_a^A N_b^A - \delta_{a,b} r_a^A) . \tag{7.7.8}$$

Then the Poisson representation FPE is

$$\boxed{\frac{\partial f(\boldsymbol{\alpha}, t)}{\partial t} = - \sum_a \frac{\partial}{\partial \alpha_a} \{A_a[\boldsymbol{J}(\boldsymbol{\alpha})] f(\boldsymbol{\alpha}, t)\} + \tfrac{1}{2} \sum_{a,b} \frac{\partial^2}{\partial \alpha_a \partial \alpha_b} \{B_{ab}[\boldsymbol{J}(\boldsymbol{\alpha})] f(\boldsymbol{\alpha}, t)\} .} \tag{7.7.9}$$

Notice also that if we use the explicit volume dependence of the parameters given in Sect. 7.5.3 (7.5.29) and define

$$\eta_a = \alpha_a/V \tag{7.7.10}$$

$$\varepsilon = V^{-1/2} \tag{7.7.11}$$

and $F(\boldsymbol{\eta}, t)$ is the quasiprobability in the $\boldsymbol{\eta}$ variable, then the FPE for the $\boldsymbol{\eta}$ variable takes the form of

$$\frac{\partial F(\boldsymbol{\eta}, t)}{\partial t} = -\sum_a \frac{\partial}{\partial \eta_a}[\hat{A}_a(\boldsymbol{\eta})F(\boldsymbol{\eta}, t)] + \frac{\varepsilon^2}{2}\sum_{a,b}\frac{\partial^2}{\partial\eta_a\partial\eta_b}[\hat{B}_{a,b}(\boldsymbol{\eta})F(\boldsymbol{\eta}, t)] \tag{7.7.12}$$

with

$$\hat{A}_a(\boldsymbol{\eta}) = \sum_A r_a^A \hat{J}_A(\boldsymbol{\eta}) \tag{7.7.13a}$$

$$\hat{J}_A(\boldsymbol{\eta}) = \kappa_A^+ \prod_a \eta_a^{N_a^A} - \kappa_A^- \prod_a \eta_a^{M_a^A} \tag{7.7.13b}$$

$$\hat{B}_{ab}(\boldsymbol{\eta}) = \sum_A \hat{J}_A(\boldsymbol{\eta})(M_a^A M_b^A - N_a^A N_b^A - \delta_{a,b} r_a^A)\,. \tag{7.7.13c}$$

In this form we see how the system size expansion in $V^{-1/2}$ corresponds exactly to a small noise expansion in $\boldsymbol{\eta}$ of the FPE (7.7.12). For such birth-death Master equations, this method is technically much simpler than a direct system size expansion.

b) Unimolecular Reactions

If for all A,

$$\sum_a M_a^A \leqslant 1 \quad \text{and} \quad \sum_a N_a^A \leqslant 1,$$

then it is easily checked that the diffusion coefficient $\hat{B}_{ab}(\boldsymbol{\eta})$ in (7.7.13) vanishes, and we have a Liouville equation. An initially Poissonian $P(\boldsymbol{x}, t_0)$, corresponds to a delta function $F(\boldsymbol{\eta}, t_0)$, and the time evolution generated by this Liouville equation will generate a delta function solution, $\delta(\boldsymbol{\eta} - \bar{\boldsymbol{\eta}}(t))$, where $\bar{\boldsymbol{\eta}}(t)$ is the solution of

$$d\boldsymbol{\eta}/dt = \boldsymbol{A}(\boldsymbol{\eta})$$

This means that $P(\boldsymbol{x}, t)$ will preserve a Poissonian form, with mean equal to $\bar{\boldsymbol{\eta}}(t)$. Thus we derive the general result, that there exist propagating multipoissonian solutions for any unimolecular reaction system. Non Poissonian solutions also exist—these correspond to initial $F(\boldsymbol{\eta}, t_0)$ which are not delta functions.

c) Example

As an example, consider the reaction pair

$$\begin{aligned} &\text{i)}\quad A + X \underset{k_4}{\overset{k_2}{\rightleftharpoons}} 2X \\ &\text{ii)}\quad B + X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} C \end{aligned} \tag{7.7.14}$$

$$N^1 = 1, \quad M^1 = 2, \quad k_1^+ = k_2 A, \quad k_1^- = k_4$$
$$N^2 = 1, \quad M^2 = 0, \quad k_2^+ = k_1 B, \quad k_2^- = k_3 C$$

so that (7.7.4) takes the form

$$\frac{\partial f}{\partial t} = \left[\left(1 - \frac{\partial}{\partial \alpha}\right)^2 - \left(1 - \frac{\partial}{\partial \alpha}\right)\right](k_2 A\alpha - k_4\alpha^2) f + \left[1 - \left(1 - \frac{\partial}{\partial \alpha}\right)\right](k_1 B\alpha - k_3 C) f \tag{7.7.15}$$

$$\frac{\partial f}{\partial t} = \left\{-\frac{\partial}{\partial \alpha}[k_3 C + (k_2 A - k_1 B)\alpha - k_4\alpha^2] + \frac{\partial^2}{\partial \alpha^2}[k_2 A\alpha - k_4\alpha^2]\right\} f \tag{7.7.16}$$

which is of the Fokker-Planck form, provided $k_2 A\alpha - k_4\alpha^2 > 0$. Furthermore, there is the simple relationship between moments, which takes the form (in the case of one variable)

$$\boxed{\begin{aligned} \langle x^r \rangle_f &= \sum_x \int d\alpha\, [x(x-1)\ldots(x-r+1)\frac{e^{-\alpha}\alpha^x}{x!} f(\alpha) \\ &= \int d\alpha\, \alpha^r f(\alpha) \equiv \langle \alpha^r \rangle . \end{aligned}} \tag{7.7.17}$$

This follows from the factorial moments of the Poisson distribution (Sect. 2.8.3). However, $f(\alpha)$ is not a probability or at least, is not guaranteed to be a probability in the simple minded way it is defined. This is clear, since any positive superposition of Poisson distributions must have a variance at least as wide as the Poisson distribution. Hence any $P(x)$ for which the variance is less than that of the Poisson distribution i.e., cannot be represented by a positive $f(\alpha)$.

A representation in terms of *distributions* is always possible, at least formally. For if we define

$$f_y(\alpha) = (-1)^y \delta^y(\alpha) e^\alpha, \qquad \text{then} \tag{7.7.18}$$

$$\int d\alpha\, f_y(\alpha) e^{-\alpha}\alpha^x/x! = \int d\alpha\, \alpha^x \left(-\frac{d}{d\alpha}\right)^y \delta(\alpha)/x! \tag{7.7.19}$$

and integrating by parts

$$= \delta_{x,y} \tag{7.7.20}$$

which means that we can write

$$P(x) = \int d\alpha (e^{-\alpha}\alpha^x/x!)[\sum_y (-1)^y P(y)\delta^y(\alpha) e^\alpha] \tag{7.7.21}$$

so that in a *formal* sense, an $f(\alpha)$ can always be found for any $P(x)$.

The rather singular form just given does not, in fact, normally arise since, for example, we can find the stationary solution of the FPE (7.7.16) as the potential solution (up to a normalisation)

$$f_s(\alpha) = e^{\alpha}(k_2A - k_4\alpha)^{(k_1B/k_4 - k_3C/k_2A - 1)}\alpha^{k_3C/k_2A - 1} \tag{7.7.22}$$

which is a relatively smooth function. However, an interpretation as a probability is only possible if $f_s(\alpha)$ is positive or zero and is normalisable.

If we define

$$\delta = (k_1B/k_4 - k_3C/k_2A), \tag{7.7.23}$$

then $f_s(\alpha)$ is normalisable on the interval $(0, k_2A/k_4)$ provided that

$$\begin{aligned} &\delta > 0 \\ &k_3 > 0. \end{aligned} \tag{7.7.24}$$

Clearly, by definition, k_3 must be positive.

It must further be checked that the integrations by parts used to derive the FPE (7.7.4) are such that under these conditions, surface terms vanish. For an interval (a, b) the surface terms which would arise in the case of the reaction (7.7.14) can be written

$$\begin{aligned} &[\{(k_2A\alpha - k_4\alpha^2 - k_1B\alpha + k_3C)f - \partial_\alpha[(k_2\alpha - k_4\alpha^2)f]\}\{e^{(s-1)\alpha}\}]_b^a \\ &\quad + [(k_2\alpha - k_4\alpha^2)f[(s-1)e^{(s-1)\alpha}]]_b^a . \end{aligned} \tag{7.7.25}$$

Because of the extra factor $(s - 1)$ on the second line, each line must vanish separately. It is easily checked that on the interval $(0, k_2A/k_4)$, each term vanishes at each end of the interval for the choice (7.7.22) of f, provided δ and k_3 are both greater than zero.

In the case where k_3 and δ are both positive, we have a genuine FPE equivalent to the stochastic differential equation

$$d\alpha = [k_3C + (k_2A - k_1B)\alpha - k_4\alpha^2]dt + \sqrt{2(k_2A\alpha - k_4\alpha^2)}dW(t) . \tag{7.7.26}$$

The motion takes place on the range $(0, k_2A/k_4)$ and both boundaries satisfy the criteria for *entrance boundaries*, which means that it is not possible to leave the range $(0, k_2A/k_4)$ (Sect.5.2.1).

If either of the conditions (7.7.24) is violated, it is found that the drift vector is such as to take the point outside the interval $(0, k_2A/k_4)$. For example, near $\alpha = 0$ we have

$$d\alpha \sim k_3C\, dt \tag{7.7.27}$$

and if k_3C is negative, α will proceed to negative values. In this case, the coefficient of $dW(t)$ in (7.7.26) becomes imaginary and interpretation is no longer possible without further explanation.

Of course, viewed as a SDE in the *complex* variable

$$\alpha = \alpha_x + i\alpha_y , \tag{7.7.28}$$

the SDE is perfectly sensible and is really a pair of stochastic differential equations for the two variables α_x and α_y. However, the corresponding FPE is no longer the

one variable equation (7.7.16) but a two-variable FPE. We can derive such a FPE in terms of variations of the Poisson representation, which we now treat.

7.7.1 Kinds of Poisson Representations

Let us consider the case of one variable and write

$$P(x) = \int_{\mathscr{D}} d\mu(\alpha)(e^{-\alpha}\alpha^x/x!)f(\alpha) . \tag{7.7.29}$$

Then $\mu(\alpha)$ is a measure which we will show may be chosen in three ways which all lead to useful representations, and $\mathscr{D}$ is the domain of integration, which can take on various forms, depending on the choice of measure.

7.7.2 Real Poisson Representations

Here we choose

$$d\mu(\alpha) = d\alpha \tag{7.7.30}$$

and $\mathscr{D}$ is a section of the real line. As noted in the preceding example, this representation does not always exist, but where it does, a simple interpretation in terms of Fokker-Planck equations is possible.

7.7.3 Complex Poisson Representations

Here,

$$d\mu(\alpha) = d\alpha \tag{7.7.31}$$

and $\mathscr{D}$ is a contour C in the complex plane. We can show that this exists under certain restrictive conditions. For, instead of the form (7.7.18), we can choose

$$f_y(\alpha) = \frac{y!}{2\pi i}\alpha^{-y-1}e^{\alpha} \tag{7.7.32}$$

and C to be a contour surrounding the origin. This means that

$$P_y(x) = \frac{1}{2\pi i}\oint_C \frac{y!}{x!}\alpha^{x-y-1} = \delta_{x,y} . \tag{7.7.33}$$

By appropriate summation, we may express a given $P(x)$ in terms of an $f(\alpha)$ given by

$$f(\alpha) = \frac{1}{2\pi i}\sum_y P(y)e^{\alpha}\alpha^{-y-1}y! . \tag{7.7.34}$$

If the $P(y)$ are such that for all y, $y!P(y)$ is bounded, the series has a finite radius of convergence outside which $f(\alpha)$ is analytic. By choosing C to be outside this circle of convergence, we can take the integration inside the summation to find that $P(x)$ is given by

$$P(x) = \oint_C d\alpha(e^{-\alpha}\alpha^x/x!)f(\alpha)\,. \tag{7.7.35}$$

a) Example: Reactions (1) $A + X \rightleftharpoons 2X$, (2) $B + X \rightleftharpoons C$

We use the notation of Sect. 7.7 and distinguish three cases, depending on the magnitude of δ. The quantity δ gives a measure of the direction in which the reaction system (7.7.14) is proceeding when a steady state exists. If $\delta > 0$, we find that when x has its steady state value, reaction (1) is producing X while reaction (2) consumes X. When $\delta = 0$, both reactions balance separately—thus we have chemical equilibrium. When $\delta < 0$, reaction (1) consumes X while reaction (2) produces X.

i) $\delta > 0$: according to (7.7.24), this is the condition for $f_s(\alpha)$ to be a valid quasiprobability on the real interval $(0, k_2A/k_4)$. In this range, the diffusion coefficient $(k_2A\alpha - k_4\alpha^2)$ is positive. The deterministic mean of α, given by

$$\alpha = \frac{k_2A - k_1B + [(k_2A - k_1B)^2 + 4k_3k_4C]^{1/2}}{2k_4} \tag{7.7.36}$$

lies *within* the interval $(0, k_2A/k_4)$. We are therefore dealing with the case of a genuine FPE and $f_s(\alpha)$ is a function vanishing at both ends of the interval and peaked near the deterministic steady state.

ii) $\delta = 0$: since both reactions now balance separately, we expect a Poissonian steady state. We note that $f_s(\alpha)$ in this case has a pole at $\alpha = k_2A/k_4$ and we choose the range of α to be a contour in the complex plane enclosing this pole. Since this is a closed contour, there are no boundary terms arising from partial integration and $P_s(x)$ given by choosing this type of Poisson representation clearly satisfies the steady state Master equation. Now using the calculus of residues, we see that

$$P_s(x) = \frac{e^{-\alpha_0}\alpha_0^x}{x!} \tag{7.7.37}$$

with

$$\alpha_0 = k_2A/k_4\,.$$

iii) $\delta < 0$: when $\delta < 0$ we meet some very interesting features. The steady state solution (7.7.22) now no longer satisfies the condition $\delta > 0$. However, if the range of α is chosen to be a contour C in the complex plane (Fig. 7.3) and we employ the complex Poisson representation, $P_s(x)$ constructed as

$$P_s(x) = \int_C d\alpha f_s(\alpha)\frac{e^{-\alpha}\alpha^x}{x!} \tag{7.7.38}$$

is a solution of the Master equation. The deterministic steady state now occurs at a point on the real axis to the right of the singularity at $\alpha = k_2A/k_4$, and asymptotic evaluations of means, moments, etc., may be obtained by choosing C to pass through the saddle point that occurs there. In doing so, one finds that the variance of α, defined as

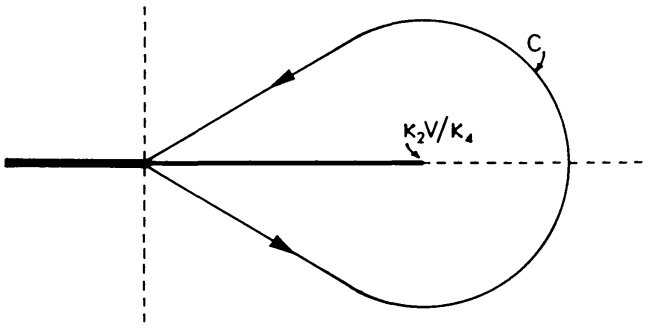


Fig. 7.3. Contour C in the complex plane for the evaluation of (7.7.38)

$$\text{var}\{\alpha\} = \langle\alpha^2\rangle - \langle\alpha\rangle^2 , \tag{7.7.39}$$

is negative, so that

$$\text{var}\{x\} = \langle x^2\rangle - \langle x\rangle^2 = \langle\alpha^2\rangle - \langle\alpha\rangle^2 + \langle\alpha\rangle < \langle x\rangle \tag{7.7.40}$$

This means that the steady state is narrower than the Poissonian. Finally, it should be noted that all three cases can be obtained from the contour C. In the case where $\delta = 0$, the cut from the singularity at $\alpha = k_2A/k_4$ to $-\infty$ vanishes and C may be distorted to a simple contour round the pole, while if $\delta > 0$, the singularity at $\alpha = k_2A/k_4$ is now integrable so the contour may be collapsed onto the cut and the integral evaluated as a discontinuity integral over the range $[0, k_2A/k_4]$. (When δ is a positive integer, this argument requires modification).

b) Example: Reactions $B \xrightarrow{k_1} X$, $2X \xrightarrow{k_2} A$
For which the Fokker-Planck equation is

$$\frac{\partial f(\alpha, t)}{\partial t} = -\frac{\partial}{\partial\alpha}[(\kappa_1 V - 2\kappa_2 V^{-1}\alpha^2)f(\alpha, t)] - \frac{\partial^2}{\partial\alpha^2}[(\kappa_2 V^{-1}\alpha^2)f(\alpha, t)] , \tag{7.7.41}$$

where $\kappa_1 V = k_1 B$, $\kappa_2 V^{-1} = k_2$ and V is the system volume. Note that the diffusion coefficient in the above FPE is negative on all the real lines.

The potential solution of (7.7.41) is (up to a normalisation factor)

$$f(\alpha) = \alpha^{-2} \exp(2\alpha + aV^2/\alpha) \tag{7.7.42}$$

with $a = 2\kappa_2/\kappa_1$ and the α integration is to be performed along a closed contour encircling the origin. Of course, in principle, there is another solution obtained by solving the stationary FPE in full. However, only the potential solution is single valued and allows us to choose an acceptable contour on which partial integration is permitted.

Thus, by putting $\alpha = \eta V$, we get

$$\langle x^r\rangle_f = \frac{V^r \oint d\eta\, e^{V(2\eta + a/\eta)}\eta^{r-2}}{\oint d\eta\, e^{V(2\eta + a/\eta)}\eta^{-2}} . \tag{7.7.43}$$

The function $(2\eta + a/\eta)$ does not have a maximum at the deterministic steady state. In fact, it has a minimum at the deterministic steady state $\eta = +(a/2)^{1/2}$.

However, in the complex η plane this point is a saddle point and provides the dominant contribution to the integral.

Thus, the negative diffusion coefficient in (7.7.41) reflects itself by giving rise to a saddle point at the deterministic steady state, which results in the variance in X being less than $\langle x \rangle$.

From (7.7.43) all the steady states moments can be calculated exactly. The results are

$$\langle x^r \rangle_f = \left[V \left(\frac{a}{2} \right)^{1/2} \right]^r \frac{I_{r-1}(2(2a)^{1/2}V)}{I_1(2(2a)^{1/2}V)}, \tag{7.7.44}$$

where $I_r(2(2a)^{1/2}V)$ are the modified Bessel functions. Using the large-argument expansion for $I_r(2(2a)^{1/2}V)$, we get

$$\begin{aligned} \langle x \rangle &= V(a/2)^{1/2} + \tfrac{1}{8} + O(1/V) \\ \text{var}\{x\} &= \tfrac{3}{4} V(a/2)^{1/2} - \tfrac{1}{16} + O(1/V) . \end{aligned} \tag{7.7.45}$$

These asymptotic results can also be obtained by directly applying the method of steepest descents to (7.7.43). In general, this kind of expansion will always be possible after explicitly exhibiting the volume dependence of the parameters.

c) Summary of Advantages

The complex Poisson representation yields stationary solutions in analytic form to which asymptotic or exact methods are easily applicable. It is not so useful in the case of time-dependent solutions. The greatest advantages, however, occur in quantum mechanical systems where similar techniques can be used for complex P representations which can give information that is not otherwise extractable. These are treated in Chap. 10.

7.7.4 The Positive Poisson Representation

Here we choose α to be a complex variable $\alpha_x + i\alpha_y$,

$$d\mu(\alpha) = d^2\alpha = d\alpha_x d\alpha_y , \tag{7.7.46}$$

and $\mathscr{D}$ is the whole complex plane. We show in Sect. 10.6.3 that *for any $P(x)$, a positive $f(\alpha)$ exists such that*

$$P(x) = \int d^2\alpha \, (e^{-\alpha}\alpha^x/x!) f(\alpha) ; \tag{7.7.47}$$

thus, the positive P representation always exists. It is not unique, however. For example, choose

$$f_p(\alpha) = (2\pi\sigma^2)^{-1} \exp(-|\alpha - \alpha_0|^2/2\sigma^2) \tag{7.7.48}$$

and note that if $g(\alpha)$ is any analytic function of α, we can write

$$g(\alpha) = g(\alpha_0) + \sum_{n=1}^{\infty} g^{(n)}(\alpha_0)(\alpha - \alpha_0)^n/n! \tag{7.7.49}$$

so that

$$\int (2\pi\sigma^2)^{-1} d^2\alpha \exp(-|\alpha - \alpha_0|^2/2\sigma^2) g(\alpha) = g(\alpha_0), \tag{7.7.50}$$

since the terms with $n \geqslant 1$ vanish when integrated in (7.7.50). Noting that the Poisson form $e^{-\alpha}\alpha^x/x!$ is itself analytic in α, we obtain for any positive value of σ^2

$$P(x) = \int d^2\alpha f_p(\alpha) e^{-\alpha}\alpha^x/x! = e^{-\alpha_0}\alpha_0^x/x! . \tag{7.7.51}$$

In practice, this nonuniqueness is an advantage rather than a problem.

a) Fokker-Planck Equations

We make use of the analyticity of the Poisson and its generating function to produce Fokker-Planck equations with positive diffusion matrices. A FPE of the form of (7.7.9) arises from a generating function equation

$$\partial_t G(\boldsymbol{s}, t) = \int d^2\boldsymbol{\alpha} f(\boldsymbol{\alpha}, t) \left(\sum_a A_a \frac{\partial}{\partial \alpha_a} + \sum_{a,b} \tfrac{1}{2} B_{ab} \frac{\partial^2}{\partial \alpha_a \partial \alpha_b} \right) \exp\left[\sum_a (s_a - 1)\alpha_a\right]. \tag{7.7.52}$$

We now take explicit account of the fact that $\boldsymbol{\alpha}$ is a complex variable

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_x + i\boldsymbol{\alpha}_y \tag{7.7.53}$$

and also write

$$\boldsymbol{A}(\boldsymbol{\alpha}) = \boldsymbol{A}_x(\boldsymbol{\alpha}) + i\boldsymbol{A}_y(\boldsymbol{\alpha}) . \tag{7.7.54}$$

We further write

$$\underset{\sim}{\boldsymbol{B}}(\boldsymbol{\alpha}) = \underset{\sim}{\boldsymbol{C}}(\boldsymbol{\alpha})\underset{\sim}{\boldsymbol{C}}^{\mathrm{T}}(\boldsymbol{\alpha}) \tag{7.7.55}$$

and

$$\underset{\sim}{\boldsymbol{C}}(\boldsymbol{\alpha}) = \underset{\sim}{\boldsymbol{C}}_x(\boldsymbol{\alpha}) + i\underset{\sim}{\boldsymbol{C}}_y(\boldsymbol{\alpha}) \tag{7.7.56}$$

For brevity we use

$$\begin{aligned} \partial_a &= \frac{\partial}{\partial \alpha_a} \\ \partial_a^x &= \frac{\partial}{\partial \alpha_{x,a}} \\ \partial_a^y &= \frac{\partial}{\partial \alpha_{y,a}} . \end{aligned} \tag{7.7.57}$$

Because of the analyticity of $\exp\left[\sum_a (s_a - 1)\alpha_a\right]$, in the generating function equation (7.7.52) we can always make the interchangeable choice

$$\partial_a \leftrightarrow \partial_a^x \leftrightarrow -i\partial_a^y . \tag{7.7.58}$$

We then substitute the form (7.7.54) for B_{ab}, and replace ∂_a by either ∂_a^x or $-i\partial_a^y$

according to whether the corresponding index on $\boldsymbol{A}$ or $\underset{\sim}{C}$ is x or y respectively. We then derive

$$\begin{aligned}\partial_t G(\boldsymbol{s}, t) = \int d^2\boldsymbol{\alpha} f(\boldsymbol{\alpha}, t)\{[\sum_a (A_{a;x}\partial_a^x + A_{a;y}\partial_a^y) \\ + \tfrac{1}{2}\sum_{a,b,c}(C_{a,c,x}C_{c,b;x}\partial_a^x\partial_b^x + C_{a,c;y}C_{c,b;y}\partial_a^y\partial_b^y \\ + 2C_{a,c;x}C_{c,b;y}\partial_a^x\partial_b^y]\exp[\sum_a (s_a - 1)\alpha_a]\}. \end{aligned} \tag{7.7.59}$$

Integrating by parts and discarding the surface terms to get a FPE in the variables $(\boldsymbol{\alpha}_x, \boldsymbol{\alpha}_y)$,

$$\begin{aligned}\partial_t f(\boldsymbol{\alpha}, t) = [-\sum_a (\partial_a^x A_{a;x} + \partial_a^y A_{a;y}) + \tfrac{1}{2}\sum_{a,b,c}(\partial_a^x\partial_b^x C_{a,c;x}C_{c,b;x} \\ + \partial_a^y\partial_b^y C_{a,c;y}C_{c,b;y} + 2\partial_a^x\partial_b^y C_{a,c;x}C_{c,b;y})]f(\boldsymbol{\alpha}, t). \end{aligned} \tag{7.7.60}$$

In the space of doubled dimensions, this is a FPE with positive semidefinite diffusion. For, we have for the variable $(\boldsymbol{\alpha}_x, \boldsymbol{\alpha}_y)$ the drift vector

$$\mathscr{A}(\boldsymbol{\alpha}) = [\boldsymbol{A}_x(\boldsymbol{\alpha}), \boldsymbol{A}_y(\boldsymbol{\alpha})] \tag{7.7.61}$$

and the diffusion matrix

$$\mathscr{B}(\boldsymbol{\alpha}) = \begin{bmatrix} \underset{\sim}{C}_x\underset{\sim}{C}_x^{\mathrm{T}} & \underset{\sim}{C}_x\underset{\sim}{C}_y^{\mathrm{T}} \\ \underset{\sim}{C}_y\underset{\sim}{C}_x^{\mathrm{T}} & \underset{\sim}{C}_y\underset{\sim}{C}_y^{\mathrm{T}} \end{bmatrix} = \mathscr{C}(\boldsymbol{\alpha})\mathscr{C}(\boldsymbol{\alpha})^{\mathrm{T}} \tag{7.7.62}$$

where

$$\mathscr{C}(\boldsymbol{\alpha}) = \begin{bmatrix} \underset{\sim}{C}_x & 0 \\ \underset{\sim}{C}_y & 0 \end{bmatrix} \tag{7.7.63}$$

so that $\mathscr{B}(\boldsymbol{\alpha})$ is explicitly positive semidefinite.

b) Stochastic Differential Equation (SDE)
Corresponding to the drift and diffusion (7.7.61, 62) we have a stochastic differential equation

$$\begin{bmatrix} d\boldsymbol{\alpha}_x \\ d\boldsymbol{\alpha}_y \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_x(\alpha) \\ \boldsymbol{A}_y(\alpha) \end{bmatrix} dt + \begin{bmatrix} \underset{\sim}{C}_x d\boldsymbol{W}(t) \\ \underset{\sim}{C}_y d\boldsymbol{W}(t) \end{bmatrix} \tag{7.7.64}$$

where $\boldsymbol{W}(t)$ is a Wiener process of the same dimension as $\boldsymbol{\alpha}_x$. Note that the same Wiener process occurs in both lines because of the two zero entries $\mathscr{C}(\boldsymbol{\alpha})$ as written in (7.7.63).

Recombining real and imaginary parts, we find the SDE for the complex variable $\boldsymbol{\alpha}$:

$$d\boldsymbol{\alpha} = \boldsymbol{A}(\boldsymbol{\alpha})dt + \underset{\sim}{C}(\boldsymbol{\alpha})d\boldsymbol{W}(t). \tag{7.7.65}$$

This is of course, exactly the same SDE which would arise if we used the usual rules for converting Fokker-Planck equations to stochastic differential equations directly on the Poisson representation FPE (7.7.9), and ignored the fact that $\underset{\sim}{C}(\boldsymbol{\alpha})$

so defined would have complex elements if $\underline{B}$ was not a positive semidefinite diffusion matrix.

c) Examples of Stochastic Differential Equations in the Complex Plane

We again consider the reactions (sect. 7.7.b)

$$\begin{aligned} A + X &\underset{k_4}{\overset{k_2}{\rightleftharpoons}} 2X \\ B + X &\underset{k_3}{\overset{k_1}{\rightleftharpoons}} C . \end{aligned} \tag{7.7.66}$$

The use of the positive Poisson representation applied to this system yields the SDE, arising from the FPE (7.7.16):

$$d\alpha = [k_3 C + (k_2 A - k_1 B)\alpha - k_4\alpha^2]dt + [2(k_2 A\alpha - k_4\alpha^2)]^{1/2} dW(t) . \tag{7.7.67}$$

In the case $\delta > 0$, we note that the noise term vanishes at $\alpha = 0$ and at $\alpha = k_2 A/k_4$, is positive between these points and the drift term is such as to return α to the range $[0, k_2 A/k_4]$ whenever it approaches the end points. Thus, for $\delta > 0$, (7.7.67) represents a real SDE on the real interval $[0, k_2 A/k_4]$.

In the case $\delta < 0$, the stationary point lies outside the interval $[0, k_2 A/k_4]$, and a point initially in this interval will move along this interval governed by (7.7.67) until it meets the right-hand end, where the noise vanishes and the drift continues to drive it towards the right. One leaving the interval, the noise becomes imaginary and the point will follow a path like that shown in Fig. 7.4 until it eventually reaches the interval $[0, k_2 A/k_4]$ again.

The case of $\delta = 0$ is not very dissimilar, except that once the point reaches the right-hand end of the interval $[0, k_2 A/k_4]$, both drift and diffusion vanish so it remains there from then on.

In the case of the system

$$\begin{aligned} B &\to X \\ 2X &\to A , \end{aligned} \tag{7.7.68}$$

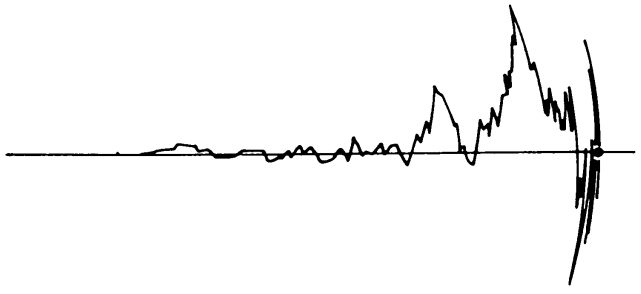


Fig. 7.4. Path followed by a point obeying the stochastic differential equation (7.7.67)

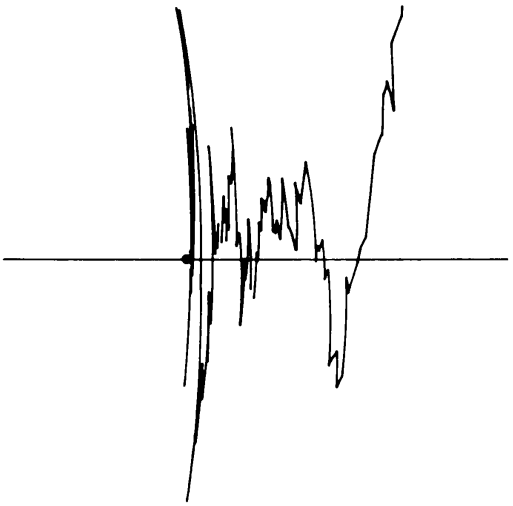


Fig. 7.5. Simulation of the path of a point obeying the stochastic differential equation (7.7.69) ▶

the SDE coming from the FPE (7.7.41) is

$$d\eta/dt = \kappa_1 - 2\kappa_2\eta^2 + i\varepsilon(2\kappa_2)^{1/2}\eta\xi(t)\,, \tag{7.7.69}$$

where $\alpha = \eta V$ and $\varepsilon = V^{-1/2}$.

The SDE (7.7.69) can be computer simulated and a plot of motion in the complex η plane generated. Figure 7.5 illustrates the behaviour. The point is seen to remain in the vicinity of $\mathrm{Re}\{\alpha\} = (a/2)^{1/2}$ but to fluctuate mainly in the imaginary direction on either side, thus giving rise to a negative variance in α.

7.7.5 Time Correlation Functions

The time correlation function of a Poisson variable $\boldsymbol{\alpha}$ is not the same as that for the variable $\boldsymbol{x}$. This can be seen, for example, in the case of a reaction $X \leftrightarrows Y$ which gives a Poisson Representation Fokker-Planck equation with no diffusion term. Hence, the Poisson variable does not fluctuate. We now show what the relationship is. For clarity, the demonstration is carried out for one variable only.

We define

$$\langle\alpha(t)\alpha(s)\rangle = \int d\mu(\alpha)d\mu(\alpha')\alpha\alpha' f(\alpha, t\,|\,\alpha', s)f(\alpha', s)\,. \tag{7.7.70}$$

We note that

$$f(\alpha, s\,|\,\alpha', s) = \delta_\mu(\alpha - \alpha')$$

which means that

$$\int d\mu(\alpha)\, e^{-\alpha}(\alpha^x/x!)f(\alpha, s\,|\,\alpha', s) = e^{-\alpha'}\alpha'^x/x! \tag{7.7.71}$$

so that

$$\int d\mu(\alpha)\, \alpha f(\alpha, t\,|\,\alpha', s) = \sum_{x,x'} xP(x, t\,|\,x', s)e^{-\alpha'}\alpha' x'/x'!$$

Hence,

$$\begin{aligned}\langle\alpha(t)\alpha(s)\rangle &= \sum_{x,x'} xP(x, t\,|\,x', s)\int d\mu(\alpha')(\alpha^{x'+1}e^{-\alpha'}/x'!)f(\alpha', s)\\ &= \sum_{x,x'} xP(x, t\,|\,x', s)\int d\mu(\alpha')\left[\left(-\alpha'\frac{\partial}{\partial\alpha'} + x'\right)(\alpha'^{x'}e^{-\alpha'}/x'!)\right]f(\alpha', s)\\ &= \sum_{x,x'} x\,x'P(x, t\,|\,x', s)P(x', s) \qquad (7.7.72)\\ &\quad - \int d\mu(\alpha')f(\alpha', s)\alpha'\frac{\partial}{\partial\alpha'}\sum_{x,x'} xP(x, t\,|\,x', s)(\alpha'^{x'}e^{-\alpha'}x!)\,. \qquad (7.7.73)\end{aligned}$$

We define

$$\langle\alpha(t)\,|\,[\alpha', s]\rangle \equiv \int d\alpha\, \alpha f(\alpha, t\,|\,\alpha', s) \tag{7.7.74}$$

as the mean of $\alpha(t)$ given the initial condition α' at s. Then the second term can be written

$$-\int d\mu(\alpha')\alpha' \frac{\partial}{\partial \alpha'} \langle \alpha(t) | [\alpha', s] \rangle f(\alpha', s) \equiv \left\langle \alpha' \frac{\partial}{\partial \alpha'} \langle \alpha(t) | [\alpha, s] \rangle \right\rangle \tag{7.7.75}$$

so we have

$$\langle x(t)x(s) \rangle = \langle \alpha(t)\alpha(s) \rangle + \left\langle \alpha' \frac{\partial}{\partial \alpha'} \langle \alpha(t) | [\alpha, s] \rangle \right\rangle . \tag{7.7.76}$$

Taking into account a many-variable situation and noting that

$$\langle \boldsymbol{x}(t) \rangle = \langle \boldsymbol{\alpha}(t) \rangle \qquad \text{always},$$

we have

$$\boxed{\langle x_a(t), x_b(s) \rangle = \langle \alpha_a(t), \alpha_b(s) \rangle + \left\langle \alpha'_b \frac{\partial}{\partial \alpha'_b} \langle \alpha_a(t) | [\boldsymbol{\alpha}', s] \rangle \right\rangle .} \tag{7.7.77}$$

This formula explicitly shows the fact that the Poisson representation gives a process which is closely related to the Birth-Death Master equation, but not isomorphic to it. The stochastic quantities of interest, such as time correlation functions, can all be calculated but are not given directly by those of the Poisson variable.

a) Interpretation in Terms of Statistical Mechanics

We assume for the moment that the reader is acquainted with the statistical mechanics of chemical systems. If we consider a system composed of chemically reacting components $A, B, C, \ldots$, the distribution function in the grand canonical ensemble is given by

$$P(I) = \exp\{\beta[\Omega + \sum_i \mu_i x_i(I) - E(I)]\}, \tag{7.7.78}$$

where I is an index describing the microscopic state of the system, $x_i(I)$ is the number of molecules of X_i in the state I, $E(I)$ is the energy of the state, μ_i is the chemical potential of component X_i, Ω is a normalization factor, and

$$\beta = 1/kT . \tag{7.7.79}$$

The fact that the components can react requires certain relationships between the chemical potentials to be satisfied, since a state I can be transformed into a state J only if

$$\sum_i \nu_i^A x_i(I) = \sum_i \nu_i^A x_i(J), \qquad A = 1, 2, 3, \ldots \tag{7.7.80}$$

where ν_a^A are certain integers. The relations (7.7.80) are the stoichiometric constraints.

The canonical ensemble for a reacting system is defined by requiring

$$\sum_i \nu_i^A x_i(I) = \tau^A , \tag{7.7.81}$$

for some τ^A, whereas the grand canonical ensemble is defined by requiring

$$\sum_i P(I) \sum_i \nu_i^A x_i(I) \equiv \sum_i \nu_i^A \langle x_i \rangle = \tau^A , \tag{7.7.82}$$

Maximization of entropy subject to the constraint (7.7.82) (and the usual constraints of fixed total probability and mean energy) gives the grand canonical form (7.7.78) in which the chemical potentials also satisfy the relation

$$\mu_i = \sum_A k_A \nu_i^A . \tag{7.7.83}$$

When one takes the ideal solution or ideal gas limit, in which interaction energies (but not kinetic or internal energies) are neglected, there is no difference between the distribution function for an ideal reacting system and an ideal nonreacting system, apart from the requirement that the chemical potentials be expressible in the form of (7.7.83).

The canonical ensemble is not so simple, since the constraints must appear explicitly as a factor of the form

$$\prod_A \delta[\sum_i \nu_i^A x_i(I), \tau^A] \tag{7.7.84}$$

and the distribution function is qualitatively different for every kind of reacting system (including a nonreacting system as a special case).

The distribution in total numbers $\boldsymbol{x}$ of molecules of reacting components in the grand canonical ensemble of an ideal reacting system is easily evaluated, namely,

$$P(\boldsymbol{x}) = \exp[\beta(\Omega + \sum_i \mu_i x_i)] \sum_I \prod_i \delta[x_i(I), x_i] \exp[-\beta E(I)] . \tag{7.7.85}$$

The sum over states is the same as that for the canonical ensemble of an ideal nonreacting mixture so that

$$P(\boldsymbol{x}) = \exp[\beta(\Omega + \sum_i \mu_i x_i)] \prod_i \frac{1}{x_i!} \{\sum_k \exp[-\beta E_k(i)]\}^{x_i} , \tag{7.7.86}$$

where $E_k(i)$ are the energy eigenstates of a single molecule of the substance A. This result is a multivariate Poisson with mean numbers given by

$$\log\langle x_i \rangle = \beta\mu_i - \log[\sum_k e^{-\beta E_k(i)}] \tag{7.7.87}$$

which, as is well known, when combined with the requirement (7.7.82) gives the law of mass action.

The canonical ensemble is obtained by maximizing entropy subject to the stronger constraint (7.7.81), which implies the weak constraint (7.7.82). Thus, the distribution function in total numbers for the canonical ensemble will simply be given by

$$P(\boldsymbol{x}) \propto \left\{\prod_i \frac{1}{x_i!}\left[\sum_k \mathrm{e}^{-\beta E_k(i)}\right]^{x_i}\right\} \sum_A \delta\left[\sum_i \nu_i^A x_i, \tau^A\right]. \tag{7.7.88}$$

In terms of the Poisson representation, we have just shown that in equilibrium situations, the quasiprobability (in a grand canonical ensemble) is

$$f(\boldsymbol{\alpha})_{eq} = \delta[\boldsymbol{\alpha} - \boldsymbol{\alpha}(eq)] \tag{7.7.89}$$

since the $\boldsymbol{x}$ space distribution is Poissonian. For the time correlation functions there are two results of this.

i) The variables $\boldsymbol{\alpha}(t)$ and $\boldsymbol{\alpha}(s)$ are nonfluctuating quantities with values $\boldsymbol{\alpha}(eq)$. Thus,

$$\langle \alpha_r(t), \alpha_s(s)\rangle_{eq} = 0\,. \tag{7.7.90}$$

ii) The equilibrium mean in the second term is trivial. Thus,

$$\langle x_a(t), x_b(s)\rangle = \left[\alpha_b' \frac{\partial}{\partial \alpha_b'} \langle \alpha_a(t) \,|\, [\boldsymbol{\alpha}', s]\rangle\right]_{\alpha' = \alpha(eq)}. \tag{7.7.91}$$

This result is, in fact, exactly that of *Bernard* and *Callen* [7.11] which relates a two-time correlation function to a derivative of the mean of a quantity with respect to a thermodynamically conjugate variable.

Consider a system in which the numbers of molecules of chemical species X_1, X_2, ... corresponding to a configuration I of the system are $x_1(I)$, $x_2(I)$, ... and it is understood that these chemical species may react with each other. Then in a grand canonical ensemble, as demonstrated above, the equilibrium distribution function is

$$Z^{-1}(\boldsymbol{\mu}) \exp\left[\{\sum_i \mu_i x_i(I) - E(I)\}/kT\right] \tag{7.7.92}$$

with

$$Z(\boldsymbol{\mu}) = \exp(-\Omega\beta)\,, \tag{7.7.93}$$

where $Z(\boldsymbol{\mu})$ is the grand canonical partition function. As pointed out above, the chemical potentials μ_i for a reacting system cannot be chosen arbitrarily but must be related by the stoichiometric constraints (7.7.82) of the allowable reactions.

Now we further define the quantities

$$\langle x_i, t \,|\, [I, s]\rangle \tag{7.7.94}$$

to be the mean values of the quantities x_i at time t under the condition that the system was in a configuration I at time s. Then a quantity of interest is the mean value of (7.7.94) over the distribution (7.7.92) of initial conditions, namely,

$$\langle x_i, t|[\mu, s]\rangle = \sum_J \langle x_i, t|[J, s]\rangle Z^{-1}(\mu) \times \exp\left\{\frac{1}{kT}\left[\sum_j \mu_j x_j(J) - E(J)\right]\right\}. \tag{7.7.95}$$

When the chemical potentials satisfy the equilibrium constraints, this quantity will be time independent and equal to the mean of x_i in equilibrium, but otherwise it will have a time dependence. Then, with a little manipulation one finds that

$$\left[kT\frac{\partial}{\partial \mu_j}\langle x_i, t|[\mu, s]\rangle\right]_{\mu=\mu(eq)} = \langle x_i(t), x_j(s)\rangle_{eq}. \tag{7.7.96}$$

The left-hand side is a reponse function of the mean value to the change in the chemical potentials around equilibrium and is thus a measure of dissipation, while the right-hand side, the two-time correlation function in equilibrum, is a measure of fluctuations.

To make contact with the Poisson representation result (7.7.91) we note that the chemical potentials μ_j in ideal solution theory are given by

$$\mu_i(\langle \boldsymbol{x}\rangle) = kT \log\langle x_i\rangle + \text{const}. \tag{7.7.97}$$

Using (7.7.97), we find that (7.7.96) becomes

$$\langle x_i(t), x_j(s)\rangle = \left[\langle x_j\rangle \frac{\partial}{\partial\langle x_j\rangle}\langle x_i, t|[\mu(\langle \boldsymbol{x}\rangle), s]\rangle\right]_{\langle \boldsymbol{x}\rangle=\langle \boldsymbol{x}\rangle_{eq}}. \tag{7.7.98}$$

Since the ideal solution theory gives rise to a distribution in x_i that is Poissonian, it follows that in that limit

$$\langle x_i, t|[\mu(\langle \boldsymbol{x}\rangle), s]\rangle = \langle \alpha_i, t|[\boldsymbol{\alpha}', s]\rangle \tag{7.7.99}$$

with $\alpha' = \langle \boldsymbol{x}\rangle$. Thus, (7.7.98) becomes

$$\langle x_i(t), x_j(s)\rangle = \left[\alpha'_j \frac{\partial}{\partial \alpha_j}\langle \alpha_i, t|[\boldsymbol{\alpha}', s]\rangle\right]_{\alpha'=\alpha(eq)}. \tag{7.7.100}$$

Thus, (7.7.91) is the ideal solution limit of the general result (7.7.98).

The general formula (7.7.77) can be considered as a generalization of the Bernard-Callen result to systems that are not in thermodynamic equilibrium.

However, it is considerably different from the equilibrium result and the two terms are directly interpretable. The second term is the equilibrium contribution, a response function, but since the system is not in a well-defined equilibrium state, we take the average of the equilibrium result over the various contributing α space states. The first term is the contribution from the α-space fluctuations themselves and is not directly related to a response function. It represents the fluctuations in excess of equilibrium.

b) Linearised Results

The general differential equation, arising from the use of the positive Poisson representation, and corresponding to the FPE (7.7.12), is

$$d\eta = \hat{A}(\eta)dt + \varepsilon\hat{\underline{C}}(\eta)dW(t)\,, \tag{7.7.101}$$

where

$$\hat{\underline{C}}\hat{\underline{C}}^{\mathrm{T}} = \hat{\underline{B}}\,. \tag{7.7.102}$$

We may now make a first-order small noise expansion about the stationary state η by following the procedure of Sect.6.3. Thus, writing

$$\eta(t) = \bar{\eta} + \varepsilon\eta_1(t) \qquad (\varepsilon = V^{-1/2}) \tag{7.7.103}$$

to lowest order we have

$$\hat{A}(\bar{\eta}) = 0 \tag{7.7.104}$$

$$d\eta_1 = -\underline{F}\eta_1 dt + \underline{G}\,dW(t)$$

where

$$F_{rs} = -\frac{\partial}{\partial\eta_s}\hat{A}_r(\bar{\eta}) \tag{7.7.105}$$

$$\underline{G} = \hat{\underline{C}}(\bar{\eta})\,.$$

Then,

$$\langle\alpha_r(t), \alpha_s(0)\rangle_s = V\sum_{r'}[\exp(-\underline{F}t)]_{rr'}\langle\eta_{r',1}, \eta_{s,1}\rangle_s \tag{7.7.106}$$

and

$$\frac{\partial}{\partial\alpha_s'}\langle\alpha_r(t)\,|\,[\boldsymbol{\alpha}', 0]\rangle = \frac{\partial}{\partial\eta_{s,1}'}\langle\eta_{r,1}(t)\,|\,[\boldsymbol{\eta}_1', 0]\rangle = [\exp(-Ft)]_{r,s}\,. \tag{7.7.107}$$

Hence,

$$\langle x_r(t), x_s(0)\rangle_s = V\sum_{r'}[\exp(-Ft)]_{rr'}[\langle\eta_{r',1}, \eta_{s,1}\rangle_s + \delta_{r',s}\bar{\eta}_s) \tag{7.7.108}$$

$$= \sum_{r'}\exp(-Ft)_{r,r'}\langle x_r, x_s\rangle_s\,. \tag{7.7.109}$$

Thus the linearised result is in agreement with the regression theorem of sect. 3.7.4 correlation functions for a variety of systems nave been computed in [7.10].

7.7.6 Trimolecular Reaction

In Sect. 7.1.3 we considered a reaction which included a part

$$A + 2X \rightleftharpoons 3X \tag{7.7.110}$$

and set up an appropriate birth-death Master equation for this. However, it is well known in chemistry that such trimolecular steps are of vanishingly small probability and proceed in stages via a short-lived intermediate. Thus, the reaction (7.7.110) presumably occurs as a two-state system

$$\text{i)}\quad A + Y \underset{1}{\overset{1}{\rightleftharpoons}} X + Y \tag{7.7.111a}$$

$$\text{ii)}\quad Y \underset{1}{\overset{\gamma}{\rightleftharpoons}} 2X\,, \tag{7.7.111b}$$

both of which are merely bimolecular, and we have set rate constants equal to one, except for γ (the decay constant of Y) which is assumed as being very large. Thus, Y is indeed a short-lived intermediate. The deterministic rate equations are

$$\begin{aligned}\frac{dx}{dt} &= ay - xy + 2(\gamma y - x^2)\\ \frac{dy}{dt} &= x^2 - \gamma y\end{aligned} \tag{7.7.112}$$

and the usual deterministic adiabatic elimination procedure sets $y = x^2/\gamma$ and gives

$$\frac{dx}{dt} = (ax^2 - x^3)/\gamma\,. \tag{7.7.113}$$

Although this procedure is straightfoward deterministically, it is not clear that the stochastic Master equation of the kind used in Sect.7.1.3 is a valid adiabatic elimination limit. The adiabatic elimination techniques used in Chap. 6 are not easily adapted to direct use on a Master equation but can be straightfowardly adapted to the case of the Poisson representation Fokker-Planck equation.

a) Fokker-Planck Equation for Trimolecular Reaction

For the reaction (7.7.110) with forward and backward rate constants equal to $1/\gamma$ to correspond to (7.7.113), the Poisson representation Fokker-Planck equation becomes, from (7.7.4),

$$\frac{\partial f}{\partial t} = \frac{1}{\gamma}\left\{\left(-\frac{\partial}{\partial \alpha} + 2\frac{\partial^2}{\partial \alpha^2} - \frac{\partial^3}{\partial \alpha^3}\right)[\alpha^2(a-\alpha)]f\right\} \tag{7.7.114}$$

and contains *third-order derivatives*. There is no truly probabilistic interpretation in terms of any real stochastic process in α space, no matter what kind of Poisson representation is chosen. The concept of *third-order noise* will be explained in the next section, which will show how probabilistic methods and stochastic differential equations can still be used.

b) Adiabatic Elimination

Using the rules developed in (7.4.9), the Fokker-Planck equation for the system (7.7.111) with the correspondence

$$\begin{bmatrix} x \\ y \end{bmatrix} \leftrightarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

is

$$\frac{\partial f}{\partial \alpha} = -\frac{\partial}{\partial \alpha}[(a-\alpha)\beta + 2(\gamma\beta - \alpha^2)] + \frac{\partial}{\partial \beta}(\gamma\beta - \alpha^2) + \frac{\partial^2}{\partial \alpha^2}(\gamma\beta - \alpha^2) + \frac{\partial^2}{\partial \alpha \partial \beta}[(a-\alpha)\beta]\,. \tag{7.7.115}$$

Adiabatic elimination now proceeds as in Sect.6.6.1. We define new variables

$$\begin{aligned} x &= \alpha \\ y &= \gamma\beta - \alpha^2 \end{aligned} \tag{7.7.116}$$

and consequently, changing variables with

$$\begin{aligned} \frac{\partial}{\partial \alpha} &= \frac{\partial}{\partial x} - 2x\frac{\partial}{\partial y} \\ \frac{\partial}{\partial \beta} &= \gamma\frac{\partial}{\partial y}, \end{aligned} \tag{7.7.117}$$

the FPE becomes

$$\frac{\partial f(x,y)}{\partial t} = \left\{-\left(\frac{\partial}{\partial x} - 2x\frac{\partial}{\partial y}\right)\left[\frac{(a-x)(y+x^2)}{\gamma} + 2y\right] + \gamma\frac{\partial}{\partial y}y + \left(\frac{\partial}{\partial x} - 2x\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} - 2x\frac{\partial}{\partial y}\right)y + \left(\frac{\partial}{\partial x} - 2x\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial y}\right)[(y+x^2)(a-x)]\right\} f. \tag{7.7.118}$$

Since y is to be eliminated, there should be a well-defined limit of the L_1 operator which governs its motion at fixed x. However, this operator is

$$\gamma\frac{\partial}{\partial y}y + \frac{\partial^2}{\partial y^2}[4x^2y - 2x(y+x^2)(a-x)] \tag{7.7.119}$$

and the large γ limit turns this into deterministic motion. Setting

$$y = v\gamma^{-1/2} \tag{7.7.120}$$

transforms(7.7.119) to

$$\begin{aligned} \gamma L_1(\gamma) &= \gamma\left\{\frac{\partial}{\partial v}v + \frac{\partial^2}{\partial v^2}[2x^3(x-a) + (4x^2 - 2x)v\gamma^{-1/2}]\right\} \\ &\xrightarrow[\gamma\to\infty]{} \gamma\left\{\frac{\partial}{\partial v}v + \frac{\partial^2}{\partial v^2}[2x^3(x-a)]\right\} \\ &\equiv \gamma L_1\,. \end{aligned} \tag{7.7.121}$$

With this substitution, we finally identify

$$\gamma^{-1}L_3 = -\gamma^{-1}\frac{\partial}{\partial x}[x^2(a-x)] \tag{7.7.122}$$

$$L_2(\gamma) = -\frac{\partial}{\partial x}[(a-x)v\gamma^{-3/2} + 2v\gamma^{-1/2}] - \frac{\partial}{\partial x}2x\frac{\partial}{\partial v}v - 2x\frac{\partial}{\partial v}\frac{\partial}{\partial x}v$$
$$+ \gamma^{-1/2}\frac{\partial^2}{\partial x^2}v + \gamma^{1/2}\frac{\partial}{\partial x}\frac{\partial}{\partial v}[(a-x)(x^2 + v\gamma^{-1/2})] \tag{7.7.123}$$

and

$$\frac{\partial f}{\partial t} = [\gamma^{-1}L_3 + L_2(\gamma) + \gamma L_1(\gamma)]f. \tag{7.7.124}$$

The projection operator P will be onto the null space of L_1 and because L_1 depends on x, we have

$$L_3P \neq PL_3\,. \tag{7.7.125}$$

This means that the equation of motion for $Pf \equiv g$ is found by similar algebra to that used in Sect. 6.5.4. We find

$$s\tilde{g}(s) = \gamma^{-1}PL_3\tilde{g}(s) + P[L_2(\gamma) + \gamma^{-1}L_3][s - \gamma L_1 - (1-P)L_2(\gamma) - \gamma^{-1}(1-P)L_3]^{-1}$$
$$\times [L_2(\gamma) + \gamma^{-1}(1-P)L_3]\tilde{g}(s) + g(0)\,. \tag{7.7.126}$$

Notice, however, since for any function of v

$$P\phi(v) = p_x(v)\int dv\,\phi(v) \tag{7.7.127}$$

where $p_x(v)$ satisfies

$$L_1p_x(v) = 0\,, \tag{7.7.128}$$

that in $PL_2(\gamma)$, all terms with $\partial/\partial v$ in them vanish. Thus, to highest order in γ,

$$PL_2(\gamma) = \gamma^{-1/2}\left(-2v\frac{\partial}{\partial x} + v\frac{\partial^2}{\partial x^2}\right). \tag{7.7.129}$$

The term $[\ ]^{-1}$ in (7.7.126) is asymptotic to $-\gamma^{-1}L_1^{-1}$ and the only term in the remaining bracket which can make the whole expression of order γ^{-1}, like the L_3 term, is the term of order $\gamma^{1/2}$ in $L_2(\gamma)$, i.e.,

$$\gamma^{1/2}\frac{\partial}{\partial x}[(a-x)x^2]\frac{\partial}{\partial v}. \tag{7.7.130}$$

Thus, the large γ limit of (7.7.126) is

$$s\tilde{g}(s) = \gamma^{-1}\left\{PL_3\tilde{g} - P\left[-2\frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2}\right]vL_1^{-1}\frac{\partial}{\partial x}\left[(a-x)x^2\frac{\partial}{\partial v}\right]p_x(v)\tilde{p}\right\} + g(0) \tag{7.7.131}$$

where we have written

$$g = p_x(v)p, \qquad \tilde{g} = p_x(v)\tilde{p}\,. \tag{7.7.132}$$

We are now lead to the central problem of the evaluation of

$$\int dv'v'L_1^{-1}\frac{\partial}{\partial x}(a-x)x^2\frac{\partial}{\partial v'}p_x(v') \tag{7.7.133}$$

which arises in the evaluation of the second part in the braces in (7.7.131). We wish to bring the $\partial/\partial x$ to the left outside the integral, but since $\partial/\partial x$ and L_1 do not commute, this requires care. Now

$$\left[L_1^{-1}, \frac{\partial}{\partial x}\right] = L_1^{-1}\left[\frac{\partial}{\partial x}, L_1\right]L_1^{-1} \tag{7.7.134}$$

and from (7.7.121),

$$= L_1^{-1}\left\{\frac{\partial^2}{\partial v^2}[8x^3 - 6ax^2]\right\}L_1^{-1}\,; \tag{7.7.135}$$

$$(7.7.133) = \frac{\partial}{\partial x}\int dv'v'L_1^{-1}(a-x)x^2\frac{\partial}{\partial v'}p_x(v')$$

$$+ \int dv'v'L_1^{-1}\frac{\partial^2}{\partial v'^2}L_1^{-1}[(8x^3 - 6ax^2)(a-x)x^2]\frac{\partial}{\partial v'}p_x(v')\,. \tag{7.7.136}$$

The second term vanishes, through the demonstration that this is so is rather specialised. For, we know that L_1 describes an Ornstein-Uhlenbeck process in v and that $p_x(v)$ is its stationary solution. The eigenfunction properties used in Sect.6.4.2 show that

$$L_1^{-1}\frac{\partial^2}{\partial v^2}L_1^{-1}\frac{\partial}{\partial v}p_x(v) \tag{7.7.137}$$

is proportional to the third eigenfunction, which is orthogonal to v, the first eigenfunction of the corresponding backward equation. The first term is now easily computed using the fact that L_1 involves the Ornstein-Uhlenbeck process. Using the same techniques as in Sect.6.6.1, we find that all the x dependence arising from $p_x(v')$ vanishes, and hence

$$(7.7.133) = \frac{\partial}{\partial x}(a-x)x^2\,. \tag{7.7.138}$$

We similarly find

$$PL_3\tilde{g} = -p_x(v)\int dv' \frac{\partial}{\partial x}[x^2(a-x)]p_x(v')\tilde{p}$$

$$= -p_x(v)\frac{\partial}{\partial x}[x^2\,(a-x)]\,\tilde{p}\,, \tag{7.7.139}$$

so in the end

$$\frac{\partial p}{\partial t} = \frac{1}{\gamma}\left[\left(-\frac{\partial}{\partial x} + 2\frac{\partial^2}{\partial x^2} - \frac{\partial^3}{\partial x^3}\right)[(a-x)x^2]p\right] \tag{7.7.140}$$

which is exactly the same as the trimolecular model Fokker-Planck equation (7.7.114). This means the trimolecular Master equation is valid in the same limit.

c) Comments

i) Notice that this system gives an end result which is *not* in a Stratonovich form but in the Ito form, with all derivatives to the left.

ii) The derivation of (7.7.140) means that techniques for understanding such non-probabilistic Fokker-Planck equations are required. We outline a possible way of doing this in the next section.

7.7.7 Third-Order Noise

To handle the third-order Fokker-Planck equations which arise with trimolecular reactions, we introduce the stochastic variable $V(t)$ whose conditional probability density $p(v, t)$ obeys the third-order partial differential equation

$$\partial p(v, t)/\partial t = -\tfrac{1}{6}\,\partial^3 p(v, t)/\partial v^3\,. \tag{7.7.141}$$

Since we have already shown in Sect.3.4 that no Markov process can possible give a third-order term like this, some fundamental requirement must be violated by $p(v, t)$. It turns out that $p(v, t)$ is not always positive, which is permissible in a quasi-probability. We will see that in spite of this, the formal probabilistic analogy is very useful.

We know that the solution of (7.7.141), subject to the boundary condition

$$p(v, t_0) = \delta(v - v_0)\,, \tag{7.7.142}$$

is given by Fourier transform methods as

$$p(v, t\,|\,v_0, t_0) = (1/2\pi)\int_{-\infty}^{\infty} dq \exp\{\mathrm{i}[q(v-v_0) + \tfrac{1}{6}\,q^3(t-t_0)]\}\,. \tag{7.7.143}$$

The moments of V can be calculated, after a partial integration, to be

$$\langle[V(t) - V_0]^n\rangle = 0, \qquad n \text{ not a multiple of 3}$$
$$\langle[V(t) - V_0]^{3m}\rangle = [(t-t_0)/6]^m(3m)!/m!\,. \tag{7.7.144}$$

Further, we assume the process (7.7.141) is some kind of generalized Markov process, for which the joint probability distribution is given by

$$p(v_2t_2\colon v_1t_1) = p(v_2t_2\,|\,v_1t_1)p(v_1, t_1) \tag{7.7.145}$$

and from (7.7.142) we see that the first factor is a function of only $v_2 - v_1$ and $t_2 - t_1$, so that the variable $V(t_2) - V(t_1)$ is statistically independent of $V(t_1)$ and that this process is a *process with independent increments.* Thus, $dV(t)$ will be independent of $V(t)$.

The rigorous definition of stochastic integration with respect to $V(t)$ is a task that we shall not attempt at this stage. However, it is clear that it will not be too dissimilar to Ito integration and, in fact, *Hochberg* [7.12] has rigorously defined higher-order noises of even degree and carried out stochastic integration with respect to them. We can show, however, that a stochastic differential equation of the form

$$dy(t) = a(y)dt + b(y)dW(t) + c(y)dV(t) \tag{7.7.146}$$

[with $W(t)$ and $V(t)$ independent processes] is equivalent to a third-order Fokker-Planck equation. It is clear that because $W(t)$ and $V(t)$ are processes with independent increments, $y(t)$ is a Markov process. We then calculate

$$\lim_{t\to t_0} \frac{\langle[y(t) - y(t_0)]^n\rangle}{t - t_0} = \lim_{dt_0\to 0} \frac{\langle[dy(t_0)]^n\rangle}{dt_0}, \tag{7.7.147}$$

where $y(t_0)$ is a numerical initial value, not a stochastic variable. From (7.7.146), $y(t)$ depends on $W(t')$ and $V(t')$ for only $t' \leqslant t$ and, since $dW(t)$ and $dV(t)$ are independent of $y(t)$, we find

$$\begin{aligned}\langle dy(t_0)\rangle &= \langle a[y(t_0)]\rangle dt_0 + \langle b[y(t_0)]\rangle\langle dW(t_0)\rangle + \langle c[y(t_0)]\rangle\langle dV(t_0)\rangle \\ &= \langle a[y(t_0)]\rangle dt_0 = a[y(t_0)]dt_0\end{aligned} \tag{7.7.148}$$

because $y(t_0)$ is a numerical initial value. Similarly, to lowest order in dt_0

$$\begin{aligned}\langle dy(t_0)^2\rangle &= b[y(t_0)]^2\langle dW(t_0)^2\rangle \\ &= b[y(t_0)]^2 dt_0\end{aligned} \tag{7.7.149}$$

$$\begin{aligned}\langle dy(t_0)^3\rangle &= c[y(t_0)]^3\langle dV(t_0)^3\rangle \\ &= c[y(t_0)]^3 dt_0\,.\end{aligned} \tag{7.7.150}$$

Thus, we find

$$\left.\begin{aligned}&\lim_{t\to t_0}[\langle y(t) - y(t_0)\rangle/(t - t_0)] = a[y(t_0)] \\ &\lim_{t\to t_0}[\langle[y(t) - y(t_0)]^2\rangle/(t - t_0)] = b[y(t_0)]^2 \\ &\lim[\langle]y(t) - y(t_0)]^3\rangle/(t - t_0)] = c[y(t_0)]^3\end{aligned}\right\} \tag{7.7.151}$$

and all higher powers give a zero result. By utilising a similar analysis to that of Sect. 3.4, this is sufficient to show that $y(t)$ is a generalized diffusion process whose generalized FPE is

$$\frac{\partial p(y,t)}{\partial t} = -\frac{\partial}{\partial y}[a(y)p] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[b(y)^2 p] - \frac{1}{6}\frac{\partial^3}{\partial y^3}[c(y)^3 p]\,. \tag{7.7.152}$$

We define a noise source $\zeta(t)$ by

$$dV(t) = \zeta(t)dt\,, \tag{7.7.153}$$

where

$$\langle\zeta(t)\rangle = \langle\zeta(t)\zeta(t')\rangle = 0 \tag{7.7.154}$$

$$\langle\zeta(t)\zeta(t')\zeta(t'')\rangle = \delta(t-t')\delta(t'-t'') \tag{7.7.155}$$

and higher moments can be readily calculated from the moments of $dV(t)$. The independence of increments means that, as with the Ito integral, integrals that have a delta-function singularity at their upper limit are to be taken as zero.

Example of the Use of Third-Order Noise. Consider the chemical process

$$A + 2X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3X \tag{7.7.156}$$

$$A \underset{k_4}{\overset{k_3}{\rightleftharpoons}} X$$

whose Poisson representation FPE is

$$\begin{aligned}\frac{\partial f(\alpha,t)}{\partial t} = &-\frac{\partial}{\partial\alpha}[\kappa_1 V^{-1}\alpha^2 - \kappa_2 V^{-2}\alpha^3 + \kappa_3 V - \kappa_4\alpha)f(\alpha,t)] \\ &+\frac{1}{2}\frac{\partial^2}{\partial\alpha^2}[4(\kappa_1 V^{-1}\alpha^2 - \kappa_2 V^{-2}\alpha^3)f(\alpha,t)] \\ &-\frac{1}{3!}\frac{\partial^3}{\partial\alpha^3}[6(\kappa_1 V^{-1}\alpha^2 - \kappa_2 V^{-2}\alpha^3)f(\alpha,t)]\,,\end{aligned} \tag{7.7.157}$$

where $\kappa_1 V^{-1} = k_1 A$, $\kappa_2 V^{-2} = k_2$, $\kappa_3 V = k_3$, $\kappa_4 = k_4$.

In the steady state, (7.7.157) reduces to a linear second-order differential equation which may be solved in terms of hypergeometric functions, and an asymptotic expansion for the various moments can be obtained using steepest descent methods. This procedure, although possible in principle, is not very practicable. It is in such cases that the method of stochastic differential equations proves to be very useful because of its ease of application.

The stochastic differential equation equivalent to (7.7.157) is

$$\begin{aligned} d\eta(t)/dt &= \kappa_1\eta(t)^2 - \kappa_2\eta(t)^3 + \kappa_3 - \kappa_4\eta(t) \\ &\quad + \mu^3\{4[\kappa_1\eta(t)^2 - \kappa_2\eta(t)^3]\}^{1/2}\xi(t) \\ &\quad + \mu^4\{6[\kappa_1\eta(t)^2 - \kappa_2\eta(t)^3]\}^{1/3}\zeta(t)\,, \end{aligned} \tag{7.7.158}$$

where $\alpha = \eta V$, $\mu = V^{-1/6}$ and the noise source $\zeta(t)$, henceforth referred to as the "third-order noise", has been defined in (7.7.153–155)

Equation (7.7.158) may be solved iteratively by expanding $\eta(t)$:

$$\eta(t) = \eta_0(t) + \mu^3\eta_3(t) + \mu^4\eta_4(t) + \mu^6\eta_6(t) + \mu^8\eta_8(t) + \mu^9\eta_9(t) + \ldots \tag{7.7.159}$$

which, when substituted in (7.7.158), yields the deterministic equation in the lowest order and linear stochastic differential equations in the higher orders which may be solved as in Sect.6.2.

In the stationary state the results are

$$\langle x\rangle = V\eta_0 + \langle\eta_6\rangle + \cdots = V\eta_0 + \frac{2ab}{c^2} + \ldots \tag{7.7.160a}$$

$$\begin{aligned} \langle x^2\rangle - \langle x^2\rangle &= V\langle\eta_3^2\rangle + [2\langle\eta_9\eta_3\rangle + 2\langle\eta_8\eta_4\rangle + \langle\eta_6^2\rangle - \langle\eta_6\rangle^2 + \langle\eta_6\rangle] + \ldots \\ &= V\left[\frac{2a}{c}\right] + \left[\frac{28}{3}\frac{a^2b^2}{c^4} + \frac{8ab^2\eta_0}{c^3} - \frac{36\kappa_2a^2}{c^3} + \frac{8ab}{c^2}\right] + \ldots \end{aligned} \tag{7.7.160b}$$

$$\begin{aligned} \langle(x - \langle x\rangle)^3\rangle &= V[\langle\eta_4^3\rangle - 3\langle\eta_3^2\rangle\langle\eta_6\rangle + 3\langle\eta_3^2\rangle + \langle\eta_0\rangle + \ldots \\ &= V\left[\frac{8a}{c} - \frac{12a^2b}{c^3} + \eta_0\right] + \ldots\,, \end{aligned} \tag{7.7.160c}$$

where $a = \kappa_1\eta_0^2 - \kappa_2\eta_0^3$, $b = 2\kappa_1 - 3\kappa_2\eta_0$, $c = \kappa_4 - 2\kappa_1\eta_0 + 3\kappa_2\eta_0^2$ and η_0 is the solution of the steady-state deterministic equation

$$\kappa_1\eta_0^2 - \kappa_2\eta_0^3 + \kappa_3 - \kappa_4\eta_0 = 0\,. \tag{7.7.161}$$

Here a few remarks are in order. The third-order noise $\zeta(t)$ contributes to $O(V^{-1})$ to the mean and to $O(1)$ to the variance, but contributes to $O(V)$ to the skewness coefficient. If one is only interested in calculating the mean and the variance to $O(V)$, the third-order noise may be dropped from (7.7158) and the expansiou carried out in powers of $\varepsilon = V^{-1/2}$. Also note that as $c \to 0$, the variance and the higher order corrections become divergent. This of, course, is due to the fact that in this limit, the reaction system exhibits a first-order phase transition type behaviour.

8. Spatially Distributed Systems

Reaction diffusion systems are treated in this chapter as a prototype of the host of spatially distributed systems that occur in nature. We introduce the subject heuristically by means of spatially dependent Langevin equations, whose inadequacies are explained. The more satisfactory multivariate master equation description is then introduced, and the spatially dependent Langevin equations formulated as an approximation to this description, based upon a system size expansion. It is also shown how Poisson representation methods can give very similar spatially dependent Langevin equations without requiring any approximation.

We next investigate the consequences of such equations in the spatial and temporal correlation structures which can arise, especially near instability points. The connection between local and global descriptions is then shown. The chapter concludes with a treatment of systems described by a distribution in phase space (i.e. the space of velocity and position). This is done by means of the *Boltzmann Master equation.*

8.1 Background

The concept of space is central to our perception of the world, primarily because well-separated objects do not, in general, have a great deal of influence on each other. This leads to the description of the world, on a macroscopic deterministic level, by *local* quantities such as local density, concentration, temperature, electromagnetic potentials, and so on. Deterministically, these are normally thought of as obeying partial differential equations such as the *Navier-Stokes* equations of hydrodynamics, the *reaction diffusion* equations of chemistry or *Maxwell's* equations of classical electromagnetism.

The simplest cases to consider are reaction diffusion equations, which describe chemical reactions and which form the main topic of this chapter. In order to get some feel of the concept, let us first consider a Langevin equation description for the time evolution of the concentration ρ of a chemical substance. Then the classical reaction-diffusion equation can be derived as follows. A diffusion *current* $\boldsymbol{j}(\boldsymbol{r}, t)$ exists such that

$$\boldsymbol{j}(\boldsymbol{r}, t) = -D\nabla \rho(\boldsymbol{r}, t) \tag{8.1.1}$$

and (8.1.1) is called Fick's law. If there is no chemical reaction, this current obeys a conservation equation. For, considering an arbitrary volume V, the total amount of chemical in this volume can only change because of transport across the boundary $\boldsymbol{S}$, of V. Thus, if N is the total amount in V,

$$\frac{dN}{dt} = \frac{d}{dt}\int_V d^3\mathbf{r}\ \rho(\mathbf{r}, t) = -\int_S d\mathbf{S}\cdot\mathbf{j}(\mathbf{r}, t) \tag{8.1.2}$$
$$= -\int_V d^3\mathbf{r}\ \nabla\cdot\mathbf{j}(\mathbf{r}, t)\,.$$

Hence, since V is arbitrary,

$$\partial_t\rho(\mathbf{r}, t) + \nabla\cdot\mathbf{j}(\mathbf{r}, t) = 0\,. \tag{8.1.3}$$

Substituting Fick's law (8.1.1) into the conservation equation (8.1.3) we get

$$\partial_t\rho(\mathbf{r}, t) = D\nabla^2\rho(\mathbf{r}, t)\,, \tag{8.1.4}$$

the diffusion equation. Now how can one add fluctuations? First notice that the conservation equation (8.1.3) is exact; this follows from its derivation. We cannot add a fluctuating term to it. However, Fick's law could well be modified by adding a stochastic source. Thus, we rewrite

$$\mathbf{j}(\mathbf{r}, t) = -D\nabla\ \rho(\mathbf{r}, t) + \mathbf{f}_{\mathrm{d}}(\mathbf{r}, t)\,. \tag{8.1.5}$$

Here, $\mathbf{f}_{\mathrm{d}}(\mathbf{r}, t)$ is a vector Langevin source. The simplest assumption to make concerning its stochastic properties is

$$\langle \mathbf{f}_{\mathrm{d}}(\mathbf{r}, t)\rangle = 0$$

and

$$\langle f_{\mathrm{d},i}(\mathbf{r}, t) f_{\mathrm{d},j}(\mathbf{r}', t')\rangle = K_{\mathrm{d}}(\mathbf{r}, t)\delta_{ij}\delta(\mathbf{r} - \mathbf{r}')\delta(t - t')\,, \tag{8.1.6}$$

that is, the different components are independent of each other at the same time and place, and all fluctuations at different times or places are independent. This is a *locality* assumption. The fluctuating diffusion equation is then

$$\partial_t\rho(\mathbf{r}, t) = D\nabla^2\rho(\mathbf{r}, t) - \nabla\cdot\mathbf{f}_{\mathrm{d}}(\mathbf{r}, t)\,. \tag{8.1.7}$$

Notice that

$$\langle\nabla\cdot\mathbf{f}_{\mathrm{d}}(\mathbf{r}, t)\nabla'\cdot\mathbf{f}_{\mathrm{d}}(\mathbf{r}', t')\rangle = \nabla\cdot\nabla'[K_{\mathrm{d}}(\mathbf{r}, t)\delta(\mathbf{r} - \mathbf{r}')]\delta(t - t')\,. \tag{8.1.8}$$

Now consider including a chemical reaction. Fick's law still applies, but instead of the conservation equation we need an equation of the form

$$\frac{dN}{dt} = \frac{d}{dt}\int_V d^3\mathbf{r}\ \rho(\mathbf{r}, t) = -\int_S d\mathbf{S}\cdot\mathbf{j}(\mathbf{r}, t) + \int_V d^3\mathbf{r}\ F[\rho(\mathbf{r}, t)]\,, \tag{8.1.9}$$

where $F[\rho(\mathbf{r}, t)]$ is a function of the concentration and represents the production of the chemical by a *local* chemical reaction.

Hence we find, before taking fluctuations into account,

$$\partial_t\rho(\mathbf{r}, t) + \nabla\cdot\mathbf{j}(\mathbf{r}, t) = F[\rho(\mathbf{r}, t)]\,. \tag{8.1.10}$$

The production of the chemical by a chemical reaction does generate fluctuations, so we can add to (8.1.10) a term $f_c(\boldsymbol{r}, t)$ which satisfies

$$\begin{aligned} &\langle f_c(\boldsymbol{r}, t)\rangle = 0 \\ &\langle f_c(\boldsymbol{r}, t) f_c(\boldsymbol{r}', t')\rangle = K_c(\boldsymbol{r}, t)\delta(\boldsymbol{r} - \boldsymbol{r}')\delta(t - t') \end{aligned} \tag{8.1.11}$$

which expresses the fact that the reaction is *local* (i.e., fluctuations at different points are uncorrelated) and *Markov* (delta correlated in time). The full reaction-diffusion chemical equation now becomes

$$\boxed{\partial_t \rho(\boldsymbol{r}, t) = D\nabla^2 \rho(\boldsymbol{r}, t) + F[\rho(\boldsymbol{r}, t)] + g(\boldsymbol{r}, t)} \tag{8.1.12}$$

where

$$g(\boldsymbol{r}, t) = -\nabla \cdot \boldsymbol{f}_d(\boldsymbol{r}, t) + f_c(\boldsymbol{r}, t) \tag{8.1.13}$$

and

$$\boxed{\langle g(\boldsymbol{r}, t) g(\boldsymbol{r}', t')\rangle = \{K_c(\boldsymbol{r} - \boldsymbol{r}', t)\delta(\boldsymbol{r} - \boldsymbol{r}') + \nabla \cdot \nabla' [K_d(\boldsymbol{r}, t)\delta(\boldsymbol{r} - \boldsymbol{r}')]\}\, \delta(t - t').} \tag{8.1.14}$$

The simplest procedure for turning a classical reaction diffusion equation into a Langevin equation yields a rather complex expression Further, we know nothing about $K_c(\boldsymbol{r})$ or $K_d(\boldsymbol{r})$, and this procedure is based on very heuristic models.

Nevertheless, the form derived is essentially correct in that it agrees with the results arising from a more microscopic approach based on Master equations, which, however, specifies all arbitrary constants precisely.

8.1.1 Functional Fokker-Planck Equations

By writing a *stochastic partial differential equation* such as (8.1.12), we immediately raise the question: what does the corresponding Fokker-Planck equation look like? It must be a partial differential equation in a continuously infinite number of variables $\rho(\boldsymbol{r})$, where $\boldsymbol{r}$ is the continuous index which distinguishes the various variables. A simple-minded way of defining functional derivatives is as follows. First, divide space into cubic cells of side l labelled i with position $\boldsymbol{r}_i$, and introduce the variables

$$x_i = l^3 \rho(\boldsymbol{r}_i) \tag{8.1.15}$$

and consider functions of the variables $\boldsymbol{x} = \{x_i\}$.

We now consider calculus of functions $F(\boldsymbol{x})$ of all these cell variables. Partial derivatives are easily defined in the usual way and we formally introduce the functional derivative by

$$\frac{\delta F(\rho)}{\delta \rho(\boldsymbol{r}_i)} = \lim_{l \to 0} l^{-3} \frac{\partial F(\boldsymbol{x})}{\partial x_i}. \tag{8.1.16}$$

In what sense this limit exists is, in most applied literature, left completely undefined. Precise definitions can be given and, as is usual in matters dealing with functionals, the precise definition of convergence is important. Further, the "obvious" definition (8.1.16) is not used.

The precise formulation of functional calculus is not within the scope of this book, but an indication of what is normally done by workers who write such equations is appropriate. Effectively, the functional derivative is formally defined by (8.1.16) and a corresponding discretised version of the stochastic differential equation such as (8.1.12) is formulated. Using the same notation, this would be

$$dx_i = [\sum_i D_{ij}x_j + \tilde{F}(x_i)]\,dt + \sum_j \tilde{g}_{ij}dW_j(t)\,. \tag{8.1.17}$$

In this equation, D_{ij} are coefficients which yield a discretised approximation to $D\nabla^2$. The coefficients $\tilde{F}$ and $\tilde{g}$ are chosen so that

$$F[\rho(\boldsymbol{r}_i, t)] = \lim_{l\to 0} \tilde{F}(x_i)l^{-3} \tag{8.1.18}$$

$$g(\boldsymbol{r}_i, t) = \lim_{l\to 0} l^{-3} \sum_j \tilde{g}_{ij}dW_j(t)\,. \tag{8.1.19}$$

More precisely, we assume a more general correlation formula than (8.1.14), i.e.,

$$\langle g(\boldsymbol{r}, t)g(\boldsymbol{r}', t)\rangle = G(\boldsymbol{r}, \boldsymbol{r}')\delta(t - t')\,, \tag{8.1.20}$$

and require

$$G(\boldsymbol{r}_i, \boldsymbol{r}_j) = \lim_{l\to 0} l^{-6} \sum_k \tilde{g}_{ik}\tilde{g}_{jk}\,. \tag{8.1.21}$$

In this case, the FPE for x_i is

$$\partial_t P(\boldsymbol{x}) = -\sum_{ij} \frac{\partial}{\partial x_i}\{[D_{ij}x_j + \delta_{ij}\tilde{F}(x_i)]P(\boldsymbol{x})\} + \sum_{ijk} \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j}\tilde{g}_{ik}\tilde{g}_{jk}P(\boldsymbol{x})\,. \tag{8.1.22}$$

Now consider the limit $l^3 \to 0$. Some manipulation gives

$$\partial_t P(\rho) = -\int d^3\boldsymbol{r}\,\frac{\delta}{\delta\rho(\boldsymbol{r})}\{[D\nabla^2(\boldsymbol{r}) + F[\rho(\boldsymbol{r})]]P(\rho)\}$$
$$+ \tfrac{1}{2}\iint d^3\boldsymbol{r}\,d^3\boldsymbol{r}'\left[\frac{\delta^2}{\delta\rho(\boldsymbol{r})\delta\rho(\boldsymbol{r}')}G(\boldsymbol{r}, \boldsymbol{r}')P(\rho)\right]. \tag{8.1.23}$$

$P(\rho)$ is now a kind of functional probability and the definition of its normalisation requires a careful statement of the probability measure on $\rho(\boldsymbol{r})$. This can be done [8.1] but what is normally understood by (8.1.23) is really the discrete version (8.1.22), and almost all calculations implicitly discretise.

The situation is clearly unsatisfactory. The formal mathematical existence of stochastic partial differential equations and their solutions has now been established, but as an everyday computational tool this has not been developed. We refer the reader to [8.1] for more information on the mathematical formulation. Since, however, most work is implicitly discretised, we will mostly formulate

matters directly in a discretised form, using continuum notations simply as a convenience in order to give a simpler notation.

8.2 Multivariate Master Equation Description

8.2.1 Diffusion

We assume that the space is divided into cubic *cells* of volume ΔV and side length l. The cells are labelled by an index i and the number of molecules of a chemical X inside cell i is called x_i. Thus we introduce a multivariate probability

$$P(\boldsymbol{x}, t) \equiv P(x_1, x_2, \ldots, t) \equiv P(x_i, \hat{\boldsymbol{x}}, t)\,. \tag{8.2.1}$$

In the last expression, $\hat{\boldsymbol{x}}$ means the vector of all x's not explicitly written.

We can model diffusion as a Markov process in which a molecule is transferred from cell i to cell j with probability per unit time $d_{ij}x_i$, i.e., the probability of transfer is proportional to the number of molecules in the cell. For a strictly *local* description, we expect that d_{ij} will be nonzero only when i and j are neighbouring cells, but this is not necessary and will not always be assumed in what follows.

In terms of the notation of Sect. 7.5, we can write a birth-death Master equation with parameters given by the replacements:

$$\begin{aligned}
&\qquad\qquad\quad i \qquad\quad j\\
\boldsymbol{N}^{(i,j)} &= (0 \ldots 0,\ 1, 0, \ldots 0, 0, 0, \ldots)\\
\boldsymbol{M}^{(i,j)} &= (0 \ldots 0,\ 0, 0, \ldots 0, 1, 0, \ldots)\\
\boldsymbol{r}^{(i,j)} &= (0 \ldots 0, -1, 0, \ldots 0, 1, 0, \ldots)\\
k^{+}_{(i,j)} &= d_{ij}\\
k^{-}_{(i,j)} &= 0\,.
\end{aligned} \tag{8.2.2}$$

Hence, the Master equation becomes

$$\partial_t P(\boldsymbol{x}, t) = \sum d_{ij}[(x_i + 1)P(\hat{\boldsymbol{x}}, x_i + 1, x_j - 1, t) - x_i P(\boldsymbol{x}, t)]\,. \tag{8.2.3}$$

This equation is a simple linear Master equation and can be solved by various means.

Notice that since

$$\begin{aligned}
\boldsymbol{r}^{(i,j)} &= -\boldsymbol{r}^{(j,i)}\\
\boldsymbol{N}^{(i,j)} &= \boldsymbol{M}^{(j,i)}\,,
\end{aligned} \tag{8.2.4}$$

we can also restrict to $i > j$ and set $k^{-}_{(i,j)} = d_{ji}$. From (7.5.15,18) we see that in this form, *detailed balance* is satisfied provided

$$d_{ij}\langle x_i\rangle_s = d_{ji}\langle x_j\rangle_s\,. \tag{8.2.5}$$

In a system which is diffusing, the stationary solution is homogenous, i.e., $\langle x_i \rangle_s = \langle x_j \rangle_s$.

Hence, detailed balance requires

$$d_{ij} = d_{ji}\,, \tag{8.2.6}$$

and (8.2.3) possesses a multivariate Poisson stationary solution.

The mean-value equation is

$$\begin{aligned} d_t\langle x_i(t)\rangle &= \sum_{j,k} r_i^{(j,k)}[\langle t_{jk}^{+}(\boldsymbol{x}) - t_{jk}^{-}(\boldsymbol{x})\rangle] \\ &= \sum_{j,k} (-\delta_{ij} + \delta_{ik})\, d_{jk}\langle x_j\rangle\,. \end{aligned}$$

Hence,

$$d_t\langle x_i(t)\rangle = \sum_j (d_{ji} - \delta_{ji} \sum_k d_{jk})\langle x_j(t)\rangle \tag{8.2.7}$$

$$\equiv \sum_j D_{ji}\langle x_j(t)\rangle\,. \tag{8.2.8}$$

8.2.2 Continuum Form of Diffusion Master Equation

Suppose the centre of cell i is located at $\boldsymbol{r}_i$ and we make the replacement

$$x_i(t) = l^3\rho(\boldsymbol{r}_i, t) \tag{8.2.9}$$

and assume that $d_{ij} = 0$ (i, j not nearest neighbours)
$= d$ (i, j adjacent).

Then (8.2.8) becomes, in the limit $l \to 0$,

$$\partial_t\langle\rho(\boldsymbol{r}, t)\rangle = D\nabla^2\langle\rho(\boldsymbol{r}, t)\rangle \quad \text{with} \tag{8.2.10}$$

$$D = l^2 d\,. \tag{8.2.11}$$

Thus, the diffusion equation is recovered. We will generalise this result shortly.

a) Kramers-Moyal or System Size Expansion Equations

We need a parameter in terms of which the numbers and transition probabilities scale appropriately. There are two limits which are possible, both of which correspond to increasing numbers of molecules:

i) limit of large cells: $l \to \infty$, at fixed concentration;
ii) limit of high concentration at fixed cell size.

The results are the same for pure diffusion. In either case,

$$t_{i,j}^{+}(\boldsymbol{x}) \to \infty \tag{8.2.12}$$

and a system size expansion is possible. To lowest order, this will be equivalent to a Kramers-Moyal expansion. From (7.5.31,32) we find

$$A_l(\boldsymbol{x}) = \sum_j D_{jl} x_j \tag{8.2.13}$$

$$B_{lm}(\boldsymbol{x}) = \delta_{lm} \sum_j (D_{lj} x_l + D_{jl} x_j) - D_{lm} x_l - D_{ml} x_m \,, \tag{8.2.14}$$

where

$$D_{jl} = d_{jl} - \delta_{jl} \sum_k d_{lk} \tag{8.2.15}$$

and thus, in this limit, $P(\boldsymbol{x}, t)$ obeys the Fokker-Planck equation

$$\partial_t P = -\sum_l \partial_l A_l(\boldsymbol{x}) P + \tfrac{1}{2} \sum_{l,m} \partial_l \partial_m B_{lm}(\boldsymbol{x}) P \,. \tag{8.2.16}$$

b) Continuum Form of Kramers-Moyal Expansion

The continuum form is introduced by associating a point $\boldsymbol{r}$ with a cell i and writing

$$l^3 \sum_i \to \int d^3\boldsymbol{r} \tag{8.2.17}$$

$$D_{ji} \to D(\boldsymbol{r}', \boldsymbol{r}) \equiv \mathscr{D}(\boldsymbol{r}', \boldsymbol{r} - \boldsymbol{r}') \tag{8.2.18}$$

$$l^{-3} \delta_{ij} \to \delta(\boldsymbol{r} - \boldsymbol{r}') \,. \tag{8.2.19}$$

At this stage we make no particular symmetry assumptions on D_{ij}, etc, so that anisotropic inhomogeneous diffusion is included.

However, there are some requirements brought about by the meaning of the concept "diffusion."

i) *Diffusion is observed only when a concentration gradient exists.* This means that the stationary state corresponds to constant concentration and from (8.2.13,15), this means that

$$\sum_j D_{jl} = 0 \,, \tag{8.2.20}$$

i.e.,

$$\sum_j d_{jl} = \sum_j d_{lj} \,. \tag{8.2.21}$$

Note that detailed balance (8.2.6) implies these.

ii) *Diffusion does not change the total amount of substance in the system,* i.e.,

$$\frac{d}{dt} \sum_l x_l = 0 \tag{8.2.22}$$

and this must be true for any value of x_l. From the equation for the mean values, this requires

$$\sum_j D_{lj} = 0 \tag{8.2.23}$$

which follows from (8.2.15) and (8.2.21)

iii) *In the continuum notation,* (8.2.20) implies that for any $\boldsymbol{r}$,

$$\int d^3\boldsymbol{\delta}\, \mathscr{D}(\boldsymbol{r}+\boldsymbol{\delta}, -\boldsymbol{\delta}) = 0 \tag{8.2.24}$$

and from (8.2.23), we also have

$$\int d^3\boldsymbol{\delta}\, \mathscr{D}(\boldsymbol{r}, \boldsymbol{\delta}) = 0\,. \tag{8.2.25}$$

iv) *If detailed balance is true*, (8.2.24) is replaced by the equation obtained by substituting (8.2.6) in the definition of D, i.e.,

$$D_{ij} = D_{ji} \tag{8.2.26}$$

which gives in the continuum form

$$\mathscr{D}(\boldsymbol{r}+\boldsymbol{\delta}, -\boldsymbol{\delta}) = \mathscr{D}(\boldsymbol{r}, \boldsymbol{\delta})\,. \tag{8.2.27}$$

The derivation of a continuum form now follows in a similar way to that of the Kramers-Moyal expansion.

We define the derivate moments

$$\boldsymbol{M}(\boldsymbol{r}) = \int d^3\boldsymbol{\delta}\, \boldsymbol{\delta}\mathscr{D}(\boldsymbol{r}, \boldsymbol{\delta}) \tag{8.2.28}$$

$$\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r}) = \tfrac{1}{2}\int d^3\boldsymbol{\delta}\, \boldsymbol{\delta}\boldsymbol{\delta}\mathscr{D}(\boldsymbol{r}, \boldsymbol{\delta}) \tag{8.2.29}$$

and it is assumed that derivate moments of higher order vanish in some appropriate limit, similar to those used in the Kramers-Moyal expansion.

The detailed balance requirement (8.2.27) gives

$$\begin{aligned}\boldsymbol{M}(\boldsymbol{r}) &= \int d^3\boldsymbol{\delta}\, \boldsymbol{\delta}\mathscr{D}(\boldsymbol{r}+\boldsymbol{\delta}, -\boldsymbol{\delta}) \\ &= \int d^3\boldsymbol{\delta}\, \boldsymbol{\delta}[\mathscr{D}(\boldsymbol{r}, -\boldsymbol{\delta}) + \boldsymbol{\delta}\cdot\nabla\mathscr{D}(\boldsymbol{r}, -\boldsymbol{\delta}) + \ldots] \\ &= -\boldsymbol{M}(\boldsymbol{r}) + 2\nabla\cdot\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r}) + \ldots\,.\end{aligned} \tag{8.2.30}$$

Hence, detailed balance requires

$$\boldsymbol{M}(\boldsymbol{r}) = \nabla\cdot\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r})\,. \tag{8.2.31}$$

The weaker requirement (8.2.24) similarly requires the weaker condition

$$\nabla\cdot[\boldsymbol{M}(\boldsymbol{r}) - \nabla\cdot\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r})] = 0\,. \tag{8.2.32}$$

We now can make the continuum form of $A_l(\boldsymbol{x})$:

$$\begin{aligned}A_l(\boldsymbol{x}) &\to \int d^3\boldsymbol{\delta}\, \mathscr{D}(\boldsymbol{r}, \boldsymbol{\delta})\rho(\boldsymbol{r}+\boldsymbol{\delta}) \\ &= \boldsymbol{M}(\boldsymbol{r})\cdot\nabla\rho(\boldsymbol{r}) + \underset{\sim}{\boldsymbol{D}}(\boldsymbol{r}) : \nabla\nabla\rho(\boldsymbol{r})\end{aligned} \tag{8.2.33}$$

If detailed balance is true, we can rewrite, from (8.2.31)

$$\boxed{A_l(x)\to\nabla\cdot[\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r})\cdot\nabla\rho(\boldsymbol{r})]\,.} \tag{8.2.34}$$

The *general* form, without detailed balance, can be obtained by defining

$$\boldsymbol{J}(\boldsymbol{r}) = \boldsymbol{M}(\boldsymbol{r}) - \nabla \cdot \underset{\sim}{\boldsymbol{D}}(\boldsymbol{r}) \tag{8.2.35}$$

from (8.2.32)

$$\nabla \cdot \boldsymbol{J}(\boldsymbol{r}) = 0 \tag{8.2.36}$$

so that we can write

$$\boldsymbol{J}(\boldsymbol{r}) = \nabla \cdot \underset{\sim}{\boldsymbol{E}}(\boldsymbol{r}) , \tag{8.2.37}$$

where $\underset{\sim}{\boldsymbol{E}}(\boldsymbol{r})$ is an antisymmetric tensor. Substituting, we find that by defining

$$\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r}) = \underset{\sim}{\boldsymbol{D}}(\boldsymbol{r}) + \underset{\sim}{\boldsymbol{E}}(\boldsymbol{r}) , \tag{8.2.38}$$

we have defined a nonsymmetric diffusion tensor $\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r})$ and that

$$A_l(\boldsymbol{x}) \rightarrow \nabla \cdot [\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r}) \cdot \nabla \rho(\boldsymbol{r}, t)] . \tag{8.2.39}$$

This means that, deterministically,

$$\boxed{\partial_t \rho(\boldsymbol{r}, t) = \nabla \cdot [\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r}) \cdot \nabla \rho(\boldsymbol{r}, t)]} \tag{8.2.40}$$

where $\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r})$ is symmetric, if detailed balance holds.

We now come to the fluctuation term, given by (8.2.14). To compute $B_{lm}(\boldsymbol{x})$, we first consider the limit of

$$l^3 \sum_m B_{lm} \phi_m \rightarrow \int d\boldsymbol{r}' B(\boldsymbol{r}, \boldsymbol{r}') \phi(\boldsymbol{r}') \tag{8.2.41}$$

where ϕ_m is an arbitrary function. By similar, but much more tedious computation, we eventually find

$$l^3 \sum_m B_{lm} \phi_m \rightarrow -2\nabla \cdot [\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r}) \rho(\boldsymbol{r}) \cdot \nabla \phi(\boldsymbol{r})] \tag{8.2.42}$$

so that

$$\boxed{B(\boldsymbol{r}, \boldsymbol{r}') = 2\nabla' \nabla : [\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r}) \rho(\boldsymbol{r}) \delta(\boldsymbol{r} - \boldsymbol{r}')] .} \tag{8.2.43}$$

The phenomenological theory of Sect. 8.1 now has a rational basis since (8.2.43) is what arises from assuming

$$\boldsymbol{j}(\boldsymbol{r}, t) = -\underline{\boldsymbol{H}}(\boldsymbol{r})\nabla\rho(\boldsymbol{r}, t) - \xi(\boldsymbol{r}, t) \tag{8.2.44}$$

in which

$$\langle\xi(\boldsymbol{r}, t)\xi(\boldsymbol{r}', t')\rangle = 2\delta(t - t')\underline{\boldsymbol{D}}(\boldsymbol{r})\delta(\boldsymbol{r} - \boldsymbol{r}')\rho(\boldsymbol{r}) \tag{8.2.45}$$

and hence,

$$\partial_t\rho(\boldsymbol{r}, t) = \nabla\cdot\underline{\boldsymbol{H}}(\boldsymbol{r})\cdot\nabla\rho + \nabla\cdot\xi(\boldsymbol{r}, t)\,. \tag{8.2.46}$$

This corresponds to a theory of inhomogeneous anisotropic diffusion without detailed balance. This is usually simplified by setting

$$\underline{\boldsymbol{D}}(\boldsymbol{r}) = D\mathbf{1} \tag{8.2.47}$$

and this gives a more familiar equation. Notice that according to (8.2.45), fluctuations in different components of the current are in general correlated, unless $\underline{\boldsymbol{D}}$ is diagonal.

Comparison with Fluctuation-Dissipation Argument. The result (8.2.43) can almost be obtained from a simple fluctuation-dissipation argument in the stationary state, where we know the fluctuations are Poissonian. In that case,

$$\langle x_i, x_j\rangle = \langle x_i\rangle\delta_{ij} \tag{8.2.48}$$

corresponding to

$$g(\boldsymbol{r}, \boldsymbol{r}') = \langle\rho(\boldsymbol{r}), \rho(\boldsymbol{r}')\rangle = \delta(\boldsymbol{r} - \boldsymbol{r}')\langle\rho(\boldsymbol{r})\rangle\,.$$

Since the theory is linear, we can apply (4.4.51) of Sect. 4.4.6. Here the matrices A and A^{T} become

$$\begin{aligned} A &\to -\nabla\cdot\underline{\boldsymbol{H}}(\boldsymbol{r})\cdot\nabla \\ A^{\mathrm{T}} &\to -\nabla'\cdot\underline{\boldsymbol{H}}(\boldsymbol{r}')\cdot\nabla'\,. \end{aligned} \tag{8.2.49}$$

Thus,

$$\begin{aligned} &B(\boldsymbol{r}, \boldsymbol{r}') \to BB^{\mathrm{T}} \\ &\quad = A\sigma + \sigma A^{\mathrm{T}} \to [-\nabla\cdot\underline{\boldsymbol{H}}(\boldsymbol{r})\cdot\nabla - \nabla'\cdot\underline{\boldsymbol{H}}(\boldsymbol{r}')\cdot\nabla']g(\boldsymbol{r}, \boldsymbol{r}')\,. \end{aligned} \tag{8.2.50}$$

We note that in the stationary state, $\langle\rho(\boldsymbol{r})\rangle = \langle\rho\rangle$, independent of $\boldsymbol{r}$. Thus,

$$\begin{aligned} B(\boldsymbol{r}, \boldsymbol{r}') &= [-\nabla\cdot\underline{\boldsymbol{H}}(\boldsymbol{r})\cdot\nabla\langle\rho\rangle\delta(\boldsymbol{r} - \boldsymbol{r}') - \nabla'\cdot\underline{\boldsymbol{H}}(\boldsymbol{r}')\cdot\nabla'\langle\rho\rangle\delta(\boldsymbol{r} - \boldsymbol{r}')] \\ &= \nabla\nabla' : \{[\underline{\boldsymbol{H}}(\boldsymbol{r}) + \underline{\boldsymbol{H}}^{\mathrm{T}}(\boldsymbol{r})]\langle\rho\rangle\delta(\boldsymbol{r} - \boldsymbol{r}')\} \\ &= 2\nabla\nabla' : [\underline{\boldsymbol{D}}(\boldsymbol{r})\langle\rho\rangle\delta(\boldsymbol{r} - \boldsymbol{r}')]\,. \end{aligned} \tag{8.2.51}$$

However, this result is not as general as (8.2.43) since it is valid in the stationary state only, nor is it the same in the stationary state because it includes $\langle\rho\rangle$, not

$\rho(\boldsymbol{r})$. However, since the Fokker-Planck formalism is valid only as a large cell size limit in which fluctuations are small, to this accuracy (8.2.51) agrees with (8.2.43)

8.2.3 Reactions and Diffusion Combined

We introduce reactions by assuming that molecules within a cell react with each other according to a Master equation like those of Chap. 7. We wish to consider several chemical components so we introduce the notation

$$\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots) \tag{8.2.52}$$

where $\boldsymbol{x}_i$ represents a vector whose components $x_{i,a}$ are the numbers of molecules of species X_a in cell i. Thus, we write

$$\begin{aligned}\partial_t P(\boldsymbol{X}, t) = \partial_t P(\text{diffusion}) + \sum_i \{\sum_A t_A^+(\boldsymbol{x}_i - \boldsymbol{r}^A)P(\boldsymbol{x}_i - \boldsymbol{r}^A, \hat{\boldsymbol{X}}) \\ + t_A^-(\boldsymbol{x}_i + \boldsymbol{r}^A)P(\boldsymbol{x}_i + \boldsymbol{r}^A, \hat{\boldsymbol{X}}) - [t_A^+(\boldsymbol{x}_i) + t_A^-(\boldsymbol{x}_i)]P(\boldsymbol{X})\}\end{aligned} \tag{8.2.53}$$

where the diffusion part has the form of (8.2.3), but is summed over the various components.

This leads via a Kramers-Moyal expansion to a Fokker-Planck equation, with the usual drift and diffusion as given by (7.5.32).

We can write equivalent stochastic partial differential equations in terms of a spatially dependent Wiener process $\boldsymbol{W}(\boldsymbol{r}, t)$ as follows: we consider an isotropic constant diffusion tensor

$$\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r}) = D\mathbf{1} \tag{8.2.54}$$

$$d\rho_a = [D_a \nabla^2 \rho_a + \sum_A r_a^A(\kappa_A^+ \prod_a \rho_a^{N_a^A} - \kappa_A^- \prod_a \rho_a^{M_a^A})]dt + dW_a(\boldsymbol{r}, t) \tag{8.2.55}$$

with

$$\begin{aligned}dW_a(\boldsymbol{r}, t)dW_b(\boldsymbol{r}', t) = \{2\delta_{a,b}\nabla' \cdot \nabla[D_a\rho_a(\boldsymbol{r})\delta(\boldsymbol{r} - \boldsymbol{r}')] \\ + \delta(\boldsymbol{r} - \boldsymbol{r}') \sum_A r_a^A r_b^A(\kappa_A^+ \prod_a \rho_a^{N_a^A} + \kappa_A^- \prod_a \rho_a^{M_a^A})\}\, dt\,.\end{aligned} \tag{8.2.56}$$

The validity of the Langevin equation depends on the system size expansion. Equations (8.2.55,56) depend on the particular scaling of the chemical rate constants with Ω, given in Sect. 7.5.3 (7.5.29). The only interpretation of Ω which is valid in this case is

$$\Omega = l^3 = \text{volume of cell.}$$

Notice, however, that at the same time, the diffusion part scales like l since $l^2 d$ must remain equal to the diffusion coefficient while the terms arising from chemical reactions scale like l^3. This means that as cell volume is increased, we have less and less effect from diffusion, but still more than the correction terms in the chemical part which will be integral powers of l^3 less than the first.

The precise method of comparing diffusion with reaction will be dealt with later.

Example: $X_1 \rightleftharpoons 2X_2$: for this chemical reaction, we find (using the methods of Sect. 7.5)

$$\boldsymbol{N} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{M} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad \boldsymbol{r} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \tag{8.2.57}$$
$$k^+ = k_1 = \kappa_1 \Omega^{-1+1}, \qquad k^- = k_2 = \kappa_1 \Omega^{-2+1} .$$

Thus, substituting

$$\begin{aligned} d\rho_1(\boldsymbol{r}) &= (D_1 \nabla^2 \rho_1 - \kappa_1 \rho_1 + \kappa_2 \rho_2^2)dt + dW_1(\boldsymbol{r}, t) \\ d\rho_2(\boldsymbol{r}) &= (D_2 \nabla^2 \rho_2 + 2\kappa_1 \rho_1 - 2\kappa_2 \rho_2^2)dt + dW_2(\boldsymbol{r}, t) \end{aligned} \tag{8.2.58}$$

where $d\boldsymbol{W}(\boldsymbol{r}, t)d\boldsymbol{W}^{\mathrm{T}}(\boldsymbol{r}', t) =$

$$\begin{bmatrix} 2\nabla \cdot \nabla'[D_1 \rho_1 \delta(\boldsymbol{r}-\boldsymbol{r}')] + \delta(\boldsymbol{r}-\boldsymbol{r}')[\kappa_1 \rho_1 + \kappa_2 \rho_2^2], & -2\delta(\boldsymbol{r}-\boldsymbol{r}')[\kappa_1 \rho_1 + \kappa_2 \rho_2^2] \\ -2\delta(\boldsymbol{r}-\boldsymbol{r}')[\kappa_1 \rho_1 + \kappa_2 \rho_2^2], & 2\nabla \cdot \nabla'[D_2 \rho_2 \delta(\boldsymbol{r}-\boldsymbol{r}')] + 4\delta(\boldsymbol{r}-\boldsymbol{r}')[\kappa_1 \rho_1 + \kappa_2 \rho_2] \end{bmatrix} dt . \tag{8.2.59}$$

This equation is valid only as a system size expansion in Ω^{-1}, that is, the cell size, and the continuum limit is to be regarded as a *notation* in which it is understood that we really mean a cell model, and are working on a sufficiently large scale for the cell size to appear small, though the cell itself is big enough to admit many molecules.

Thus, this kind of equation is really only valid as a linearised equation about the deterministic state which is the form in which *Keizer* [8.2] has formulated chemical reactions. In this respect, the Poisson representation is better since it gives equations exactly equivalent to the Master equation.

8.2.4 Poisson Representation Methods

Corresponding to a reaction with no more than bimolecular steps, we have from (7.7.9) a rather simplified Fokker-Planck equation since for the spatial diffusion [using the formulation (8.2.2)], we find the *diffusion matrix vanishes*. The generalisation to a spatially dependent system is then carried out in the density variable

$$\eta_a(\boldsymbol{r}) = \alpha_a(\boldsymbol{r})/l^3 \tag{8.2.60}$$

and we find

$$d\eta_a(\boldsymbol{r}) = [D_a \nabla^2 \eta_a(\boldsymbol{r}) + \sum_A r_a^A (\kappa_A^+ \prod_a \eta_a^{N_a^A} - \kappa_A^- \prod_a \eta_a^{M_a^A})]dt + dW_a(\boldsymbol{r}, t) \tag{8.2.61}$$

$$\begin{aligned} dW_a(\boldsymbol{r}, t)dW_b(\boldsymbol{r}', t) = dt\, \delta(\boldsymbol{r}-\boldsymbol{r}') \sum_A (\kappa_A^+ \prod_a \eta_a^{N_a^A} - \kappa_A^- \prod_a \eta_a^{M_a^A}) \\ \times (M_a^A M_b^A - N_a^A N_b^A - \delta_{ab} r_a^A) . \end{aligned} \tag{8.2.62}$$

These equations are very similar to (8.2.55,56). When explicitly written out they are simpler. For example, considering again $X_1 \rightleftharpoons 2X_2$, we get

$$d\eta_1(\boldsymbol{r}) = (D_1\nabla^2\eta_1 - \kappa_1\eta_1 + \kappa_2\eta_2^2)dt + dW_1(\boldsymbol{r}, t) \tag{8.2.63}$$

$$d\eta_2(\boldsymbol{r}) = (D_2\nabla^2\eta_2 + 2\kappa_1\eta_1 - 2\kappa_2\eta_2^2)dt + dW_2(\boldsymbol{r}, t) \tag{8.2.64}$$

$$d\boldsymbol{W}(\boldsymbol{r}, t)d\boldsymbol{W}^{\mathrm{T}}(\boldsymbol{r}', t) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} (\kappa_1\eta_1 - \kappa_2\eta_2^2)\delta(\boldsymbol{r} - \boldsymbol{r}')dt\,. \tag{8.2.65}$$

The simplicity of (8.2.63–65) when compared to their counterparts (8.2.57, 58) is quite striking, and it is especially noteworthy that they are exactly equivalent (in a continuum formulation) to the Master equation.

8.3 Spatial and Temporal Correlation Structures

We want to consider here various aspects of spatial, temporal and spatio-temporal correlations in linear systems, which are of course all exactly soluble. The correlations that are important are the *factorial correlations* which are defined in terms of factorial moments in the same way as ordinary correlations are defined in terms of moments. The equations which arise are written much more naturally in terms of factorial moments, as we shall see in the next few examples.

8.3.1 Reaction $X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} Y$

We assume homogenous isotropic diffusion with the same diffusion constant for X and Y, and since both the reaction and the diffusion are linear we find Poisson representation Langevin equations for the *concentration* variables η, μ (corresponding, respectively, to X and Y) *with no stochastic source*, i.e.,

$$\begin{aligned} \partial_t\eta(\boldsymbol{r}, t) &= D\nabla^2\eta - k_1\eta + k_2\mu \\ \partial_t\mu(\boldsymbol{r}, t) &= D\nabla^2\mu + k_1\eta - k_2\mu\,. \end{aligned} \tag{8.3.1}$$

a) Spatial Correlations
We now note that

$$\begin{aligned} &\langle\eta(\boldsymbol{r}, t)\rangle = \langle\rho_x(\boldsymbol{r}, t)\rangle \\ &\langle\mu(\boldsymbol{r}, t)\rangle = \langle\rho_y(\boldsymbol{r}, t)\rangle \\ &\langle\eta(\boldsymbol{r},t),\ \eta(\boldsymbol{r}', t)\rangle = \langle\rho_x(\boldsymbol{r},t),\ \rho_x(\boldsymbol{r}', t)\rangle - \delta(\boldsymbol{r} - \boldsymbol{r}')\langle\rho_x(\boldsymbol{r}, t)\rangle \equiv g(\boldsymbol{r}, \boldsymbol{r}', t) \\ &\langle\eta(\boldsymbol{r},t),\ \mu(\boldsymbol{r}', t)\rangle = \langle\rho_x(\boldsymbol{r},t),\ \rho_y(\boldsymbol{r}', t)\rangle \equiv f(\boldsymbol{r}, \boldsymbol{r}', t) \\ &\langle\mu(\boldsymbol{r},t),\ \mu(\boldsymbol{r}', t)\rangle = \langle\rho_y(\boldsymbol{r},t),\ \rho_x(\boldsymbol{r}', t)\rangle - \delta(\boldsymbol{r} - \boldsymbol{r}')\langle\rho_y(\boldsymbol{r}, t)\rangle = h(\boldsymbol{r}, \boldsymbol{r}', t)\,, \end{aligned} \tag{8.3.2}$$

which are all continuum notation versions of the fact that the Poissonian moments are equal to the factorial moments of the actual numbers.

The equations for the mean concentrations are obviously exactly the same as (8.3.1). Assuming now a homogeneous situation, so that we can assume

$$\left.\begin{aligned} &\langle \rho_x(\boldsymbol{r}, t)\rangle = \langle \rho_x(t)\rangle \\ &\langle \rho_y(\boldsymbol{r}, t)\rangle = \langle \rho_y(t)\rangle \\ &g(\boldsymbol{r}, \boldsymbol{r}', t) = g(\boldsymbol{r} - \boldsymbol{r}', t) \\ &f(\boldsymbol{r}, \boldsymbol{r}', t) = f(\boldsymbol{r} - \boldsymbol{r}', t) \\ &h(\boldsymbol{r}, \boldsymbol{r}', t) = h(\boldsymbol{r} - \boldsymbol{r}', t) \end{aligned}\right\} \tag{8.3.3}$$

and compute equations of motion for $\langle \eta(\boldsymbol{r}, t)\eta(\boldsymbol{0}, t)\rangle$, etc, we quickly find

$$\begin{aligned} \frac{\partial g(\boldsymbol{r}, t)}{\partial t} &= 2D\nabla^2 g(\boldsymbol{r}, t) - 2k_1 g(\boldsymbol{r}, t) + 2k_2 f(\boldsymbol{r}, t) \\ \frac{\partial f(\boldsymbol{r}, t)}{\partial t} &= 2D\nabla^2 f(\boldsymbol{r}, t) - (k_1 + k_2) f(\boldsymbol{r}, t) + k_2 h(\boldsymbol{r}, t) + k_1 g(\boldsymbol{r}, t) \\ \frac{\partial h(\boldsymbol{r}, t)}{\partial t} &= 2D\nabla^2 h(\boldsymbol{r}, t) - 2k_2 h(\boldsymbol{r}, t) + 2k_1 f(\boldsymbol{r}, t)\,, \end{aligned} \tag{8.3.4}$$

The stationary solution of these equations has the form

$$g(\boldsymbol{r}) = \xi k_2^2, \qquad f(\boldsymbol{r}) = \xi k_1 k_2, \qquad h(\boldsymbol{r}) = \xi k_1^2 \tag{8.3.5}$$

where ξ is an arbitrary parameter. The corresponding stationary solutions for the means are

$$\langle x(\boldsymbol{r}_i)\rangle = \lambda k_2, \qquad \langle y(\boldsymbol{r}_i)\rangle = \lambda k_1 \tag{8.3.6}$$

where λ is another arbitrary parameter. If $\xi = 0$, we recover the Poissonian situation where

$$\begin{aligned} &\langle \rho_x(\boldsymbol{r}), \rho_x(\boldsymbol{r}')\rangle = \langle \rho_x(\boldsymbol{r})\rangle\, \delta(\boldsymbol{r} - \boldsymbol{r}') \\ &\langle \rho_y(\boldsymbol{r}), \rho_y(\boldsymbol{r}')\rangle = \langle \rho_y(\boldsymbol{r})\rangle\, \delta(\boldsymbol{r} - \boldsymbol{r}') \\ &\langle \rho_x(\boldsymbol{r}), \rho_y(\boldsymbol{r}')\rangle = 0\,. \end{aligned} \tag{8.3.7}$$

(By choosing other values of λ, different solutions corresponding to various distributions over the total number of molecules in the system are obtained).

Time-dependent solutions for any initial condition can easily be developed. In the case where the solutions are initially homogeneous, uncorrelated and Poissonian, (8.3.7) is satisfied as an initial condition and thus f, g, and h are initially all zero, and will remain so. Thus, an uncorrelated Poissonian form is preserved in time, as has already been deduced in Sect. 7.7b.

The problem of relaxation to the Poisson is best dealt with by assuming a specific form for the initial correlation function For example, an initially uncorrelated but non-Poissonian system represented by

$$g(\boldsymbol{r}, 0) = \alpha\, \delta(\boldsymbol{r}), \qquad f(\boldsymbol{r}, 0) = \beta\, \delta(\boldsymbol{r}), \qquad h(\boldsymbol{r}, 0) = \gamma\, \delta(\boldsymbol{r}). \tag{8.3.8}$$

Time-dependent solutions are

$$\begin{bmatrix} g(\mathbf{r}, t) \\ f(\mathbf{r}, t) \\ h(\mathbf{r}, t) \end{bmatrix} = \frac{\exp(-r^2/8Dt)}{(8\pi Dt)^{3/2}} \begin{bmatrix} k_2^2\varepsilon_1 - 2k_2\varepsilon_2 e^{-(k_1+k_2)t} + \varepsilon_3 e^{-2(k_1+k_2)t} \\ k_1k_2\varepsilon_1 + (k_2 - k_1)\varepsilon_2 e^{-(k_1+2)t} - \varepsilon_3 e^{-2(k_1+k_2)t} \\ k_1^2\varepsilon_1 + 2k_1\varepsilon_2 e^{-(k_1+k_1)t} + \varepsilon_3 e^{-2(k_1+k_2)t} \end{bmatrix} \tag{8.3.9}$$

where

$$\begin{aligned} \varepsilon_1 &= (\alpha + 2\beta + \gamma)/(k_1 + k_2)^2 \\ \varepsilon_2 &= [k_2(\beta + \gamma) - k_1(\alpha + \beta)]/(k_1 + k_2)^2 \\ \varepsilon_3 &= [k_1^2\alpha + k_2^2\gamma - 2k_1k_2\beta]/(k_1 + k_2)^2 . \end{aligned} \tag{8.3.10}$$

Comments

i) The terms ε_1, ε_2, and ε_3 correspond, respectively, to deviations from an uncorrelated Poissonian of the quantities $\langle (x_i + y_i), (x_j + y_j)\rangle$, $\langle (x_i + y_i), (k_1y_j - k_2x_j)\rangle$, and $\langle (k_1y_i - k_2x_i), (k_1y_j - k_2x_j)\rangle$, which are essentially density fluctuations, correlation between density fluctuation and chemical imbalance, and fluctuations in chemical imbalance. We notice a characteristic diffusion form multiplying a chemical time dependence appropriate to the respective terms.

ii) The time taken for the deviation from a Poissonian uncorrelated form given by (8.3.8) to become negligible compared to the Poissonian depends, of course, on the magnitude of the initial deviation. Assuming, however, that α, β, γ, $\langle x_i\rangle$, and $\langle y_i\rangle$ are all of comparable size, one can make a rough estimate as follows. We consider a small spherical volume of radius R much larger, however, than our basic cells. Then in this small volume V, we find that

$$\text{var}\{x[V, 0]\} = \int_V d^3\mathbf{r} \int_V d^3\mathbf{r}' \langle \rho_x(\mathbf{r}, 0), \rho_x(\mathbf{r}', 0)\rangle , \qquad \text{i.e.,}$$

$$\text{var}\{x[V, 0]\} = \langle x[V, 0]\rangle + \alpha V \tag{8.3.11}$$

and similarly,

$$\begin{aligned} \text{var}\{y[V, 0]\} &= \langle y[V, 0]\rangle + \gamma V \\ \langle x[V, 0], y[V, 0]\rangle &= \beta V \end{aligned}$$

while after a time $t \gg R^2/4D$, these quantities satisfy approximately

$$\left.\begin{aligned} \text{var}\{x[V, t]\} &\simeq \langle x[V, t]\rangle \\ &\quad + \frac{V^2}{(8\pi Dt)^{3/2}}(k_1^2\varepsilon_1 - 2k_2\varepsilon_2 e^{-(k_1+k_2)t} + \varepsilon_3 e^{-2(k_1+k_2)t}) \\ \text{var}\{y[V, t]\} &\simeq \langle y[V, t]\rangle \\ &\quad + \frac{V^2}{(8\pi Dt)^{3/2}}(k_1^2\varepsilon_1 + 2k_1\varepsilon_2 e^{-(k_1+k_2)t} + \varepsilon_3 e^{-2(k_1+k_2)t}) \\ \langle x[V, t], y[V, t]\rangle &\simeq \frac{V^2}{(8\pi Dt)^{3/2}} \\ &\quad \times [k_1k_2\varepsilon_1 + (k_2 - k_1)\varepsilon_2 e^{-(k_1+k_2)t} - \varepsilon_3 e^{-2(k_1+k_2)t}] . \end{aligned}\right\} \tag{8.3.12}$$

Thus, the diffusion has reduced the overall deviation from Poissonian uncorrelated behaviour by a factor of the order of magnitude of $R^3/(Dt)^{3/2}$. However, notice that in the case of an initial non-Poissonian, but also uncorrelated, situation, corresponding to $\beta = 0$, we find that a correlation has appeared between X and Y, which if the chemical rate constants are sufficiently large, can be quite substantial.

b) Space-Time Correlations

Since the equations of motion here are linear, we may use the linear theory developed in Sect. 3.7.4. Define the stationary two-time correlation matrix $\underline{G}(\boldsymbol{r}, t)$ by

$$\underline{G}(\boldsymbol{r}, t) = \begin{bmatrix} \langle \rho_x(\boldsymbol{r}, t), \rho_x(\boldsymbol{0}, 0)\rangle_s & \langle \rho_y(\boldsymbol{r}, t), \rho_x(\boldsymbol{0}, 0)\rangle \\ \langle \rho_x(\boldsymbol{r}, t), \rho_y(\boldsymbol{0}, 0)\rangle_s & \langle \rho_y(\boldsymbol{r}, t), \rho_y(\boldsymbol{0}, 0)\rangle \end{bmatrix}. \tag{8.3.13}$$

Then the equation corresponding to (3.7.63) is

$$\partial_t \underline{G}(\boldsymbol{r}, t) = \begin{bmatrix} D\nabla^2 - k_1 & k_2 \\ k_1 & D\nabla^2 - k_2 \end{bmatrix} \underline{G}(\boldsymbol{r}, t) . \tag{8.3.14}$$

The solution can be obtained by Fourier transforming and solving the resultant 1st order differential matrix equation by standard methods, with boundary conditions at $t = 0$ given by (3.3.7). The result is

$$\underline{G}(\boldsymbol{r}, t) = \left(\frac{\langle \rho_x \rangle + \langle \rho_y \rangle}{(k_1 + k_2)^2}\right) \frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}} \begin{bmatrix} k_2^2 + k_1 k_2 e^{-(k_1+k_2)t} & k_1 k_2 (1 - e^{-(k_1+k_2)t}) \\ k_1 k_2 (1 - e^{-(k_1+k_2)t}) & k_1^2 + k_1 k_2 e^{-(k_1+k_2)t} \end{bmatrix}. \tag{8.3.15}$$

If we define variables

$$n(\boldsymbol{r}, t) = \rho_x(\boldsymbol{r}, t) + \rho_y(\boldsymbol{r}, t) \tag{8.3.16}$$

$$c(\boldsymbol{r}, t) = [k_1 \rho_x(\boldsymbol{r}, t) - k_2 \rho_y(\boldsymbol{r}, t)]/(k_1 + k_2) , \tag{8.3.17}$$

the solution (8.3.15) can be written as

$$\langle n(\boldsymbol{r}, t), n(\boldsymbol{0}, 0)\rangle_s = \langle n \rangle_s \frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}} \tag{8.3.18}$$

$$\langle n(\boldsymbol{r}, t), c(\boldsymbol{0}, 0)\rangle_s = \langle c(\boldsymbol{r}, t), n(\boldsymbol{0}, 0)\rangle_s = 0 \tag{8.3.19}$$

$$\langle c(\boldsymbol{r}, t), c(\boldsymbol{0}, 0)\rangle_s = \frac{k_1 k_2 \langle n \rangle_s}{(k_1 + k_2)^2} \frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}} e^{-(k_1+k_2)t} . \tag{8.3.20}$$

The variables n and c correspond to total density and chemical imbalance density, i.e., $\langle c \rangle_s = 0$. Thus we see that (8.3.18) gives the correlation for density fluctuations which is the same as that arising from pure diffusion: (8.3.20) gives the correlation of fluctuations in chemical imbalance, and (8.3.19) shows that these are independent. A characteristic diffusion term multiplies all of these. The simplicity of this result depends on the identity of diffusion constants for the different species.

8.3.2 Reactions $B + X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} C$, $A + X \xrightarrow[k_2]{} 2X$

This reaction has already been treated in Sect. 7.6.4a without spatial dependence. We find the Poisson representation equation for the concentration variable $\eta(\boldsymbol{r}, t)$ is [from (8.2.60–62), (7.6.55,56)]

$$d\eta(\boldsymbol{r}, t) = [D\nabla^2\eta(\boldsymbol{r}, t) + (K_2 - K_1)\eta(\boldsymbol{r}, t) + K_3]dt + dW(\boldsymbol{r}, t) \tag{8.3.21}$$

$$dW(\boldsymbol{r}, t)dW(\boldsymbol{r}', t) = 2dt\,\delta(\boldsymbol{r} - \boldsymbol{r}')K_2\eta(\boldsymbol{r}, t)\,, \tag{8.3.22}$$

where

$$\left.\begin{aligned} k_3C &= K_3l^3 \\ k_1B &= K_1 \\ k_2A &= K_2\,. \end{aligned}\right\} \tag{8.2.23}$$

This system can be solved exactly since it is linear, but since the second reaction *destroys* X at a rate proportional to X, we do obtain a noise term in the Poisson representation.

[In the Kramers-Moyal method, we would find an equation of the same form as (8.3.21) for $\rho(\boldsymbol{r}, t)$, but the $dW(\boldsymbol{r}, t)$ would satisfy

$$\begin{aligned} dW(\boldsymbol{r}, t)\,dW(\boldsymbol{r}', t) = dt\,\{&2\nabla'\cdot\nabla[D\rho(\boldsymbol{r})\delta(\boldsymbol{r}-\boldsymbol{r}')] \\ &+ \delta(\boldsymbol{r}-\boldsymbol{r}')[(K_2 + K_1)\eta(\boldsymbol{r}, t) + K_3]\}]\,. \end{aligned} \tag{8.3.24}$$

a) Spatial Correlations

Define now

$$\begin{aligned} g(\boldsymbol{r}, \boldsymbol{r}', t) &= \langle\rho(\boldsymbol{r}, t),\, \rho(\boldsymbol{r}', t)\rangle - \delta(\boldsymbol{r} - \boldsymbol{r}')\langle\rho(\boldsymbol{r}, t)\rangle \\ &= \langle\eta(\boldsymbol{r}, t),\, \eta(\boldsymbol{r}', t)\rangle = \langle\eta(\boldsymbol{r}, t)\eta(\boldsymbol{r}', t)\rangle - \langle\eta(\boldsymbol{r}, t)\rangle\langle\eta(\boldsymbol{r}', t)\rangle \end{aligned} \tag{8.3.25}$$

We consider the stationary homogeneous situation in which clearly

$$\langle\eta(\boldsymbol{r}, t)\rangle_s = \langle\rho(\boldsymbol{r}, t)\rangle_s = \langle\rho\rangle_s\,. \tag{8.3.26}$$

Then,

$$\begin{aligned} dg(\boldsymbol{r}, \boldsymbol{r}', t) &= d\langle\eta(\boldsymbol{r}, t)\eta(\boldsymbol{r}', t)\rangle \\ &= \langle\eta(\boldsymbol{r}, t)d\eta(\boldsymbol{r}', t)\rangle + \langle d\eta(\boldsymbol{r}, t)\eta(\boldsymbol{r}'\,t)\rangle + \langle d\eta(\boldsymbol{r}, t)d\eta(\boldsymbol{r}', t)\rangle \end{aligned} \tag{8.3.27}$$

and using the usual Ito rules and (8.3.21,22),

$$\begin{aligned} &= [D\nabla^2 + D\nabla'^2 + 2(K_2 - K_1)]\langle\eta(\boldsymbol{r}, t)\eta(\boldsymbol{r}', t)\rangle \\ &\quad + 2K_3\langle\rho\rangle_s dt + 2K_2\delta(\boldsymbol{r} - \boldsymbol{r}')\langle\rho\rangle_s dt\,. \end{aligned} \tag{8.3.28}$$

Note that

$$\langle \rho \rangle_s = K_3/(K_1 - K_2)\,, \tag{8.3.29}$$

and that in a spatially homogeneous situation, $g(\boldsymbol{r}, \boldsymbol{r}', t)$ can only be a function of $\boldsymbol{r} - \boldsymbol{r}'$, which we call $g(\boldsymbol{r}, t)$.

Substitute using (8.3.25,26) to obtain

$$\partial_t g(\boldsymbol{r}, t) = 2[D\nabla^2 + (K_2 - K_1)]g(\boldsymbol{r}, t) + 2K_2\langle \rho \rangle_s \delta(\boldsymbol{r})\,. \tag{8.3.30}$$

The stationary solution $g_s(\boldsymbol{r})$ is best obtained by representing it as a Fourier integral:

$$g_s(\boldsymbol{r}) = \int d^3\boldsymbol{q}\, e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} \tilde{g}_s(\boldsymbol{q}) \tag{8.3.31}$$

which, on using

$$\delta(\boldsymbol{r}) = (2\pi)^{-3} \int d^3\boldsymbol{q}\, e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}\,, \tag{8.3.32}$$

gives

$$\tilde{g}_s(\boldsymbol{q}) = \frac{K_2\langle \rho \rangle_s}{(2\pi)^3} \frac{1}{Dq^2 + K_1 - K_2} \tag{8.3.33}$$

whose inverse is given by

$$g_s(\boldsymbol{r}) = \frac{K_2\langle \rho \rangle_s}{4\pi D r} \exp\left[-r\left(\frac{K_1 - K_2}{D}\right)^{1/2}\right]. \tag{8.3.34}$$

Hence,

$$\langle \rho(\boldsymbol{r}, t), \rho(\boldsymbol{r}', t) \rangle_s = \delta(\boldsymbol{r} - \boldsymbol{r}')\langle \rho \rangle_s + \frac{K_2\langle \rho \rangle_s}{4\pi D|\boldsymbol{r} - \boldsymbol{r}'|} \exp\left[-|\boldsymbol{r} - \boldsymbol{r}'|\left[\frac{K_1 - K_2}{D}\right]^{1/2}\right]. \tag{8.3.35}$$

Comments

i) We note two distinct parts: a δ correlated part which corresponds to Poissonian fluctuations, independent at different space points, and added to this, a correlation term with a characteristic correlation length

$$l_c = \sqrt{D/(K_1 - K_2)}\,. \tag{8.3.36}$$

ii) *Approach to equilibrium*: when $K_2 \to 0$, we approach a simple reversible reaction $B + X \rightleftharpoons C$; one sees that the correlation in (8.3.34) becomes zero. However, the correlation length itself does *not* vanish.

iii) *Local and global fluctuations*: a question raised originally by *Nicolis* [8.3] is the following. Consider the total number of molecules in a volume V:

$$x[V, t] = \int_V d^3\boldsymbol{r}\, \rho(\boldsymbol{r}, t)\,. \tag{8.3.37}$$

Then we would like to know what is the variance of this number. We can easily see

$$\mathrm{var}\{x(V)\}_s = \int_V d^3\mathbf{r} \int_V d^3\mathbf{r}' \langle \rho(\mathbf{r}, t), \rho(\mathbf{r}', t)\rangle_s \tag{8.3.38}$$

and using (8.3.35),

$$= \langle x(V)\rangle_s + \frac{K_2\langle\rho\rangle_s}{4\pi D} \int_V d^3\mathbf{r} \int_V d^3\mathbf{r}' \, |\mathbf{r} - \mathbf{r}'|^{-1} \exp(-|\mathbf{r} - \mathbf{r}'|/l_c) \tag{8.3.39}$$

We can compute two limits to this. If V is much larger than l^3, then, noting that $g(\mathbf{r}) \to 0$ as $\mathbf{r} \to \infty$, we integrate the stationary version of (8.3.30) and drop the surface terms arising from integrating the Laplacian by parts:

$$0 = 2(K_2 - K_1) \int_V d^3\mathbf{r} \int_V d^3\mathbf{r}' g(\mathbf{r} - \mathbf{r}') + 2K_2\langle\rho\rangle_s \int_V d^3\mathbf{r}' \,. \tag{8.3.40}$$

Thus,

$$\iint_{VV} d^3\mathbf{r}\, d^3\mathbf{r}'\, g(\mathbf{r} - \mathbf{r}') \sim \frac{K_2\langle x(V)\rangle_s}{K_1 - K_2} \tag{8.3.41}$$

so that

$$\mathrm{var}\{x(V)\}_s \sim \frac{K_1\langle x(V)\rangle_s}{K_1 - K_2} \qquad (V \gg l_c^3) \,. \tag{8.3.42}$$

The small volume limit is somewhat more tricky. However, for a spherical volume of radius $R \ll l_c$, we can neglect the exponential in (8.3.39) and

$$\begin{aligned} \iint_{VV} d^3\mathbf{r}\, d^3\mathbf{r}' \, |\mathbf{r} - \mathbf{r}'|^{-1} &= \int_V d^3\mathbf{r}' \int_0^R 4\pi r^2 dr \, d(\cos\theta)(r^2 + r'^2 - 2rr'\cos\theta)^{-1/2} \\ &= \int_V d^3\mathbf{r}' \int_0^R \frac{4\pi r}{r'} \, dr(|r + r'| - |r - r'|) \\ &= 2R^5(4\pi)^2/5 \end{aligned} \tag{8.3.43}$$

so that

$$\mathrm{var}\{x(V)\}_s \sim \langle x(V)\rangle_s \left(1 + \frac{2R^2K_2}{5D}\right) \qquad \left(V = \frac{4}{3}\pi R^3 \ll l_c^3\right). \tag{8.3.44}$$

Hence, we see in a small volume that the fluctuations are Poissonian, but in a large volume the fluctuations are non-Poissonian, and in fact the variance is exactly (7.6.72) given by the global Master equation for the same reaction.

In fact, for an arbitrary spherical volume, the variance can be evaluated directly (by use of Fourier transform methods) and is given by

$$\begin{aligned} \mathrm{var}\{x(V)\} = \langle x(V)\rangle \Bigg\{1 + \frac{3K_2 l_c^5}{2DR^3}\Bigg[1 - \left(\frac{R}{l_c}\right)^2 + \frac{2}{3}\left(\frac{R}{l_c}\right)^3 \\ - e^{-2R/l_c}\left(1 + \frac{R}{l_c}\right)^2\Bigg]\Bigg\} \end{aligned} \tag{8.3.45}$$

and this gives the more precise large volume asymptotic form

$$\text{var}\{x(V)\} \sim \langle x(V)\rangle \left(\frac{K_1}{K_1 - K_2} - \frac{3K_1 l_c}{2R(K_1 - K_2)}\right) \qquad (V \gg l_c^3)\,. \tag{8.3.46}$$

The result can also be illustrated by noting that the Fourier representation of $\langle \rho(\boldsymbol{r}, t), \rho(\boldsymbol{r}', t)\rangle$ is given by adding those of $g_s(\boldsymbol{r})$ and $\delta(\boldsymbol{r})\langle\rho\rangle_s$ and is clearly

$$\langle\rho\rangle_s(2\pi)^{-3}\frac{Dq^2 + K_1}{Dq^2 + K_1 - K_2}\,. \tag{8.3.47}$$

This means that for small $\boldsymbol{q}$, i.e., long wavelength, this is approximately

$$\langle\rho\rangle_s(2\pi)^{-3}\frac{K_1}{K_1 - K_2} \tag{8.3.48}$$

which is the Fourier transform of

$$\langle\rho\rangle_s \frac{K_1}{K_1 - K_2}\delta(\boldsymbol{r} - \boldsymbol{r}') \tag{8.3.49}$$

which corresponds to the same variance as the global Master equation.

For large $\boldsymbol{q}$, i.e., small wavelengths, we fine

$$\langle\rho\rangle_s(2\pi)^{-3} \tag{8.3.50}$$

corresponding to

$$\langle\rho\rangle_s\delta(\boldsymbol{r} - \boldsymbol{r}')\,, \tag{8.3.51}$$

i.e., Poissonian fluctuations. Physically, the difference arises from the different scaling of the diffusion and reaction parts of the master equation which was noted in Sect. 8.2.3. Thus, in a small volume the diffusion dominates, since the fluctuations arising from diffusion come about because the molecules are jumping back and forth across the boundary of V. This is a surface area effect which becomes relatively larger than the chemical reaction fluctuations, which arise from the bulk reaction, and are proportional to volume. Conversely, for larger V, we find the surface effect is negligible and only the bulk effect is important.

b) Space-Time Correlations

Since the system is linear, we can, as in Sect. 8.3.1, use the method of Sect. 3.7.4 Define

$$G(\boldsymbol{r}, t) = \langle\rho(\boldsymbol{r}, t), \rho(\boldsymbol{0}, 0)\rangle_s\,. \tag{8.3.52}$$

Then the linear equation corresponding to (3.7.63) is

$$\partial_t G(\boldsymbol{r}, t) = D\nabla^2 G(\boldsymbol{r}, t) - (K_1 - K_2)G(\boldsymbol{r}, t) \tag{8.3.53}$$

with an initial condition given by

$$G(\boldsymbol{r}, 0) = \langle\rho(\boldsymbol{r}), \rho(\boldsymbol{0})\rangle_s \tag{8.3.54}$$

which is given by (8.3.35). Representing $G(\boldsymbol{r}, t)$ as a Fourier integral

$$\tilde{G}(\boldsymbol{r}, t) = \int d^3\boldsymbol{q}\, e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}\, \tilde{G}(\boldsymbol{q}, t)\,, \tag{8.3.55}$$

(8.3.53) becomes

$$\partial_t \tilde{G}(\boldsymbol{q}, t) = -(Dq^2 + K_1 - K_2)\tilde{G}(\boldsymbol{q}, t) \tag{8.3.56}$$

whose solution is, utilising the Fourier representation (8.3.47) of the initial condition (8.3.54),

$$\tilde{G}(\boldsymbol{q}, t) = (2\pi)^{-3} \frac{Dq^2 + K_1}{Dq^2 + K_1 - K_2} \exp[-(Dq^2 + K_1 - K_1)t]\,. \tag{8.3.57}$$

If desired, this can be inverted by quite standard means though, in fact, the Fourier transform correlation function is often what is desired in practical cases and being usually easier to compute, it is favoured.

Thus, we have

$$\begin{aligned} G(\boldsymbol{r}, t) = {} & \frac{\langle\rho\rangle \exp(-r^2/4Dt - Dt/l_c^2)}{(4\pi Dt)^{3/2}} \\ & + \frac{K_2\langle\rho\rangle}{8\pi Dr}\left\{[\exp(-r/l_c)]\,\mathrm{erfc}\left[\frac{(Dt)^{1/2}}{l_c} - \frac{r}{2(Dt)^{1/2}}\right]\right. \\ & \left. + [\exp(r/l_c)]\,\mathrm{erfc}\left[\frac{(Dt)^{1/2}}{l_c} + \frac{r}{2(Dt)^{1/2}}\right]\right\}\,. \end{aligned} \tag{8.3.58}$$

For small t we find

$$G(r, t) \to \langle\rho\rangle \exp\left(\frac{-r^2}{4Dt}\right)\left[(4\pi Dt)^{-3/2} - \frac{4K_2(Dt)^{3/2}}{r^3\sqrt{\pi}}\right] + \frac{K_2\langle\rho\rangle e^{-r/l_c}}{4\pi Dr}\,, \tag{8.3.59}$$

while for large t,

$$G(r, t) \sim \left(\frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}} + \frac{K_2 l_c}{4\pi r(\pi D^2 t)^{1/2}}\right)\langle\rho\rangle \exp(-Dt/l_c^2)\,. \tag{8.3.60}$$

c) Behaviour at the Instability Point

As $K_1 \to K_2$, the reaction approaches instability, and when $K_1 = K_2$ there are no longer any stationary solutions. We see that simultaneously,

i) the correlation length $l_c \to \infty$ (8.3.36);

ii) the variance of the fluctuations in a volume $V \gg l_c^3$ will become infinite [(8.3.42)]. However, as $l_c \to \infty$ at finite V, one reaches a point at which $l_c^3 \sim V$ and (8.3.42) is no longer valid. Eventually $l_c^3 \gg V$; the volume now appears small and the fluctuations within it take the Poissonian form;

iii) the correlation time is best taken to be

$$\tau_c = l_c^2/D \tag{8.3.61}$$

being the coefficient of t in the exponent of the long-time behaviour of the space time correlation function (8.3.60) (ignoring the diffusive part). Thus, $\tau_c \to \infty$ also, near the instability. We thus have a picture of long-range correlated slow fluctuations. This behaviour is characteristic of many similar situations.

8.3.3 A Nonlinear Model with a Second-Order Phase Transition

We consider the reactions

$$\begin{aligned} A + X &\underset{k_4}{\overset{k_2}{\rightleftharpoons}} 2X \\ B + X &\underset{k_4}{\overset{k_1}{\rightleftharpoons}} C \end{aligned} \tag{8.3.62}$$

considered previously in Sects. 7.7c, 7.7.3a, 7.7.4c. This model was first introduced by *Schlögl* [8.4] and has since been treated by many others.

The deterministic equation is

$$\partial_t \rho(\boldsymbol{r}, t) = D\nabla^2 \rho + \kappa_3 + (\kappa_2 - \kappa_1)\rho - \kappa_4 \rho^2 \tag{8.3.63}$$

whose stationary solution is given by

$$\rho(\boldsymbol{r}) = \rho_s = [\kappa_2 - \kappa_1 + \sqrt{(\kappa_2 - \kappa_1)^2 + 4\kappa_4\kappa_3}]/2\kappa_4 \,. \tag{8.3.64}$$

[The κ's are as defined by (7.5.29), with $\Omega = l^3$]. For small κ_3, this stationary solution represents a transition behaviour as illustrated in Fig. 8.1.

The appropriate stochastic differential equation in the Poisson representation is

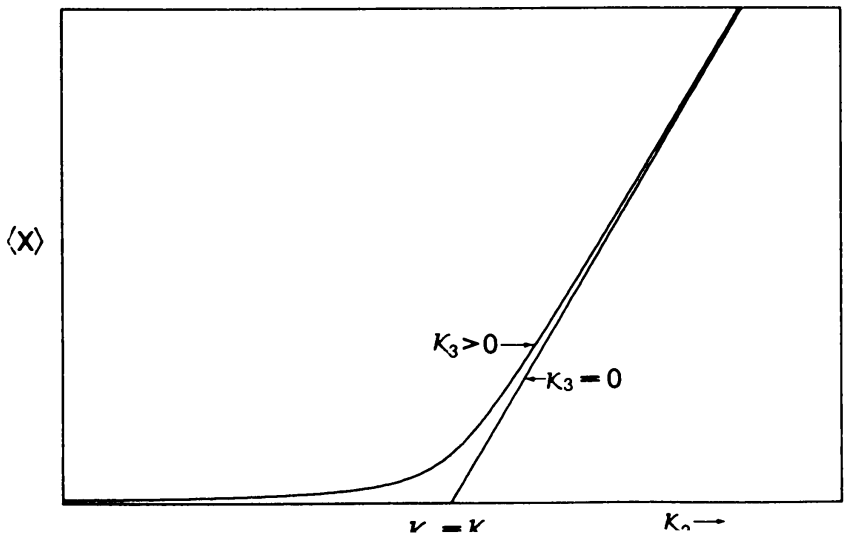


Fig. 8.1. Plot of $\langle x \rangle$ vs κ_2 for the second-order phase transition model

$$d\eta(\mathbf{r}, t) = [(\mathscr{D}\eta)(\mathbf{r}, t) + \kappa_3 + (\kappa_2 - \kappa_1)\eta(\mathbf{r}, t) - \kappa_4\eta^2(\mathbf{r}, t)]dt + \sqrt{2}\,[\kappa_2\eta(\mathbf{r}, t) - \kappa_4\eta^2(\mathbf{r}, t)]^{1/2}dW(\mathbf{r}, t)\,, \tag{8.3.65}$$

where

$$dW(\mathbf{r}, t)dW(\mathbf{r}', t) = \delta(\mathbf{r} - \mathbf{r}')dt\,. \tag{8.3.66}$$

Here we have *not* taken the continuum limit developed in Sect. 8.2.1 but have preserved a possible nonlocality of diffusion, i.e.,

$$(\mathscr{D}\eta)(\mathbf{r}, t) = -\int \mathscr{D}(|\mathbf{r} - \mathbf{r}'|)\eta(\mathbf{r}', t)d^3\mathbf{r}' \tag{8.3.67}$$

but we have considered only homogeneous isotropic diffusion.

In this form, there is no small parameter multiplying the noise term $dW(\mathbf{r}, t)$ and we are left without any obvious expansion parameter. We formally introduce a parameter λ in (8.3.65) as

$$d\eta(\mathbf{r}, t) = [(\mathscr{D}\eta)(\mathbf{r}, t) + \kappa_3 + (\kappa_2 - \kappa_1)\eta(\mathbf{r}, t) - \kappa_4\eta^2(\mathbf{r}, t)]dt + \lambda\sqrt{2}\,[\kappa_2\eta(\mathbf{r}, t) - \kappa_4\eta^2(\mathbf{r}, t)]^{1/2}dW(\mathbf{r}, t) \tag{8.3.68}$$

and expand $\eta(\mathbf{r}, t)$ in powers of λ to

$$\eta(\mathbf{r}, t) = \eta_0(\mathbf{r}, t) + \lambda\eta_1(\mathbf{r}, t) + \lambda^2\eta_2(\mathbf{r}, t) + \ldots \tag{8.3.69}$$

and set λ equal to one at the end of the calculation. However, if it is understood that all Fourier variable integrals have a cutoff l^{-1}, this will still be in fact a $(l^3)^{-1}$ expansion.

Substituting (8.3.69) in (8.3.65), we get

$$\frac{d\eta_0(\mathbf{r}, t)}{dt} = (\mathscr{D}\eta_0)(\mathbf{r}, t) + \kappa_3 + (\kappa_2 - \kappa_1)\eta_0(\mathbf{r}, t) - \kappa_4\eta_0^2(\mathbf{r}, t) \tag{8.3.70}$$

$$d\eta_1(\mathbf{r}, t) = \{(\mathscr{D}\eta_1)(\mathbf{r}, t) + [\kappa_2 - \kappa_1 - 2\kappa_4\eta_0(\mathbf{r}, t)]\eta_1(\mathbf{r}, t)\}\,dt + \sqrt{2}\,[\kappa_2\eta_0(\mathbf{r}, t) - \kappa_4\eta_0^2(\mathbf{r}, t)]^{1/2}dW(\mathbf{r}, t) \tag{8.3.71}$$

$$d\eta_2(\mathbf{r}, t) = \{(\mathscr{D}\eta_2)(\mathbf{r}, t) + [\kappa_2 - \kappa_1 - 2\kappa_4\eta_0(\mathbf{r}, t)]\eta_2(\mathbf{r}, t) - \kappa_4\eta_1^2(\mathbf{r}, t)\}\,dt + \frac{[\kappa_2 - 2\kappa_4\eta_0(\mathbf{r}, t)]\eta_1(\mathbf{r}, t)}{\sqrt{2}\,[\kappa_2\eta_0(\mathbf{r}, t) - \kappa_4\eta_0^2(\mathbf{r}, t)]^{1/2}}\,dW(\mathbf{r}, t)\,. \tag{8.3.72}$$

Equation (8.3.70) has a homogeneous steady-state solution (8.3.64) so that

$$\kappa_2 - \kappa_1 - 2\kappa_4\eta_0 \equiv \kappa = [(\kappa_1 - \kappa_2)^2 + 4\kappa_4\kappa_3]^{1/2}\,. \tag{8.3.73}$$

Substituting this in (8.3.70–72) and taking the Fourier transforms, we get

$$\eta_1(\mathbf{q}, t) = (2\kappa_1\eta_0)^{1/2}\int_0^t [\exp\{-[\mathscr{D}(q^2) + \kappa](t - t')\}]\,d\tilde{W}(\mathbf{q}, t') \tag{8.3.74}$$

$$\tilde{\eta}_2(\boldsymbol{q}, t) = -\kappa_4 \int d^3\boldsymbol{q}_1 \int_0^t dt'[\exp\{-[\mathscr{D}(q^2) + \kappa](t - t')\}]\tilde{\eta}_1(\boldsymbol{q} - \boldsymbol{q}_1, t')\tilde{\eta}_1(\boldsymbol{q}_1, t') + \frac{\kappa_1 + \kappa}{[2(\kappa_1\eta_0 - \kappa_3]^{1/2}} \int d^3\boldsymbol{q}_1 \int_0^t [\exp\{-[\mathscr{D}(q^2) + \kappa](t - t')\}]\tilde{\eta}_1(\boldsymbol{q} - \boldsymbol{q}_1, t')d\tilde{W}(\boldsymbol{q}_1, t) \tag{8.3.75}$$

where

$$d\tilde{W}(\boldsymbol{q}, t) = \left(\frac{1}{2\pi}\right)^{3/2} \int \exp(-i\boldsymbol{q}\cdot\boldsymbol{r})dW(\boldsymbol{r}, t)d^3\boldsymbol{r} \tag{8.3.76}$$

and

$$\tilde{\eta}_1(\boldsymbol{q}, t) = \left(\frac{1}{2\pi}\right)^{3/2} \int \exp(-i\boldsymbol{q}\cdot\boldsymbol{r})\eta_1(\boldsymbol{r}, t)d^3\boldsymbol{r}\,, \tag{8.3.77}$$

etc., and $\mathscr{D}(q^2)$ is the Fourier transform $\mathscr{D}(|\boldsymbol{r} - \boldsymbol{r}'|)$. We have left out the trivial initial value terms in (8.3.74, 75).

To the lowest order, the mean concentration and the correlation function are given by (in the stationary state)

$$\langle\rho(\boldsymbol{r}, t)\rangle = \eta_0 \tag{8.3.78}$$

$$\langle\rho(\boldsymbol{r}, t)\rho(\boldsymbol{r}', t)\rangle - \langle\rho(\boldsymbol{r}, t)\rangle\langle\rho(\boldsymbol{r}', t)\rangle = \eta_0\delta(\boldsymbol{r} - \boldsymbol{r}') + \langle\eta_1(\boldsymbol{r}, t)\eta_1(\boldsymbol{r}', t)\rangle\,. \tag{8.3.79}$$

From (8.3.74) it follows that

$$\langle\tilde{\eta}_1(\boldsymbol{q}, t)\tilde{\eta}_1(\boldsymbol{q}', t)\rangle = \frac{\kappa_1\eta_0\delta(\boldsymbol{q} + \boldsymbol{q}')}{[\mathscr{D}(q^2) + \kappa]}[1 - \exp\{-2[\mathscr{D}(q^2) + \kappa]t\}]\,. \tag{8.3.80}$$

Hence, the lowest-order contributions to the correlation function in the steady state are given by

$$\langle\rho(\boldsymbol{r})\rho(\boldsymbol{r}')\rangle - \langle\rho(\boldsymbol{r})\rangle\langle\rho(\boldsymbol{r}')\rangle = \eta_0\delta(\boldsymbol{r} + \boldsymbol{r}') + \frac{\kappa_1\eta_0}{(2\pi)^3} \int d^3\boldsymbol{q} \frac{\exp[i\boldsymbol{q}\cdot(\boldsymbol{r} - \boldsymbol{r}')]}{\mathscr{D}(q^2) + \kappa}\,. \tag{8.3.81}$$

If we assume that

$$(\mathscr{D}\eta)(\boldsymbol{r}) = D\nabla^2\eta(\boldsymbol{r})\,, \tag{8.3.82}$$

then

$$\mathscr{D}(q^2) = Dq^2\,. \tag{8.3.83}$$

Equation (8.3.81) then has exactly the same form as derived previously in (8.3.35), namely,

$$\langle\rho(\boldsymbol{r}),\ \rho(\boldsymbol{r}')\rangle_s = \langle\rho\rangle_0\delta(\boldsymbol{r} - \boldsymbol{r}') + \frac{\kappa_1\langle\rho_0\rangle}{4\pi D|\boldsymbol{r} - \boldsymbol{r}'|}\exp(-|\boldsymbol{r} - \boldsymbol{r}'|/l_c) \tag{8.3.84}$$

where

$$l_c = (D/\kappa)^{1/2} . \tag{8.3.85}$$

All the same conclusions will follow concerning local and global fluctuations, Poissonian and non-Poissonian fluctuations, since these are all consequences of the form (8.3.84) and not of the particular values of the parameters chosen.

Notice that if $\kappa_1 = \kappa_2$, as $\kappa_3 \to 0$ so does $\kappa \to 0$, and hence $l_c \to \infty$, and the same long-range correlation phenomena occur here as in that example.

Higher-Order Corrections: Divergence Problems. These are only lowest-order results, which can be obtained equally well by a system size expansion, Kramers-Moyal method, or by factorisation assumptions on correlation function equations.

The next order correction to the mean concentration is $\langle \eta_2(\boldsymbol{r}, t)\rangle$. Now from (8.3.75)

$$\langle \tilde{\eta}_2(\boldsymbol{q}, t)\rangle = -\int d^3\boldsymbol{q}_1 \int_0^t dt' \exp\{-[\mathscr{D}(q^2) + \kappa](t - t')\} \times \langle \eta_1(\boldsymbol{q} - \boldsymbol{q}_1, t')\eta_1(\boldsymbol{q}_1, t')\rangle . \tag{8.3.86}$$

In the steady state (8.3.86) gives

$$\langle \tilde{\eta}_2(\boldsymbol{q})\rangle = -\kappa_4\kappa_1\eta_0\delta(\boldsymbol{q}) \int \frac{d^3\boldsymbol{q}_1}{\mathscr{D}(q^2) + \kappa} . \tag{8.3.87}$$

For the choice of $\mathscr{D}$ given by (8.3.83,87) gives

$$\langle \eta_2\rangle = -\frac{\kappa_4\kappa_1\eta_0}{(2\pi)^{3/2}} \int \frac{d^3\boldsymbol{q}_1}{Dq^2 + \kappa} \tag{8.3.88}$$

and this integral is divergent. The problem lies in the continuum form chosen, for if one preserves the discrete cells at all stages, one has two modifications:

i) $\int d^3\boldsymbol{q}$ is a sum over all allowed $\boldsymbol{q}$ which have a maximum value $|\boldsymbol{q}| \sim 1/l$.

ii) $\mathscr{D}(q^2)$ is some trigonometric function of q and l. If

$$d_{ij} = 0 \quad (i \text{ not adjacent to } j)$$
$$= d \text{ otherwise,}$$

then $\mathscr{D}(q^2)$ has the form

$$\frac{4D}{l^2}\left[\sin^2\left(\frac{lq_x}{2}\right) + \sin^2\left(\frac{lq_y}{2}\right) + \sin^2\left(\frac{lq_z}{2}\right)\right].$$

In this form, no divergence arises since the sum is finite and for

$$l|\boldsymbol{q}| \ll 1 ,$$
$$\mathscr{D}(q^2) \to Dq^2 .$$

Nonlocal Reactions. The divergence, however, can also be avoided by preserving the continuum but modifying the $\eta^2(\boldsymbol{r})$ term in the equations. This term represents a collision picture of chemical reactions and will be nonlocal for a variety of reasons.
i) Molecules have a finite size, so a *strictly* local form ignores this.
ii) The time scale on which the system is Markovian is such that molecules will have travelled a certain distance in this time; the products would be the result of encounters at a variety of previous positions, and would be produced at yet another position. By making the replacement

$$\eta(\boldsymbol{r})^2 \to \int d^3\boldsymbol{r}' d^3\boldsymbol{r}'' \, m(\boldsymbol{r} - \boldsymbol{r}', \boldsymbol{r} - \boldsymbol{r}'')\eta(\boldsymbol{r}')\eta(\boldsymbol{r}'') \,, \tag{8.3.89}$$

one finds that instead of (8.3.88) one has

$$\langle \eta_2 \rangle = -\frac{\kappa_4 \kappa_1 \eta_0}{(2\pi)^{3/2}} \int \frac{d^3\boldsymbol{q}\, \tilde{m}(\boldsymbol{q}, -\boldsymbol{q})}{D\boldsymbol{q}^2 + \kappa} \tag{8.3.90}$$

where $\tilde{m}(\boldsymbol{q}, \boldsymbol{q}')$ is the fourier transform of $m(\boldsymbol{r}, \boldsymbol{r}')$. If m is sufficiently nonlocal, at high $\boldsymbol{q}$, $\tilde{m}(\boldsymbol{q}, -\boldsymbol{q})$ will decay and $\langle \eta_2 \rangle$ will be finite.

8.4 Connection Between Local and Global Descriptions

We saw in Sect. 8.3.2 that the variance of the fluctuations in a volume element V is Poissonian for small V and for sufficiently large V, approaches that corresponding to the master equation without diffusion for the corresponding reaction. This arises because the reactions give rise to fluctuations which add to each other whereas diffusion, as a mere transfer of molecules from one place to another, has an effect on the fluctuations in a volume which is effective only on the surface of the volume.

There is a precise theorem which expresses the fact that if diffusion is fast enough, the molecules in a volume V will travel around rapidly and meet each other very frequently. Hence, any molecule will be equally likely to meet any other molecule. The results are summarised by *Arnold* [8.5] and have been proved quite rigorously.

Basically, the method of proof depends on adiabatic elimination techniques as developed in Chap. 5 and can be easily demonstrated by Poisson representation methods.

8.4.1 Explicit Adiabatic Elimination of Inhomogeneous Modes

We suppose we are dealing with a single chemical species which diffuses and reacts according to the cell model and thus has a Poisson representation Fokker-Planck equation:

$$\frac{\partial P}{\partial t} = \left\{ \sum_{i,j} \left[-\frac{\partial}{\partial x_i} D_{ij} x_j \right] + \sum_i \left[\frac{\partial}{\partial x_i} a(x_i) + \frac{\partial^2}{\partial x_i^2} b(x_i) \right] \right\} P \,. \tag{8.4.1}$$

We introduce the eigenvectors $f_i(\boldsymbol{q})$ of D_{ij}, which satisfy

$$\sum_j D_{ij} f_j(q) = -D\lambda(q) f_i(q) . \tag{8.4.2}$$

The precise nature of the eigenfunctions is not very important and will depend on D_{ij}. For the simplest form, with transfer only between adjacent cells and with reflecting boundaries at the walls of the total system, assumed to be one dimensional and of length $L = Nl$ (with l the cell length), one has

$$f_j(q) \propto \cos(qlj) \tag{8.4.3}$$

and the reflecting boundary condition requires that q has the form

$$q = \frac{n\pi}{Nl} \qquad (n = 0, 1, \ldots, N) \tag{8.4.4}$$

and

$$\lambda(q) = 4 \sin^2 \left(\frac{ql}{2}\right) \Big/ l^2 . \tag{8.4.5}$$

Appropriate modifications must be made to take care of more dimensions, but the basic structure is the same, namely,

$$\begin{aligned} &\lambda(0) = 0 \\ &\lambda(q) > 0 \qquad (q \neq 0) \qquad \text{and} \end{aligned} \tag{8.4.6}$$

$$f_i(0) = \text{constant} = N^{-1/2} . \tag{8.4.7}$$

This last result represents the homogeneous stationary state of diffusion with the normalisation $N^{1/2}$ fixed by the completeness and orthogonality relations

$$\begin{aligned} &\sum_j f_j(q) f_j(q') = \delta_{q,q'} , \\ &\sum_q f_j(q) f_j(q) = \delta_{i,j} . \end{aligned} \tag{8.4.8}$$

We now introduce the variables

$$x(q) = \sum_j f_j(q) x_j . \tag{8.4.9}$$

The variable of interest is proportional to $x(0)$, since

$$x(0) = N^{-1/2} \sum_j x_j \tag{8.4.10}$$

is proportional to the total amount of substance in the system. The other variables are to be adiabatically eliminated. In anticipation of an appropriate choice of operators L_1, L_2, L_3, we define

$$y(q) = x(q)\sqrt{D} \qquad (q \neq 0) \tag{8.4.11}$$

$$x = x(0)/\sqrt{N} = N^{-1}\sum_j x_j\,. \tag{8.4.12}$$

The various terms in the FPE can now be written

$$\sum_{i,j}\frac{\partial}{\partial x_i}D_{ij}x_j = -D\sum_q \lambda(q)\frac{\partial}{\partial y(q)}y(q). \tag{8.4.13}$$

Define now

$$v_i = \sum_{q\neq 0} f_i(q)y(q) \tag{8.4.14}$$

then

$$\sum_i \frac{\partial}{\partial x_i}a(x_i) = \frac{1}{N}\sum_i \frac{\partial}{\partial x}a\left(x + \frac{1}{\sqrt{D}}v_i\right) + \sqrt{D}\sum_{q\neq 0}\frac{\partial}{\partial y(q)}\sum_i f_i(q)a\left(x + \frac{1}{\sqrt{D}}v_i\right) \tag{8.4.15}$$

$$\sum_i \frac{\partial^2}{\partial x_i^2}b(x_i) = \frac{1}{N^2}\sum_i \frac{\partial^2}{\partial x^2}b\left(x + \frac{1}{\sqrt{D}}v_i\right) + \frac{1}{N}\sqrt{D}\frac{\partial}{\partial x}\sum_{\substack{i\\ q\neq 0}}\frac{\partial}{\partial y(q)}f_i(q)b\left(x + \frac{1}{\sqrt{D}}v_i\right) + D\sum_{\substack{i\\ q',q\neq 0}}\frac{\partial^2}{\partial y(q)\partial y(q')}f_i(q)f_i(q')b\left(x + \frac{1}{\sqrt{D}}v_i\right). \tag{8.4.16}$$

We now write this in decreasing powers of $\sqrt{D}$. Thus, define

$$DL_1 = D\sum_{q\neq 0}\left[-\lambda(q)\frac{\partial}{\partial y(q)}y(q) + \frac{\partial^2}{\partial y(q)^2}b(x)\right] \tag{8.4.17}$$

which is the coefficient of D in an expansion of the terms above (lower-order terms are absorbed in L_2). We also set

$$L_3 = \left\langle \sum_i \left[\frac{1}{N}\frac{\partial}{\partial x}a\left(x + \frac{1}{\sqrt{D}}v_i\right) + \frac{1}{N^2}\frac{\partial^2}{\partial x^2}b\left(x + \frac{1}{\sqrt{D}}v_i\right)\right]\right\rangle \tag{8.4.18}$$

where the average is over the stationary distribution of L_1. As D becomes large,

$$L_3 = \left[\frac{\partial}{\partial x}a(x) + \frac{1}{N}\frac{\partial^2}{\partial x^2}b(x)\right] + O\left(\frac{1}{\sqrt{D}}\right). \tag{8.4.19}$$

We define L_2 to order $\sqrt{D}$ which can be computed to involve only terms in which $\partial/\partial y(q)$ stands to the left. Thus, in carrying out an adiabatic elimination of the

$y(q)$ variables as $D \to \infty$ to lowest order, there will be no $PL_2L_1^{-1}L_2$ contribution and we ultimately find the FPE

$$\frac{\partial P}{\partial t} = \left[\frac{\partial}{\partial x} a(x) + \frac{1}{N}\frac{\partial^2}{\partial x^2} b(x)\right] P \tag{8.4.20}$$

which corresponds to the global Master equation, since x is by (8.4.12) a concentration. Notice that the condition for validity is that

$$D\lambda(1) \gg K \tag{8.4.21}$$

where K represents a typical rate of the chemical reaction, or noticing that $\lambda(1) \simeq (\pi/Nl^2)$,

$$D \gg \frac{KN^2l^2}{\pi^2} \tag{8.4.22}$$

or

$$\sqrt{\frac{D}{K}} \gg \frac{L}{\pi} . \tag{8.4.23}$$

The left-hand side, by definition of diffusion, represents the root mean square distance travelled in the time scale of the chemical reaction, and the inequality says this must be very much bigger than the length of the system. Thus, diffusion must be able to homogenise the system. In [8.9] this result has been extended to include homogenisation on an arbitrary scale.

8.5 Phase-Space Master Equation

The remainder of this chapter deals with a stochastic version of kinetic theory and draws on the works of *van Kampen* [8.6] and *van den Broek* et al. [8.7].

We consider a gas in which molecules are described in terms of their positions $\boldsymbol{r}$ and velocities $\boldsymbol{v}$. The molecules move about and collide with each other. When a collision between two molecules occurs, their velocities are altered by this collision but their positions are not changed. However, between the collisions, the particles move freely with constant velocity.

Thus, there are two processes—collision and flow. Each of these can be handled quite easily in the absence of the other, but combining them is not straightforward.

8.5.1 Treatment of Flow

We suppose there is a large number of particles of mass m with coordinates $\boldsymbol{r}_n$ and velocities $\boldsymbol{v}_n$, which do not collide but move under the influence of an *external* force field $m\boldsymbol{A}(\boldsymbol{r})$. The equations of motion are

$$\begin{aligned} \dot{\boldsymbol{r}}_n &= \boldsymbol{v}_n \\ \dot{\boldsymbol{v}}_n &= \boldsymbol{A}(\boldsymbol{r}_n) . \end{aligned} \tag{8.5.1}$$

Then a phase-space density can be defined as

$$f(\boldsymbol{r}, \boldsymbol{v}, t) = \sum_n \delta(\boldsymbol{r} - \boldsymbol{r}_n(t))\delta(\boldsymbol{v} - \boldsymbol{v}_n(t)) \tag{8.5.2}$$

so that

$$\partial_t f(\boldsymbol{r}, \boldsymbol{v}, t) = \sum_n [\dot{\boldsymbol{r}}_n \cdot \nabla \delta(\boldsymbol{r} - \boldsymbol{r}_n(t))\delta(\boldsymbol{v} - \boldsymbol{v}_n(t)) + \delta(\boldsymbol{r} - \boldsymbol{r}_n(t))\boldsymbol{v}_n \cdot \nabla_{\boldsymbol{v}} \delta(\boldsymbol{v} - \boldsymbol{v}_n(t))]$$

and using the properties of the delta function and the equations of motion (8.5.1), we get

$$\boxed{[\partial_t + \boldsymbol{v} \cdot \nabla + \boldsymbol{A}(\boldsymbol{r}) \cdot \nabla_{\boldsymbol{v}}] f(\boldsymbol{r}, \boldsymbol{v}, t) = 0.} \tag{8.5.3}$$

This is a deterministic flow equation for a phase-space density. If the particles are distributed according to some initial probability distribution in position and velocity, (8.5.3) is unaltered but $f(\boldsymbol{r}, \boldsymbol{v}, t)$ is to be interpreted as the average of $f(\boldsymbol{r}, \boldsymbol{v}, t)$ as defined in (8.5.2) over the initial positions and velocities.

Equation (8.5.3) is exact. The variable $f(\boldsymbol{r}, \boldsymbol{v}, t)$ can be regarded as a random variable whose time development is given by this equation, which can be regarded as a stochastic differential equation with zero noise. The number of particles in a phase cell, i.e., a six-dimensional volume element of volume Δ_i centred on the phase-space coordinate $(\boldsymbol{r}_i, \boldsymbol{v}_i)$ is, of course,

$$X(\boldsymbol{r}_i, \boldsymbol{v}_i) = \int_{\Delta_i} d^3\boldsymbol{r}\, d^3\boldsymbol{v}\, f(\boldsymbol{r}, \boldsymbol{v}) . \tag{8.5.4}$$

8.5.2 Flow as a Birth-Death Process

For the purpose of compatibility with collisions, represented by a birth-death Master equation, it would be very helpful to be able to represent flow as a birth-death process in the cells Δ_i. This cannot be done for arbitrary cell size but, in the limit of vanishingly small cell size, there is a birth-death representation of any flow process.

To illustrate the point, consider a density $\rho(r, t)$ in a one-dimensional system which obeys the flow equation

$$\partial_t \rho(r, t) = -\kappa \partial_r \rho(r, t) . \tag{8.5.5}$$

This deterministic equation is the limit of a discrete equation for x_i, defined as

$$x_i(t) = \int_{\Delta_i} dr\, \rho(r, t) \simeq \lambda \rho(r_i, t) \tag{8.5.6}$$

where Δ_i is an interval of length λ around a cell point at r_i. The flow equation is then the limit as $\lambda \to 0$ of

$$\partial_t x_i(t) = \frac{\kappa}{\lambda} [x_{i-1}(t) - x_i(t)] , \tag{8.5.7}$$

i.e.,

$$\lambda \partial_t \rho(r_i, t) = \frac{\kappa}{\lambda} [\lambda \rho(r_i - \lambda, t) - \lambda \rho(r_i, t)] \tag{8.5.8}$$

whose limit as $\lambda \to 0$ is the flow equation (8.5.5). A stochastic version is found by considering particles jumping from cell i to cell $i + 1$ with transition probabilities per unit time:

$$\begin{aligned} t^+(\boldsymbol{x}) &= \kappa x_i / \lambda \\ t^-(\boldsymbol{x}) &= 0 \,. \end{aligned} \tag{8.5.9}$$

This is of the form studied in Sect. 7.5, whose notation (7.5.4, 5) takes the form here of

$$\left.\begin{aligned} &A \to i \\ &N^i_j \to \delta_{i,j} \\ &M^i_j \to \delta_{i+1,j} \\ &r^i_j \to \delta_{j,i+1} - \delta_{j,i} \,. \end{aligned}\right\} \tag{8.5.10}$$

We consider the Kramers-Moyal expansion and show that in the limit $\lambda \to 0$, all derivate moments except those of first order vanish. The drift coefficient is given by (7.5.32), i.e.,

$$\begin{aligned} A_a(\boldsymbol{x}) &= \sum_i (\delta_{a,i+1} - \delta_{a,i}) \kappa x_i / \lambda \\ &= \frac{\kappa}{\lambda} (x_{a-1} - x_a) \end{aligned} \tag{8.5.11}$$

and the diffusion matrix

$$B_{a,b}(\boldsymbol{x}) = \frac{\kappa}{\lambda} [(\delta_{a,b} - \delta_{a-1,b}) x_{a-1} + (\delta_{a,b} - \delta_{a,b-1}) x_a)] \,. \tag{8.5.12}$$

We now set $x_i = \rho(r_i)$ as in (8.5.6) and take a small λ limit; one finds

$$A_a(\boldsymbol{x}) \to \frac{\kappa}{\lambda} [\lambda \rho(r_a - \lambda, t) - \lambda \rho(r_a, t)] \tag{8.5.13}$$

$$\to - \lambda \kappa \partial_r \rho(r, t) \tag{8.5.14}$$

and the limiting value of $B_{a,b}(\boldsymbol{x})$ is found similarly, but also using

$$\delta_{a,b} \to \lambda \, \delta(r_a - r_b) \tag{8.5.15}$$

$$B_{a,b}(\boldsymbol{x}) \to \kappa \lambda^3 \partial_r \partial_{r'} [\delta(r - r') \rho(r)] \,. \tag{8.5.16}$$

Thus, in this limit, $\rho(r, t)$ obeys a stochastic differential equation

$$d\rho(r, t) = -\kappa\partial_r\rho(r, t)dt + \lambda^{1/2}dW(r, t) \tag{8.5.17}$$

where

$$dW(r, t)dW(r', t) = \kappa\, dt\partial_r\partial_{r'}[\delta(r - r')\rho(r, t)] . \tag{8.5.18}$$

We see that in the limit $\lambda \to 0$, the noise term in (8.5.17) vanishes, leading to the deterministic result as predicted.

It is interesting to ask why this deterministic limit occurs. It is not a system size expansion in the usual sense of the word, since neither the numbers of particles x_i nor the transition probabilities $t^+(\boldsymbol{x})$ become large in the small λ limit. However, the transition probability for a single particle at r_i to jump to the next cell does become infinite in the small λ limit, so in this sense, the motion becomes deterministic.

The reader can also check that this result is independent of dimensionality of space. Thus, we can find a representation of flow in phase space which, in the limit of small cell size, does become equivalent to the flow equation (8.5.3).

Let us now consider a full phase-space formulation, including both terms of the flow equation. The cells in phase space are taken to be six-dimensional boxes with side length λ in position coordinates and ξ in velocity coordinates. We define the phase-space density in terms of the total number of molecules X in a phase cell by

$$f(\boldsymbol{r}^q, \boldsymbol{v}^q) = X(\boldsymbol{r}^q, \boldsymbol{v}^q)/\xi^3\lambda^3 . \tag{8.5.19}$$

We consider transitions of two types. For simplicity, consider these only in the x-direction and define

$$\boldsymbol{\lambda}_x = (\lambda, 0, 0) . \tag{8.5.20}$$

Type 1: Flow in position space:

$$X(\boldsymbol{r}^i, \boldsymbol{v}^i) \to X(\boldsymbol{r}^i, \boldsymbol{v}^i) - 1$$

and either

$$X(\boldsymbol{r}^i + \boldsymbol{\lambda}_x, \boldsymbol{v}^i) \to X(\boldsymbol{r}^i + \boldsymbol{\lambda}_x, \boldsymbol{v}^i) + 1 \qquad (v_x > 0) \tag{8.5.21}$$

or

$$X(\boldsymbol{r}^i - \boldsymbol{\lambda}_x, \boldsymbol{v}^i) \to X(\boldsymbol{r}^i - \boldsymbol{\lambda}_x, \boldsymbol{v}^i) + 1 \qquad (v_x < 0) .$$

Then we make the labelling $A \to (i, x, 1)$:

$$\left.\begin{aligned} N_a^{(i,x,1)} &= \delta(\boldsymbol{r}^i, \boldsymbol{r}^a)\delta(\boldsymbol{v}^i, \boldsymbol{v}^a) \\ M_a^{(i,x,1)} &= \delta(\boldsymbol{r}^i - \boldsymbol{\lambda}_x, \boldsymbol{r}^a)\delta(\boldsymbol{v}^i, \boldsymbol{v}^a) \\ r_a^{(i,x,1)} &= M_a^{(i,x,1)} - N_a^{(i,x,1)} \end{aligned}\right\} \tag{8.5.22}$$

$$t^{+}_{(i,x,1)}(\boldsymbol{X}) = \lambda^2 |v_x| X(\boldsymbol{r}^i, \boldsymbol{v}^i)/\lambda^3$$
$$t^{-}_{(i,x,1)}(\boldsymbol{X}) = 0 . \tag{8.5.23}$$

The form (8.5.23) is written to indicate explicitly that the transition probability is the product of the end surface area of the cell (λ^2), the number of particles per unit volume, and the x component of velocity, which is an approximation to the rate of flow across this cell face.

Consequently, assuming $v_x > 0$,

$$A^{(x,1)}_a = v_x[X(\boldsymbol{r}^a + \boldsymbol{\lambda}_x, \boldsymbol{v}^a) - X(\boldsymbol{r}^a, \boldsymbol{v}^a)]/\lambda \tag{8.5.24}$$
$$= \xi^3\lambda^3 v_x[f(\boldsymbol{r}^a + \boldsymbol{\lambda}_x, \boldsymbol{v}^a) - f(\boldsymbol{r}^a, \boldsymbol{v}^a)]/\lambda \tag{8.5.25}$$
$$\to \xi^3\lambda^3 v_x \frac{\partial}{\partial x} f(\boldsymbol{r}, \boldsymbol{v}) . \tag{8.5.26}$$

In a similar way to the previous example, the diffusion matrix $B_{a,b}$ can be shown to be proportional to λ and hence vanishes as $\lambda \to 0$.

Type 2: Flow in velocity space:
Define

$$\boldsymbol{\xi}_x = (\xi, 0, 0) . \tag{8.5.27}$$

Then we have

$$X(\boldsymbol{r}^i, \boldsymbol{v}^i) \to X(\boldsymbol{r}^i, \boldsymbol{v}^i) - 1$$

and either

$$X(\boldsymbol{r}^i, \boldsymbol{v}^i + \boldsymbol{\xi}_x) \to X(\boldsymbol{r}^i, \boldsymbol{v}^i + \boldsymbol{\xi}_x) + 1 \qquad (A_x > 0) \tag{8.5.28}$$

or

$$X(\boldsymbol{r}^i, \boldsymbol{v}^i - \boldsymbol{\xi}_x) \to X(\boldsymbol{r}^i, \boldsymbol{v}^i - \boldsymbol{\xi}_x) + 1 \qquad (A_x < 0) .$$

The labelling is

$$\left.\begin{aligned}
&A \to (i, x, 2)\\
&N^{(i,x,2)}_a = \delta(\boldsymbol{v}^i, \boldsymbol{v}^a)\delta(\boldsymbol{r}^i, \boldsymbol{r}^a)\\
&M^{(i,x,2)}_a = \delta(\boldsymbol{v}^i - \boldsymbol{\xi}_x, \boldsymbol{v}^a)\delta(\boldsymbol{r}^i, \boldsymbol{r}^a)\\
&r^{(i,x,2)}_a = [\delta(\boldsymbol{v}^i - \boldsymbol{\xi}_x, \boldsymbol{v}^a) - \delta(\boldsymbol{v}^i, \boldsymbol{v}^a)]\delta(\boldsymbol{r}^i, \boldsymbol{r}^a)\\
&t^{+}_{(i,x,2)}(\boldsymbol{X}) = \xi^2 |A_x(\boldsymbol{r}^i)| X(\boldsymbol{r}^i, \boldsymbol{v}^i)/\xi^3\\
&t^{-}_{(i,x,2)}(\boldsymbol{X}) = 0 .
\end{aligned}\right\} \tag{8.5.29}$$

Consequently, assuming $A_x(\boldsymbol{r}^a) > 0$, the drift coefficient is

$$A_a = [X(r^a, v^a + \xi_x)A_x(r^a) - X(r^a, v^a)A_x(r^a)]/\xi \tag{8.5.30}$$

$$= \xi^3\lambda^3 A_x(r^a)[f(r^a, v^a + \xi_x) - f(r^a, v^a)]/\xi \tag{8.5.31}$$

$$\to \xi^3\lambda^3 A_x(r)\frac{\partial}{\partial v_x} f(r, v)\,. \tag{8.5.32}$$

Again, similar reasoning shows the diffusion coefficient vanishes as $\xi \to 0$.

Putting together (8.5.26, 32), one obtains the appropriate flow equation (8.5.3) in the $\lambda, \xi, \to 0$ limit.

8.5.3 Inclusion of Collisions—the Boltzmann Master Equation

We consider firstly particles in velocity space only and divide the velocity space into cells as before. Let $X(v)$ be the number of molecules with velocity v (where the velocity is, of course, discretised).

A collision is then, represented by a "chemical reaction",

$$X(v_i) + X(v_j) \to X(v_k) + X(v_l)\,. \tag{8.5.33}$$

The collision is labelled by the index (i, j, k, l) and we have (using the notation of Sect. 7.5)

$$\begin{aligned} N_a^{ij,kl} &= \delta_{a,i} + \delta_{a,j} \\ M_a^{ij,kl} &= \delta_{a,k} + \delta_{a,l} \\ r_a^{ij,kl} &= -\delta_{a,i} - \delta_{a,j} + \delta_{a,k} + \delta_{a,l} \end{aligned} \tag{8.5.34}$$

and the transition probability per unit time is taken in the form

$$\begin{aligned} t^+_{ij,kl}(X) &= R(ij, kl)X(v_i)X(v_j) \\ t^-_{ij,kl}(X) &= 0\,. \end{aligned} \tag{8.5.35}$$

There are five collisional invariants, that is, quantities conserved during a collision, which arise from dynamics. These are:

i) the number of molecules—there are two on each side of (8.5.33);

ii) the three components of momentum: since all molecules here are the same, this means in a collision

$$v_i + v_j = v_k + v_l\,; \tag{8.5.36}$$

iii) the total energy: this means that

$$v_i^2 + v_j^2 = v_k^2 + v_l^2\,. \tag{8.5.37}$$

The quantities v, and v^2 are known as additive invariants and it can be shown that any function which is similarly additively conserved is a linear function of them (with a possible constant term) [8.8].

In all molecular collisons we have time reversal symmetry, which in this case implies

$$R(ij, kl) = R(kl, ij) \,. \tag{8.5.38}$$

Finally, because of the identity of all particles, we have

$$R(ij, kl) = R(ji, kl) = R(ij, lk), \text{ etc.} \tag{8.5.39}$$

We now have a variety of possible approaches. These are:

i) Attempt to work directly with the Master equation.

ii) *System size expansion:* we assume ξ^3, the volume in phase space of the phase cells is large, and we can write a FPE using the Kramers-Moyal expansion. From (7.5.32), we can write a FPE with a drift term

$$A_a(\boldsymbol{X}) = \sum_{i,j,k,l} (-\delta_{a,i} - \delta_{a,j} + \delta_{a,k} + \delta_{a,l}) X(\boldsymbol{v}_i) X(\boldsymbol{v}_j) R(ij, kl) \tag{8.5.40}$$

and utilising all available symmetries,

$$= 2 \sum_{j,k,l} R(aj, kl) \, [X(\boldsymbol{v}_k) X(\boldsymbol{v}_l) - X(\boldsymbol{v}_a) X(\boldsymbol{v}_j)] \,. \tag{8.5.41}$$

The diffusion coefficient can also be deduced;

$$B_{ab}(\boldsymbol{X}) = \sum_{i,j,k,l} r_a^{ij,kl} r_b^{ij,kl} \, R(ij, kl) X(\boldsymbol{v}_i) X(\boldsymbol{v}_j)$$

and again, utilising to the full all available symmetries,

$$\begin{aligned} B_{ab}(\boldsymbol{X}) = {} & 2\delta_{a,b} \sum_{j,k,l} R(aj, kl)[X(\boldsymbol{v}_a) X(\boldsymbol{v}_j) + X(\boldsymbol{v}_k) X(\boldsymbol{v}_l)] \\ & + 2 \sum_{i,j} R(ij, ab)[X(\boldsymbol{v}_i) X(\boldsymbol{v}_j) + X(\boldsymbol{v}_a) X(\boldsymbol{v}_b)] \\ & - 4 \sum_{j,l} R(aj, bl)[X(\boldsymbol{v}_a) X(\boldsymbol{v}_j) + X(\boldsymbol{v}_b) X(\boldsymbol{v}_l)] \,. \end{aligned} \tag{8.5.42}$$

These imply a stochastic differential equation

$$dX(\boldsymbol{v}_a) = \{2 \sum_{j,k,l} R(aj, kl) \, [X(\boldsymbol{v}_k) X(\boldsymbol{v}_l) - X(\boldsymbol{v}_a) X(\boldsymbol{v}_j)]\} \, dt + dW(\boldsymbol{v}_a, t) \tag{8.5.43}$$

where

$$dW(\boldsymbol{v}_a, t) dW(\boldsymbol{v}_b, t) = dt \; B_{ab}(\boldsymbol{X}) \,. \tag{8.5.44}$$

Neglecting the stochastic term, we recover the *Boltzmann Equation* for $X(\boldsymbol{v}_a)$ in a discretised form. As always, this Kramers-Moyal equation is only valid in a small noise limit which is equivalent to a system size expansion, the size being ξ^3, the volume of the momentum space cells.

iii) *Poisson representation:* the Boltzmann master equation is an ideal candidate for a Poisson-representation treatment. Using the variable $\alpha(\boldsymbol{v}_a)$ as usual, we can follow

through the results (7.7.6–9) to obtain a Poisson representation FPE with a drift term

$$A_a(\boldsymbol{\alpha}) = 2\sum_{j,k,l} R(aj, kl)\,[\alpha(v_k)\alpha(v_l) - \alpha(v_a)\alpha(v_j)] \tag{8.5.45}$$

and diffusion matrix

$$\begin{aligned} B_{ab}(\boldsymbol{\alpha}) = {} & 2\delta_{ab}\sum_{i,j,l} R(ij, kl)[\alpha(v_i)\alpha(v_j) - \alpha(v_a)\alpha(v_l)] \\ & + 2\sum_{i,j} R(ij, ab)[\alpha(v_i)\alpha(v_j) - \alpha(v_a)\alpha(v_b)] \\ & - \delta_{a,b}\sum_{j,l} R(aj, bl)[\alpha(v_a)\alpha(v_j) - \alpha(v_b)\alpha(v_l)] \end{aligned} \tag{8.5.46}$$

and the corresponding SDE is

$$d\alpha(v_a) = 2\sum_{j,k,l} R(aj, kl)\,[\alpha(v_k)\alpha(v_l) - \alpha(v_a)\alpha(v_j)] + d\tilde{W}(v_a, t) \tag{8.5.47}$$

where

$$d\tilde{W}(v_a, t)d\tilde{W}(v_b, t) = dt\, B_{ab}(\boldsymbol{\alpha})\,. \tag{8.5.48}$$

As emphasised previously, *this Poisson representation stochastic differential equation is exactly equivalent to the Boltzmann master equation assumed.* Unlike the Kramers-Moyal or system size expansions, it is valid for all sizes of velocity space cell ξ.

iv) *Stationary solution of the Boltzmann master equation:* we have chosen to write the Boltzmann master equation with $t^-_{ij,kl}(\boldsymbol{X})$ zero, but we can alternatively write

$$t^-_{ij,kl}(\boldsymbol{X}) = t^+_{ij,kl}(\boldsymbol{X}) \tag{8.5.49}$$

and appropriately divide all the R's by 2, since everything is now counted twice.

The condition for detailed balance (7.5.18) is trivially satisfied. Although we have set $t^-(ij, kl) = 0$, the reversed transition is actually given by $t^+(kl, ij)$. Hence,

$$k^+_{ij,kl} = k^-_{ij,kl} = R(ij, kl)\,, \tag{8.5.50}$$

provided the time-reversal symmetry (8.5.38) is satisfied.

Under these conditions, the stationary state has a mean $\langle \boldsymbol{X} \rangle$ satisfying

$$\alpha(v_i)\alpha(v_j) = \alpha(v_k)\alpha(v_l)\,. \tag{8.5.51}$$

This means that $\log\,[\alpha(v_i)]$ is additively conserved and must be a function of the invariants (8.5.36, 37) Hence,

$$\alpha(v) = \exp\,[-(v - U)^2/mkT]\,. \tag{8.5.52}$$

Here m is the mass of the molecules and U and kT are parameters which are of course identified with the mean velocity of the molecules, and the absolute temperature multiplied by Boltzmann's constant.

The stationary distribution is then a multivariate Poisson with mean values given by (8.5.52). The fluctuations in number are uncorrelated for different velocities.

8.5.4 Collisions and Flow Together

There is a fundamental difficulty in combining the treatment of flow and that of collisions. It arises because a stochastic treatment of flow requires infinitesimally small cells, whereas the Boltzmann master equation is better understood in terms of cells of finite size. This means that it is almost impossible to write down explicitly an exact stochastic equation for the system, except in the Poisson representation which we shall shortly come to.

To formally write a multivariate phase-space master equation is, however, straightforward when we assume we have phase-space cells of finite size $\lambda^3\xi^3$. We simply include all transitions available, i.e., those leading to flow in position space, flow in velocity space and collisions. The resultant Master equation thus includes the possible transitions specified in (8.5.22, 23, 29) and in a modified form (8.5. 34, 35). Here, however, we have collisions within each cell defined by the transition probability per unit time

$$t^{+}_{ij,kl}(\boldsymbol{X}) = \delta(\boldsymbol{r}_i, \boldsymbol{r}_j)\,\delta(\boldsymbol{r}_k, \boldsymbol{r}_l)\,\delta(\boldsymbol{r}_i, \boldsymbol{r}_k)\,R(ij, kl)X_iX_j\ . \tag{8.5.53}$$

For finite $\lambda^3\xi^3$, there will be an extra stochastic effect arising from the finite cell size as pointed out in Sect. 8.5.2, which disappears in the limit of small λ and ξ when transfer from flow is purely deterministic.

The resulting master equation is rather cumbersome to write down and we shall not do this explicitly. Most work that has been done with it has involved a system size expansion or equivalently, a Kramers-Moyal approximation. The precise limit in which this is valid depends on the system size dependence of $R(ij, kl)$. The system size in this case is the six-dimensional phase-space volume $\lambda^3\xi^3$. In order to make the deterministic equation for the density, defined by

$$f(\boldsymbol{r}_i, \boldsymbol{v}_i) = X(\boldsymbol{r}_i, \boldsymbol{v}_i)/\lambda^3\xi^3\ . \tag{8.5.54}$$

independent of cell size, $R(ij, kl)$ as defined in (8.5.53) must scale like $(\lambda^3\xi^3)^4$, i.e.

$$R(ij, kl) = \bar{R}(ij, kl)\,(\lambda^3\xi^3)^4\ . \tag{8.5.55}$$

This is interpretable as meaning that $\bar{R}(ij, kl)$ is the mean collision rate per phase-space volume element in each of the arguments.

Taking the conservation law (8.5.36, 37) into account, we can then write

$$\begin{aligned} R(ij, kl) = {} & 8\sigma[(\boldsymbol{v}_i - \boldsymbol{v}_j)^2, (\boldsymbol{v}_i - \boldsymbol{v}_j)\cdot(\boldsymbol{v}_k - \boldsymbol{v}_l)] \\ & \times \delta(\boldsymbol{v}_i^2 + \boldsymbol{v}_j^2 - \boldsymbol{v}_k^2 - \boldsymbol{v}_l^2)\,\delta(\boldsymbol{v}_i + \boldsymbol{v}_j - \boldsymbol{v}_k - \boldsymbol{v}_l) \end{aligned} \tag{8.5.56}$$

and we have assumed that σ is a function only of scalars. [The fact that σ is only a

function of $(v_i - v_j)^2$ and $(v_i - v_j)\cdot(v_k - v_l)$ is a result of scattering theory, and it follows from invariance with respect to the Galilean group of transformations, i.e., rotational invariance, and the fact that the laws of physics do not depend on the particular choice of unaccelerated frame of reference. We choose to keep the dependence in terms of scalar products for simplicity of expression in the fluctuation terms.]

a) Kramers-Moyal Expression

We now replace the summations by integrations according to

$$(\xi^3\lambda^3)\sum_j \to \int d^3\boldsymbol{r}_j d^3\boldsymbol{v}_j \tag{8.5.57}$$

and change the variables by

$$\begin{aligned} \boldsymbol{v}_j &= \boldsymbol{v}_i \\ \boldsymbol{v}_k &= \tfrac{1}{2}(\boldsymbol{p} + \boldsymbol{q}) \\ \boldsymbol{v}_l &= \tfrac{1}{2}(\boldsymbol{p} - \boldsymbol{q}) . \end{aligned} \tag{8.5.58}$$

After a certain amount of manipulation in the collision term, the deterministic equation comes out in the form [from (8.5.3.41)]

$$\begin{aligned} df(\boldsymbol{r}, \boldsymbol{v}) = \Big\{ & -\boldsymbol{v}\cdot\nabla f(\boldsymbol{r}, \boldsymbol{v}) - \boldsymbol{A}\cdot\nabla_{\boldsymbol{v}} f(\boldsymbol{r}, \boldsymbol{v}) \\ & + \int d^3\boldsymbol{v}_1 \int \frac{d^3\boldsymbol{q}}{|\boldsymbol{q}|}\, \delta(|\boldsymbol{q}| - |\boldsymbol{v} - \boldsymbol{v}_1|)\sigma[\boldsymbol{q}^2, \boldsymbol{q}\cdot(\boldsymbol{v} - \boldsymbol{v}_1)] \\ & \times [f[\boldsymbol{r}, \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_1 - \boldsymbol{q})] f[\boldsymbol{r}, \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_1 + \boldsymbol{q})] - f(\boldsymbol{r}, \boldsymbol{v})\, f(\boldsymbol{r}, \boldsymbol{v}_1)] \Big\} dt. \end{aligned} \tag{8.5.59}$$

The stochastic differential equation, arising from a Kramers-Moyal expansion, is obtained by adding a stochastic term $dW(\boldsymbol{r}, \boldsymbol{v}, t)$ satisfying

$$\begin{aligned} dW(\boldsymbol{r}, \boldsymbol{v}, t)\, dW(\boldsymbol{r}', \boldsymbol{v}', t) = \delta(\boldsymbol{r} - \boldsymbol{r}')dt \times \Big\{ & \delta(\boldsymbol{v} - \boldsymbol{v}') \int d^3\boldsymbol{v}_1 \frac{d^3\boldsymbol{q}}{|\boldsymbol{q}|}\, \delta(|\boldsymbol{q}| - |\boldsymbol{v} - \boldsymbol{v}_1|) \\ & \times \sigma(\boldsymbol{q}^2, \boldsymbol{q}\cdot(\boldsymbol{v} - \boldsymbol{v}_1)] f[\boldsymbol{r}, \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_1 - \boldsymbol{q})] f[\boldsymbol{r}, \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_1 + \boldsymbol{q})] + f(\boldsymbol{r}, \boldsymbol{v}) f(\boldsymbol{r}, \boldsymbol{v}_1)] \\ & - 2 \int d^3\boldsymbol{k}\, \delta[(\boldsymbol{v}_a - \boldsymbol{v}_b)\cdot\boldsymbol{k}]\, \sigma[(\boldsymbol{v}_a - \boldsymbol{v}_b)^2 + \tfrac{1}{4}\boldsymbol{k}^2, -(\boldsymbol{v}_a - \boldsymbol{v}_b)^2 + \tfrac{1}{4}\boldsymbol{k}^2] \\ & \times [f(\boldsymbol{v}_a) f(\boldsymbol{v}_b - \boldsymbol{k}) + f(\boldsymbol{v}_b) f(\boldsymbol{v}_a - \boldsymbol{k})] \Big\} . \end{aligned} \tag{8.5.60}$$

Using such a Kramers-Moyal method, *van den Broek* et al. [8.7] have been able to apply a Chapman-Enskog expansion and have obtained fluctuating hydrodynamics from it.

The validity of this method, which depends on the largeness of the cells for the validity of the Kramers-Moyal expansion and the smallness of the cells for the validity of the modelling of flow as a birth-death process, is naturally open to question. However, since the Chapman-Enskog method is probably equivalent to the adiabatic elimination of the variables governed by the Boltzmann collision term,

the result of adiabatic elimination is not likely to be very sensitive to the precise form of this operator. Thus, the Kramers-Moyal approximation may indeed be sufficiently accurate, even for very small cells.

b) Poisson Representation

The Poisson representation stochastic differential equation can be similarly obtained from (8.5.45, 46). We use the symbol $\phi(\boldsymbol{r}, \boldsymbol{v})$ defined by

$$\phi(\boldsymbol{r}, \boldsymbol{v}) = \alpha(\boldsymbol{r}, \boldsymbol{v})/(\lambda^3 \xi^3) \,. \tag{8.5.61}$$

We find

$$\begin{aligned} d\phi(\boldsymbol{r}, \boldsymbol{v}) = dt \Big\{ & -\boldsymbol{v}\cdot\nabla\phi(\boldsymbol{r}, \boldsymbol{v}) - \boldsymbol{A}\cdot\nabla_{\boldsymbol{v}}\phi(\boldsymbol{r}, \boldsymbol{v}) + \int d^3\boldsymbol{v}_1 \frac{d^3\boldsymbol{q}}{|\boldsymbol{q}|}\, \delta(|\boldsymbol{q}| - |\boldsymbol{v} - \boldsymbol{v}_1|) \\ & \times \sigma[\boldsymbol{q}^2, \boldsymbol{q}\cdot(\boldsymbol{v} - \boldsymbol{v}_1)]\,[f[\boldsymbol{r}, \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_1 - \boldsymbol{q})] f[\boldsymbol{r}, \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_1 + \boldsymbol{q})] \\ & - f(\boldsymbol{r}, \boldsymbol{v}) f(\boldsymbol{r}, \boldsymbol{v}_1)] \Big\} + dW(\boldsymbol{r}, \boldsymbol{v}, t), \end{aligned} \tag{8.5.62}$$

where

$$\begin{aligned} dW(\boldsymbol{r}, \boldsymbol{v}, t) dW(\boldsymbol{r}', \boldsymbol{v}', t) = \delta(\boldsymbol{r} - \boldsymbol{r}') dt \times \Big(& \delta(\boldsymbol{v} - \boldsymbol{v}') \int d^3\boldsymbol{v}_1 \frac{d^3\boldsymbol{q}}{|\boldsymbol{q}|}\, \delta(|\boldsymbol{q}| - |\boldsymbol{v} - \boldsymbol{r}_1|) \\ & \times \sigma[\boldsymbol{q}^2, \boldsymbol{q}\cdot(\boldsymbol{v} - \boldsymbol{v}_1)]\, \{\phi[\boldsymbol{r}, \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_1 - \boldsymbol{q})]\, \phi[\boldsymbol{r}, \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}_1 + \boldsymbol{q})] - \phi(\boldsymbol{r}, \boldsymbol{v})\, \phi(\boldsymbol{r}, \boldsymbol{v}_1)\} \\ & + \int \frac{d^3\boldsymbol{q}}{|\boldsymbol{q}|}\, \delta(|\boldsymbol{q}| - |\boldsymbol{v} - \boldsymbol{v}_1|)\, \sigma[\boldsymbol{q}^2, \boldsymbol{q}\cdot(\boldsymbol{v} - \boldsymbol{v}_1)] \\ & \times \{\phi[\boldsymbol{r}, \tfrac{1}{2}(\boldsymbol{r} + \boldsymbol{r}' - \boldsymbol{q})]\, \phi[\boldsymbol{r}, \tfrac{1}{2}(\boldsymbol{v} + \boldsymbol{v}' + \boldsymbol{q})] - \phi(\boldsymbol{r}, \boldsymbol{v})\, \phi(\boldsymbol{r}, \boldsymbol{v}')\} \Big). \end{aligned} \tag{8.5.63}$$

The terms in (8.5.63) correspond to the first two terms in (8.5.46). The final term gives zero contribution in the limit that $\xi^3\lambda^3 \to 0$.

As always, we emphasise that this Poisson representation stochastic differential equation is exactly equivalent to the small cell size limit of the Boltzmann Master equation with flow terms added. Equations (8.5.62, 63) have not previously been written down explicitly, and so far, have not been applied. By employing Chapman-Enskog or similar techniques, one could probably deduce exact fluctuating hydrodynamic equations.

9. Bistability, Metastability, and Escape Problems

This chapter is devoted to the asymptotic study of systems which can exist in at least two stable states, and to some closely related problems. Such systems are of great practical importance, e.g., switching and storage devices in computers are systems which have this property. So do certain molecules which can isomerise, and more recently, a large number of electronic, chemical and physical systems which demonstrate related properties in rich variety have been investigated.

The problems of interest are:

i) How stable are the various states relative to each other?
ii) How long does it take for a system to switch spontaneously from one state to another?
iii) How is the transfer made, i.e., through what path in the relevant state space?
iv) How does a system relax from an unstable state?

These questions can all be answered relatively easily for one-dimensional diffusion processes—but the extension to several, but few dimensions is only recent. The extension to infinitely many variables brings us to the field of the liquid-gas transition and similar phase transitions where the system can be in one of two phases and arbitrarily distributed in space. This is a field which is not ready to be written down systematically from a stochastic point of view, and it is not treated here.

The chapter is divided basically into three parts: single variable bistable diffusion processes, one-step birth-death bistable systems and many-variable diffusion processes. The results are all qualitatively similar, but a great deal of effort must be invested for quantitative precision.

9.1 Diffusion in a Double-Well Potential (One Variable)

We consider once more the model of Sect. 5.2.7 where the probability density $p(x, t)$ of a particle obeys the Fokker-Planck equation

$$\partial_t p(x, t) = \partial_x[U'(x)p(x, t)] + D\partial_x^2 p(x, t) . \tag{9.1.1}$$

The shape of $U(x)$ is as shown in Fig. 9.1. There are two minima at a and c and in between, a local maximum. The stationary distribution is

$$p_s(x) = \mathcal{N} \exp[-U(x)/D] \tag{9.1.2}$$

and it is this that demonstrates the bistability. Corresponding to a, c and b are

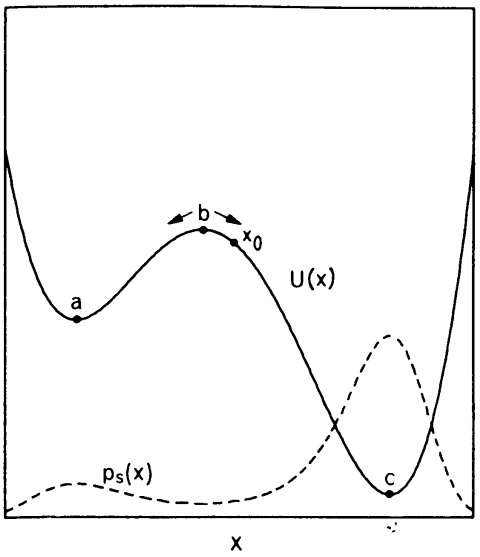


Fig. 9.1. Plot of $p_s(x)$ and $U(x)$ for a double well potential

two maxima, and a central minimum as plotted in Fig. 9.1. The system is thus most likely to be found at a or c.

9.1.1 Behaviour for $D = 0$

In this case, $x(t)$ obeys the differential equation

$$\frac{dx}{dt} = -U'(x), \qquad x(0) = x_0 \,. \tag{9.1.3}$$

Since

$$\frac{dU(x)}{dt} = U'(x)\frac{dx}{dt} = -[U'(x)]^2 < 0 \,,$$

$x(t)$ always moves in such a way as to minimise $U(x)$, and stops only when $U'(x)$ is zero. Thus, depending on whether x_0 is greater than or less than b, the particle ends up at c or a, respectively. The motion follows the arrows on the figure.

Once the particle is at a or c it stays there. If it starts exactly at b, it also stays there, though the slightest perturbation drives it to a or c. Thus, b is an unstable stationary point and a and c are stable. There is no question of *relative* stability of a and c.

9.1.2 Behaviour if D is Very Small

With the addition of noise, the situation changes. The stationary state can be approximated asymptotically as follows. Assuming $U(x)$ is everywhere sufficiently smooth, we can write

$$\begin{aligned} U(x) &\simeq U(a) + \tfrac{1}{2}U''(a)\,(x-a)^2 \qquad |x-a| \text{ small} \\ &\simeq U(c) + \tfrac{1}{2}U''(c)\,(x-c)^2 \qquad |x-c| \text{ small} \,. \end{aligned} \tag{9.1.4}$$

If D is very small, then we may approximate

$$\begin{aligned} p_s(x) &\simeq \mathcal{N} \exp\left[-U(a)/D - \tfrac{1}{2}U''(a)\,(x-a)^2/D\right] \qquad |x-a| \text{ small} \\ &\simeq \mathcal{N} \exp\left[-U(c)/D - \tfrac{1}{2}U''(c)\,(x-c)^2/D\right] \qquad |x-c| \text{ small} \\ &\simeq \quad 0 \quad \text{(elsewhere)} \end{aligned} \tag{9.1.5}$$

so that

$$\mathcal{N}^{-1} \simeq e^{-U(a)/D}\sqrt{2\pi D/U''(a)} + e^{-U(c)/D}\sqrt{2\pi D/U''(c)}\,. \tag{9.1.6}$$

Suppose, as drawn in the figure,

$$U(a) > U(c)\,. \tag{9.1.7}$$

Then for small enough D, the second term is overwhelmingly larger than the first and $\mathcal{N}^{-1}$ can be approximated by it alone. Substituting into (9.1.5) we find

$$\begin{aligned} p_s(x) &= \sqrt{\frac{U''(c)}{2\pi D}} \exp\left[-\tfrac{1}{2}U''(c)\,(x-c)^2/D\right] \qquad |x-c| \sim \sqrt{D} \\ &= 0 \qquad \text{(otherwise)}\,. \end{aligned} \tag{9.1.8}$$

This means that in the limit of very small D, the deterministic stationary state at which $U(x)$ has an absolute minimum is the more stable state in the sense that in the stochastic stationary state, $p_s(x)$ is very small everywhere except in its immediate vicinity.

Of course this result disagrees with the previous one, which stated that each state was equally stable. The distinction is one of time, and effectively we will show that the deterministic behaviour is reproduced stochastically if we start the system at x_0 and consider the limit $D \to 0$ of $p(x, t)$ for any finite t. The methods of Sect. 6.3 show this as long as the expansion about the deterministic equation is valid. Equation (6.3.10) shows that this will be the case provided $U'(x_0)$ is nonzero, or, in the case of any finite D, $U'(x_0)$ is of order D^0 (here, D replaces the ε^2 in Sect. 6.3.) This is true provided x_0 is not within a neighbourhood of width of order $D^{1/2}$ of a, c, or b. This means that in the case of a and c, fluctuations take over and the motion is given approximately by linearising the SDE around a or c. Around b, the linearised SDE is unstable. The particle, therefore, follows the Ornstein-Uhlenbeck Process until it leaves the immediate neighbourhood of $x = b$, at which stage the asymptotic expansion in $\sqrt{D}$ takes over.

However, for $t \to \infty$, the asymptotic expansion is no longer valid. Or, in other words, the $t \to \infty$ limit of the small noise perturbation theory does not reproduce the $D \to 0$ limit of the stationary state.

The process that can occur is that of escape over the central barrier. The noise $dW(t)$ can cause the particle to climb the barrier at b and reach the other side. This involves times of order $\exp(-\text{const}/D)$, which do not contribute to an asymptotic expansion in powers of D since they go to zero faster than any power of D as $D \to 0$.

9.1.3 Exit Time

This was investigated in Sect. 5.2.7c. The time for the particle, initially near a, to reach the central point b is

$$T(a \to b) = \pi[|U''(b)|\,U''(a)]^{-1/2} \exp\{[U(b) - U(a)]/D\} \tag{9.1.9}$$

(half as long as the time for the particle to reach a point well to the right of b). For $D \to 0$, this becomes exponentially large. The time taken for the system to reach the stationary state will thus also become exponentially large, and on such a time scale development of the solutions of the corresponding SDE in powers of $D^{1/2}$ cannot be expected to be valid.

9.1.4 Splitting Probability

Suppose we put the particle at x_0: what is the probability that it reaches a before c, or c before a? This can be related to the problem of exit through a particular end of an interval, studied in Sect. 5.2.8 We put absorbing barriers at $x = a$ and $x = c$, and using the results of that section, find π_a and π_c, the "*splitting probabilities*" for reaching a or c first. These are (noting that the diffusion coefficient is D, and hence independent of x):

$$\boxed{\begin{aligned} \pi_a(x_0) &= \left[\int_{x_0}^{c} dx\, p_s(x)^{-1}\right] \Big/ \left[\int_{a}^{c} dx\, p_s(x)^{-1}\right] \\ \pi_c(x_0) &= \left[\int_{a}^{x_0} dx\, p_s(x)^{-1}\right] \Big/ \left[\int_{a}^{c} dx\, p_s(x^{-1})\right]. \end{aligned}} \tag{9.1.10}$$

The splitting probabilities π_a and π_c can be viewed more generally as simply the probability that the particle, started at x_0, will fall into the left or right-hand well, since the particle, having reached a, will remain on that side of the well for a time of the same order as the mean exit time to b.

We now consider two possible asymptotic forms as $D \to 0$.

a) x_0 a Finite Distance from b

We first evaluate

$$\int_a^c dx\, p_s(x)^{-1} = \int_a^c dx\, \mathscr{N}^{-1} \exp[U(x)/D]\,. \tag{9.1.11}$$

This is dominated by the behaviour at $x \sim b$. An asymptotic evaluation is correctly obtained by setting

$$U(x) \simeq U(b) - \tfrac{1}{2}|U''(b)|/(b - x)^2\,. \tag{9.1.12}$$

As $D \to 0$, the limits at $x = a, c$ effectively recede to $\pm\infty$ and we find

$$\int_a^c dx\, p_s(x)^{-1} \sim \mathcal{N}^{-1} \sqrt{\frac{2\pi D}{|U''(b)|}} \exp\left[U(b)/D\right]. \tag{9.1.13}$$

Now suppose $x_0 < b$. Then $\int_a^{x_0} dx\, p(x)^{-1}$ can be evaluated by the substitution

$$y = U(x) \tag{9.1.14}$$

with an inverse $x = W(y)$ and is asymptotically

$$\begin{aligned} \mathcal{N}^{-1} \int_{-\infty}^{U(x_0)} e^{y/D} W'(y) dy &\sim \mathcal{N}^{-1} D\, e^{U(x_0)/D} W'[U(x_0)] \\ &= \mathcal{N}^{-1} \frac{D\, e^{U(x_0)/D}}{U'(x_0)}. \end{aligned} \tag{9.1.15}$$

Thus,

$$\pi_c \sim \frac{1}{U'(x_0)} \sqrt{\frac{|U''(b)|D}{2\pi}} \exp\left[\frac{U(x_0) - U(b)}{D}\right] \tag{9.1.16}$$

and

$$\pi_a = 1 - \pi_c. \tag{9.1.17}$$

We see here that the splitting probability depends only on x_0 and b. Thus, the probability of reaching c in this limit is governed entirely by the probability of jumping the barrier at b. The points at a and c are effectively infinitely distant.

b) x_0 Infinitesimally Distant from b

Suppose

$$x_0 = b - y_0 \sqrt{D}. \tag{9.1.18}$$

In this case, we can make the approximation (9.1.12) in both integrals. Defining [9.1]

$$\operatorname{erf}(x) = \sqrt{\frac{\pi}{2}} \int_0^x dt\, e^{-t^2}, \tag{9.1.19}$$

we find

$$\pi_c = 1 - \pi_a \sim \tfrac{1}{2}\{1 - \operatorname{erf}[y_0 \sqrt{|U''(b)|}]\} \tag{9.1.20}$$

$$= \tfrac{1}{2}\left\{1 - \operatorname{erf}\left[(b - x_0)\sqrt{\frac{|U''(b)|}{D}}\right]\right\}. \tag{9.1.21}$$

Equation (9.1.21) is the result that would be obtained if we replaced $U(x)$ by its quadratic approximation (9.1.12) over the whole range.

c) Comparison of Two Regions

The two regions give different results, and we find that a simple linearisation of the SDE [which is what replacing $U(x)$ by a quadratic approximation amounts to] gives the correct result only in the limit of large D and in a region of order of magnitude $\sqrt{D}$ around the maximum b.

9.1.5 Decay from an Unstable State

The mean time for a particle placed at a point on a potential to reach one well or the other is an object capable of being measured experimentally. If we use (9.1.1) for the process, then the mean time to reach a or c from b can be computed exactly using the formulae of Sect. 5.2.8. The mean time to reach a from b is the solution of

$$-U'(x)\partial_x[\pi_a(x)T(a,x)] + D\partial_x^2[\pi_a(x)T(a,x)] = -\pi_a(x) \tag{9.1.22}$$

with the boundary conditions

$$\pi_a(a)T(a,a) = \pi_a(c)T(a,c) = 0 \tag{9.1.23}$$

and $\pi_a(x)$ given by (9.1.10)

The solution to (9.1.22) is quite straightforward to obtain by direct integration, but it is rather cumbersome. The solution technique is exactly the same as that used for (5.2.158) and the result is similar. Even the case covered by (5.2.158) where we do not distinguish between exit at the right or at the left, is very complicated.

For the record, however, we set down that

$$T(a,x) = \frac{\pi_c(x)\int_x^c dx' p_s(x')^{-1}\int_a^{x'}\pi_a(z)p_s(z)dz - \pi_a(x)\int_a^x dx' p_s(x')^{-1}\int_a^{x'}\pi_a(z)p_s(z)dz}{D\pi_a(x)} \tag{9.1.24}$$

where one considers that $\pi_a(x)$ is given by (9.1.10) and $p_s(z)$ by (9.1.2). It can be seen that even for the simplest possible situation, namely,

$$U(x) = -\tfrac{1}{2}kx^2\,, \tag{9.1.25}$$

the expression is almost impossible to comprehend. An asymptotic treatment is perhaps required. Fortunately, in the cases where $p_s(z)$ is sharply peaked at a and c with a sharp minimum at b, the problem reduces essentially to the problem of relaxation to a or to c with a reflecting barrier at b.

To see this note that

i) the explicit solution for $\pi_a(x)$ (9.1.10) means

$$\begin{aligned} \pi_a(x) &\simeq 1 \quad (x < b) \\ &\simeq \tfrac{1}{2} \quad (x = b) \\ &\simeq 0 \quad (x > b) \end{aligned} \tag{9.1.26}$$

and the transition from 1 to 0 takes place over a distance of order $\sqrt{D}$, the width of the peak in $p_s(x)^{-1}$.

ii) In the integrals with integrand $\pi_a(z)p_s(z)$, we distinguish two cases.

$x' > b$: in this case the estimates allow us to say (9.1.27)

$$\int_a^{x'} \pi_a(z)p_s(z)dz \simeq n_a/2 , \tag{9.1.29}$$

where $n_a = \int_{-\infty}^{b} p_s(z)dz$ and represents the probability of being in the left-hand well.

However, when $x' < a$, we may still approximate $\pi_a(z)$ by 1, so we get

$$\int_a^{x'} \pi_a(z)p_s(z)dz \simeq \int_a^{x'} p_s(z)dz = \frac{n_a}{2} - \int_{x'}^{b} p_s(z)dz . \tag{9.1.30}$$

Substituting these estimates into (9.1.24) we obtain

$$T(a, b) \simeq D^{-1} \int_a^b dx' p_s(x)^{-1} \int_{x'}^b p_s(z)dz . \tag{9.1.31}$$

which is the exact mean exit time from b to a in which there is a reflecting barrier at b. Similarly,

$$T(c, b) \simeq D^{-1} \int_b^c dx' p_s(x)^{-1} \int_b^{x'} p_s(z)dz . \tag{9.1.32}$$

9.2 Equilibration of Populations in Each Well

Suppose we start the system with the particle initially in the left-hand well at some position x_i so that

$$p(x, 0) = \delta(x-x_i) . \tag{9.2.1}$$

If D is very small, the time for x to reach the centre is very long and for times small compared with the first exit time, there will be no effect arising from the existence of the well at c. We may effectively assume that there is a reflecting barrier at b.

The motion inside the left-hand well will be described simply by a small noise expansion, and thus the typical relaxation time will be the same as that of the deterministic motion. Let us approximate

$$U(x) \simeq U(a) + \tfrac{1}{2}U''(a)x^2 ;$$

then the system is now an Ornstein-Uhlenbeck process and the typical time scale is of the order of $[U''(a)]^{-1}$.

Thus, we expect a two time scale description. In the short term, the system relaxes to a quasistationary state in the well in which it started. Then on a longer time scale, it can jump over the maximum at b and the long-time bimodal stationary distribution is approached.

9.2.1 Kramers' Method

In 1940, *Kramers* [9.2] considered the escape problem from the point of view of molecular transformation. He introduced what is called the Kramers equation (Sect. 5.3.6a) in which he considered motion under the influence of a potential $V(x)$ which was double welled. In the case of large damping, he showed that a corresponding Smoluchowski equation of the form found in (6.4.18) could be used, and hence, fundamentally, the escape problem is reduced to the one presently under consideration.

Kramers' method has been rediscovered and reformulated many times [9.3]. It will be presented here in a form which makes its precise range of validity reasonably clear.

Using the notation of Fig. 9.1, define

$$M(x, t) = \int_{-\infty}^{x} dx' p(x', t) \tag{9.2.2}$$

$$N_a(t) = 1 - N_c(t) = M(b, t)$$

and (9.2.3)

$$N_0(t) = (c - a)\, p(x_0, t)\,.$$

Further, define the corresponding stationary quantities by

$$\begin{aligned} n_a &= 1 - n_c = \int_{-\infty}^{c} p_s(x')dx' \\ n_0 &= (c-a)\, p_s(x_0) \end{aligned} \tag{9.2.4}$$

From the FPE (9.1.1) and the form of $p_s(x)$ given in (9.1.2) we can write

$$d_t M(x, t) = D\, p_s(x)\partial_x[\, p(x, t)/p_s(x)] \tag{9.2.5}$$

which can be integrated to give

$$d_t \int_{a}^{x_0} dx\; M(x, t)/p_s(x) = D[\, p(x_0, t)/p_s(x_0) - p(a, t)/p_s(a)]\,. \tag{9.2.6}$$

This equation contains no approximations. We want to introduce some kind of approximation which would be valid at long times.

We are forced to introduce a somewhat less rigorous argument than desirable in order to present the essence of the method. Since we believe relaxation within each well is rather rapid, we would expect the distribution in each well (in a time of order of magnitude which is finite as $D \to 0$) to approach the same shape as the stationary distribution, but the relative weights of the two peaks to be different. This can be formalised by writing

$$\begin{aligned} p(x, t) &= p_s(x)N_a(t)/n_a \qquad x < b \\ &= p_s(x)N_c(t)/n_c \qquad x > b\,. \end{aligned} \tag{9.2.7}$$

This would be accurate to lowest order in D, except in a region of magnitude $\sqrt{D}$ around b.

If we substitute these into (9.2.6), we obtain

$$\boxed{\begin{aligned}\kappa(x_0)\dot{N}_a(t) &= D[N_0(t)/n_0 - N_a(t)/n_a]\\ \mu(x_0)\dot{N}_c(t) &= D[N_0(t)/n_0 - N_c(t)/n_c]\end{aligned}} \tag{9.2.8}$$

with

$$\begin{aligned}\kappa(x_0) &= \int_a^{x_0} p_s(x)^{-1}[1 - \psi(x)]dx\\ \mu(x_0) &= \int_{x_0}^{c} p_s(x)^{-1}[1 - \psi(x)]dx\end{aligned} \tag{9.2.9}$$

and

$$\begin{aligned}\psi(x) &= n_a^{-1}\int_x^b p_s(z)dz \qquad x < b\\ &= n_c^{-1}\int_b^x p_s(z)dz \qquad x > b\,.\end{aligned} \tag{9.2.10}$$

Note that if x is finitely different from a or c, then $\psi(x)$ vanishes exponentially as $D \to 0$, as follows directly from the explicit form of $p_s(x)$. Hence, since x in both integrals (9.2.9) satisfies this condition over the whole range of integration, we can set

$$\psi(x) = 0$$

and use

$$\boxed{\begin{aligned}\kappa(x_0) &= \int_a^{x_0} p_s(x)^{-1}\,dx\\ \mu(x_0) &= \int_{x_0}^{c} p_s(x)^{-1}\,dx\,.\end{aligned}} \tag{9.2.11}$$

a) Three State Interpretation

Equations (9.2.8) correspond to a process able to be written as a chemical reaction of the kind

$$X_a \rightleftharpoons X_0 \rightleftharpoons X_c \tag{9.2.12}$$

except that there is no equation for N_0, the number of X_0. By noting that $N_a + N_c = 1$, we find that

$$N_0(t) = n_0[\mu(x_0)N_a(t) + \kappa(x_0)N_c(t)]/[\kappa(x_0) + \mu(x_0)]\,. \tag{9.2.13}$$

This is the same equation as would be obtained by adiabatically eliminating the variable $N_0(t)$ from (9.2.8) and the further equation for $N_0(t)$

$$\dot{N}_0(t) = D\{N_a(t)/[n_a\kappa(x_0)] + N_c(t)/[n_c\mu(x_0)]\} - N_0(t)\{[n_0\kappa(x_0)]^{-1} + [n_0\mu(x_0)]^{-1}\} . \tag{9.2.14}$$

Since

$$n_0 = p_s(x_0)(c - a) = \mathcal{N}\exp[-U(x_0)/D](c - a) . \tag{9.2.15}$$

we see that the limit $D \to 0$ corresponds to $n_0 \to 0$, and hence the rate constant in (9.2.14) multiplying $N_0(t)$ becomes exponentially large. Hence, adiabatic elimination is valid.

This three-state interpretation is essentially the transition state theory of chemical reactions proposed by *Eyring* [9.4].

b) Elimination of Intermediate States

Eliminating $N_0(t)$ from (9.2.8) by adding the two equations, we get

$$\dot{N}_a(t) = -\dot{N}_c(t) = r_a N_a(t) + r_c N_c(t) \tag{9.2.16}$$

with

$$r_a = D[n_a \int_a^c dx\, p_s(x)^{-1}]^{-1} \qquad r_c = D[n_c \int_a^c dx\, p_s(x)^{-1}]^{-1} , \tag{9.2.17}$$

where r_a and r_c are independent of x_0. Thus, the precise choice of x_0 does not affect the interpeak relaxation.

Since $N_a + N_c = 1$, the relaxation time constant, τ_r, is givenby

$$\tau_r^{-1} = r_a + r_c = \frac{D}{n_a n_c \int_a^c dx\, p_s(x)^{-1}} . \tag{9.2.18}$$

c) The Escape Probability Per Unit Time for a particle initially near a to reach x_0 is the decay rate of $N_a(t)$ under the condition that an absorbing barrier is at x_0. This means that in (9.2.8) we set $N_0(t) = 0$ [but note that $p_s(x)$ is defined by (9.1.2)]. Similar reasoning gives us

$$\dot{N}_a(t) = -DN_a(t)/n_a\kappa(x_0) \tag{9.2.19}$$

so that the mean first passage time is given by

$$\tau_a = n_a D^{-1} \int_a^{x_0} dx\, p_s(x)^{-1} . \tag{9.2.20}$$

This result is essentially that obtained in (5.2.166) by more rigorous reasoning.

d) Dependence of Relaxation Time on Peak Populations

Equation (9.2.18) looks at first glance like a simple formula relating the relaxation time to n_a and $n_c = 1 - n_a$. One might think that all other factors were independent of n_a and $\tau_r \propto n_a(1 - n_a)$. However, a more careful evaluation shows this is not so. If we use the asymptotic evaluation (9.1.13) we find

$$\tau_r = \frac{n_a n_c}{\mathcal{N} D}\sqrt{\frac{2\pi D}{|U''(b)|}}\exp[U(b)/D] \tag{9.2.21}$$

and similarly, $\mathcal{N}$ can be evaluated asymptotically by taking the contribution from each peak. We obtain

$$\mathcal{N}^{-1} = \sqrt{2\pi D}\{[U''(a)]^{-1/2}\exp[-U(a)/D] + [U''(c)]^{-1/2}\exp[-U(c)/D]\} \tag{9.2.22}$$

and similarly, by definition (9.2.4) of n_a and n_c,

$$n_a/n_c = \sqrt{U''(c)/U''(a)}\exp\left[\frac{U(c) - U(a)}{D}\right]. \tag{9.2.23}$$

After a certain amount of algebra, one can rewrite (9.2.21) as

$$\tau_r = 2\pi\, H(b; a, c)[n_a n_c]^{1/2} \tag{9.2.24}$$

with $H(b; a, c)$ a function which depends on the height of $U(b)$ compared to the average of $U(a)$ and $U(c)$: explicitly,

$$H(b; a, c) = [|U''(b)|^{-1/2}U''(a)^{-1/4}U''(c)^{-1/4}]\exp\left[\frac{2U(b) - U(a) - U(c)}{2D}\right]. \tag{9.2.25}$$

9.2.2 Example: Reversible Denaturation of Chymotrypsinogen

Chymotrypsinogen is a protein which can be transformed into a denatured form by applying elevated pressures of up to several thousand atmospheres, as demonstrated by *Hawley* [9.5]. Presumably the molecule collapses suddenly if sufficient pressure is imposed.

A somewhat unrealistic, but simple, model of this process is given by the equation.

$$dx = \frac{-U'(x)}{\gamma}dt + \sqrt{\frac{2kT}{\gamma}}\,dW(t), \tag{9.2.26}$$

where x is the volume of a molecule, $U(x)$ is the Gibbs Free energy per molecule and kT/γ takes the place of D. Here γ is a friction constant, which would arise by an adiabatic elimination procedure like that used to derive the Smoluchowski equation in Sect. 6.4. The stationary distribution is then

$$p_s(x) = \mathcal{N}\exp[-U(x)/kT] \tag{9.2.27}$$

as required by statistical mechanics.

The explicit effect of the variation of pressure is included by writing

$$U(x) = U_0(x) + x\,\delta p \tag{9.2.28}$$

where δp is the pressure difference from the state in which n_a and n_c are equal. The term $x\delta p$ changes the relative stability of the two minima and is equivalent to the work done against the pressure δp. From (9.2.23) this requires $U_0(x)$ to satisfy

$$\sqrt{U_0''(a)}\,\exp\,[U_0(a)/kT] = \sqrt{U_0''(c)}\,\exp\,[U_0(c)/kT]\,. \tag{9.2.29}$$

The maxima and minima of $U(x)$ are slightly different from those of $U_0(x)$. If we assume that higher derivatives of $U_0(x)$ are negligible, then the maxima and minima of $U(x)$ are at points where $U'(x) = 0$ and are given by $a + \delta a$, $b + \delta b$, $c + \delta c$, where

$$\left.\begin{aligned} \delta a &= -\delta p/U_0''(a) = \beta_a \delta p \\ \delta b &= \delta p/|\,U_0''(b)\,| = \beta_b \delta p \\ \delta c &= -\delta p/U_0''(c) = \beta_c \delta p\,. \end{aligned}\right\} \tag{9.2.30}$$

We thus identify β_a and β_c as the *compressibilities* $\partial x/\partial p$ of the states a and c, which are negative, as required by stability. The quantity β_b is some kind of *incompressibility* of the transition state. Since this is unstable, β_b is positive.

The values of $U(a)$, $U(b)$, $U(c)$ of these minima are

$$\left.\begin{aligned} U(a + \delta a) &= U_0(a + \delta a) + (a + \delta a)\delta p \\ &= U_0(a) + a\delta p + \tfrac{1}{2}\beta_a(\delta p)^2 \\ U(b + \delta b) &= U_0(b) + b\delta p + \tfrac{1}{2}\beta_b(\delta p)^2 \\ U(c + \delta c) &= U_0(c) + c\delta p + \tfrac{1}{2}\beta_c(\delta p)^2 \end{aligned}\right\} \tag{9.2.31}$$

and from (9.2.23) we obtain

$$\boxed{\frac{n_a}{n_c} = \exp\left[\frac{-(a-c)}{kT}\delta p - \frac{(\beta_a - \beta_c)}{2kT}(\delta p)^2\right].} \tag{9.3.32}$$

This formula is exactly that obtained by thermodynamic reasoning and fits the data well.

The relaxation time τ_r to the stationary distribution has also been measured. We compute it using (9.2.24, 25). We find that

$$\boxed{\begin{aligned} \tau_r(0) &= \gamma^{-1}\pi\beta_b{}^{1/2}\exp\left[\frac{U(b)}{kT}\right] \\ \tau_r(\delta p) &= (n_a n_c)^{1/2}\tau_r(0)\exp\left[-\frac{a+c-2b}{2kT}\delta p + \frac{\delta p^2}{2kT}(\beta_b - \tfrac{1}{2}\beta_a - \tfrac{1}{2}\beta_c)\right]. \end{aligned}} \tag{9.2.33}$$

Notice that, in principle, a and c, the volumes of the two states and β_a and β_c, their compressibilities, are all measurable directly. The transition state data b, $U(b)$ and β_b are left as free parameters to be determined from lifetime measurements. Of course, the quadratic terms will only be valid for sufficiently small δp and for applications, it may be necessary to use a more sophisticated method.

In Fig. 9.2, $\tau_r(\delta p)$ and n_a/n_c are plotted for a set of possible values of parameters, as computed from (9.2.33).

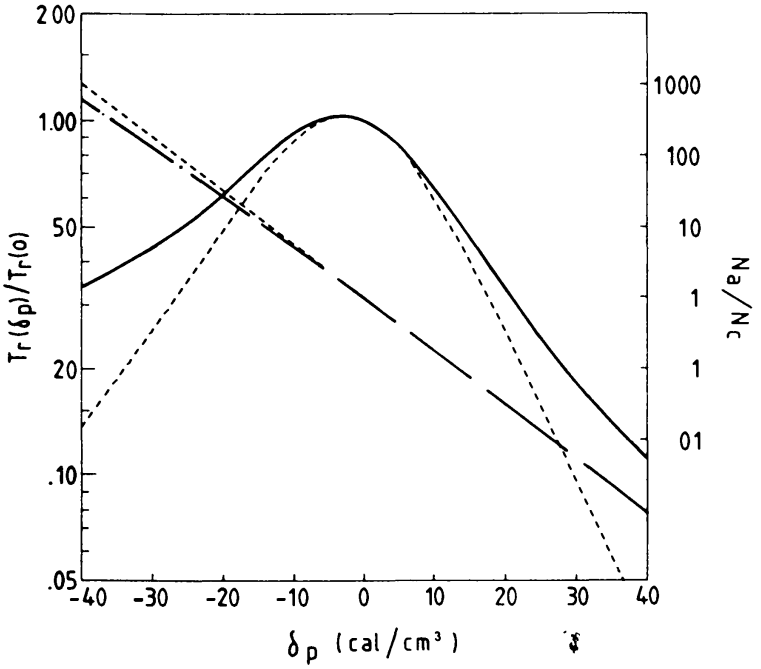


Fig. 9.2. Relaxation rate $\tau_r(\delta p)$ and ratio of concentrations of natural and denatured forms of chymotrypsinogen according to (9.2.33). Dashed line omits compressibility corrections

Notice that the equilibration time reaches a maximum near the point at which natural and denatured forms are in equal concentration. Some skewing, however, can be induced by making the potential asymmetric. Hawley, in fact, notes that measurements in this region are limited by "instrumental stability and the patience of the investigator."

Finally, the curves with zero compressibility are given for comparison. The difference is so large at the wings that it is clear that the quadratic correction method is not valid for the $\tau_r(\delta p)$ curve. However, the corrections almost cancel for the n_a/n_c curve. A realistic treatment in which more variables are included preserves the qualitative nature of this description, but permits as well the possibility $\beta_b < 0$, which is not possible here, as is shown in [9.9].

9.2.3 Bistability with Birth-Death Master Equations (One Variable)

The qualitative behaviour of bistable systems governed by one-step birth-death Master equations is almost the same as that for Fokker-Planck equations.

Consider a one step process with transition probabilities $t^+(x)$, $t^-(x)$ so that the Master equation can be written as

$$\frac{dP(x)}{dt} = J(x+1, t) - J(x, t) \tag{9.2.34}$$

with

$$J(x, t) = t^{-}(x)P(x, t) - t^{+}(x-1)P(x-1, t)\,. \tag{9.2.35}$$

Suppose now that the stationary distribution has maxima at a, c and a minimum at b, and in a similar way to that in Sect. 9.2.1, define

$$M(x, t) = \sum_{z=0}^{x-1} P(z, t) \tag{9.2.36}$$

$$N_a(t) = 1 - N_c(t) = M(b, t) \tag{9.2.37}$$

and if x_0 is a point near b,

$$N_0(t) = P(x_0, t)\,. \tag{9.2.38}$$

The corresponding stationary quantities are

$$n_a = 1 - n_c = \sum_{z=0}^{b-1} P_s(z) \tag{9.2.39}$$

$$n_0 = P_s(x_0)\,. \tag{9.2.40}$$

We now sum (9.2.34) from 0 to $x - 1$ to obtain

$$\frac{dM(x, t)}{dt} = J(x, t) \tag{9.2.41}$$

[since $J(0, t) = 0$]. We now use the fact that the stationary solution $P_s(x)$ is in a one-step process obtained from the detailed balance equation

$$t^{-}(x)P_s(x) = t^{+}(x-1)P_s(x-1) \tag{9.2.42}$$

to introduce an "integrating factor" for (9.2.41). Namely, define

$$\beta(x, t) = P(x, t)/P_s(x) \tag{9.2.43}$$

then (9.2.41) can be written

$$\frac{dM(x, t)}{dt} = P_s(x)t^{-}(x)[\beta(x, t) - \beta(x-1, t)] \tag{9.2.44}$$

so that

$$\begin{aligned} \frac{d}{dt}\sum_{z=a+1}^{x_0}\left[\frac{M(z, t)}{P_s(z)t^{-}(z)}\right] &= \beta(x_0, t) - \beta(a, t) \\ &= \frac{P(x_0, t)}{P_s(x_0)} - \frac{P(a, t)}{P_s(a)}\,. \end{aligned} \tag{9.2.45}$$

Equation (9.2.45) is now in almost precisely the form of (9.2.6) for the corresponding Fokker-Planck process. It depends on the solution being obtained via a detailed balance method (9.2.42).

We make the same assumptions as Kramers; namely, we assume that only relaxation between peaks is now relevant and write

$$\begin{aligned} P(x, t) &= P_s(x)N_a(t)/n_a \qquad x < b \\ &= P_s(x)N_c(t)/n_c \qquad x > b \end{aligned} \tag{9.2.46}$$

and obtain relaxation equations exactly like those in (9.2.8)

$$\begin{aligned} \kappa(x_0)\dot{N}_a(t) &= N_0(t)/n_0 - N_a(t)/n_a \\ \mu(x_0)\dot{N}_c(t) &= N_0(t)/n_0 - N_c(t)/n_c \, . \end{aligned} \tag{9.2.47}$$

where

$$\begin{aligned} \kappa(x_0) &= \sum_{z=a+1}^{x_0} [P_s(z)t^-(z)]^{-1}[1 - \psi(z)] \\ \mu(x_0) &= \sum_{z=x_0+1}^{c} [P_s(z)t^-(z)]^{-1}[1 - \psi(z)] \end{aligned} \tag{9.2.48}$$

with

$$\begin{aligned} \psi(z) &= n_a^{-1} \sum_{y=z}^{b} P_s(y) \qquad z < b \\ &= n_c^{-1} \sum_{y=b+1}^{z+1} P_s(y) \qquad z > b \, . \end{aligned} \tag{9.2.49}$$

The only significant difference is that D appears on the right of (9.2.8) but is here replaced by a factor $t^-(z)^{-1}$ in the definitions of $\kappa(x_0)$ and $\mu(x_0)$.

All the same approximations can be made, the only difficulty being a precise reformulation of the $D \to 0$ limit, which must here correspond to a large number limit, in which all functions change smoothly as x changes by ± 1. This is just the limit of the system size expansion, in which a Fokker-Planck description can be used anyway.

We shall not go into details, but merely mention that exact mean exit times are obtainable by the method of Sect. 7.4. By adapting the methods of Sect. 5.2.8 to this system one finds the splitting probabilities that the system initially at x_0, reaches points a, c are

$$\begin{aligned} \pi_a &= \{ \sum_{z=x_0+1}^{c} [P_s(z)t^-(z)]^{-1} \} / \{ \sum_{z=a+1}^{c} [P_s(z)t^-(z)]^{-1} \} \\ \pi_c &= \{ \sum_{z=a+1}^{x_0} [P_s(z)t^-(z)]^{-1} \} / \{ \sum_{z=a+1}^{c} [P_s(z)t^-(z)]^{-1} \} \, . \end{aligned} \tag{9.2.50}$$

Thus, for all practical considerations one might just as well model by means of a

Fokker-Planck description. It is rare that one knows exactly what the underlying mechanisms are, so that any equation written down can be no more than an educated guess, for which purpose the simplest is the most appropriate.

9.3 Bistability in Multivariable Systems

There is a wide variety of possibilities when one deals with multivariable systems. If the system is described by a Master equation, the possible variety of kinds of transition and state space is so bewilderingly rich, that one can hardly imagine where to start. However, since we saw in Sect. 9.2.4 that a Master equation description is not very different from a Fokker-Planck description, it seems reasonable to restrict oneself to these, which, it turns out, are already quite sufficiently complicated.

The heuristic treatments of these problems, as developed mainly by the physicists *Langer, Landauer* and *Swanson* [9.6] are now in the process of being made rigorous by mathematical treatments by *Schuss* and *Matkowsky* [9.7] and others. The first rigorous treatment was by *Ventsel* and *Freidlin* [9.8] which, however, does not seem to have attracted much attention by applied workers since the rigour is used only to confirm estimates that have long been guessed, rather than to give precise asymptotic expansions, as do the more recent treatments.

We will consider here systems described in a space of l dimensions by a Fokker-Planck equation which is conveniently written in the form

$$\partial_t p = \nabla \cdot [-\boldsymbol{v}(\boldsymbol{x})p + \varepsilon \underset{\sim}{\boldsymbol{D}}(\boldsymbol{x}) \cdot \nabla p] \tag{9.3.1}$$

whose stationary solution is called $p_s(\boldsymbol{x})$ and which is assumed to be known in much of what follows. It can, of course, be estimated asymptotically in the small ε limit by the method of Sect. 6.3.3.

9.3.1 Distribution of Exit Points

We will treat here only a simplified case of (9.3.1) in which $\underset{\sim}{\boldsymbol{D}}(\boldsymbol{x})$ is the identity:

$$\underset{\sim}{\boldsymbol{D}}(\boldsymbol{x}) = \mathbf{1} . \tag{9.3.2}$$

This does conceal features which can arise from strongly varying D but mostly, the results are not greatly changed.

We suppose that the system is confined to a region R with boundary $\boldsymbol{S}$, and that the velocity field $\boldsymbol{v}(\boldsymbol{x})$ points inwards to a stationary point $\boldsymbol{a}$. The problem is the asymptotic estimate of the distribution of points $\boldsymbol{b}$ on $\boldsymbol{S}$ at which the point escapes from R. We use (5.4.49) for $\pi(\boldsymbol{b}, \boldsymbol{x})$ (the distribution of escape points on $\boldsymbol{S}$, starting from the point $\boldsymbol{x}$), which in this case takes the form

$$\boldsymbol{v}(\boldsymbol{x}) \cdot \nabla \pi(\boldsymbol{b}, \boldsymbol{x}) + \varepsilon \nabla^2 \pi(\boldsymbol{b}, \boldsymbol{x}) = 0 \tag{9.3.3}$$

with boundary condition

$$\pi(\boldsymbol{b}, \boldsymbol{u}) = \delta_s(\boldsymbol{b} - \boldsymbol{u}) \qquad (\boldsymbol{u} \in \boldsymbol{S}) . \tag{9.3.4}$$

An asymptotic solution, valid for $\varepsilon \to 0$, is constructed, following the method of *Matkowsky and Schuss* [9.7].

a) Solution Near $x = u$ and in the Interior of R

Firstly one constructs a solution valid inside R. For $\varepsilon = 0$ we have

$$\boldsymbol{v}(\boldsymbol{x}) \cdot \nabla \pi(\boldsymbol{b}, \boldsymbol{x}) = 0 \tag{9.3.5}$$

which implies that $\pi(\boldsymbol{b}, \boldsymbol{x})$ is constant along the flow lines of $\boldsymbol{v}(\boldsymbol{x})$, since it simply states that the derivative of $\pi(\boldsymbol{b}, \boldsymbol{x})$ along these lines is zero. Since we assume all the flow lines pass through $\boldsymbol{a}$, we have

$$\pi(\boldsymbol{b}, \boldsymbol{x}) = \pi(\boldsymbol{b}, \boldsymbol{a}) \tag{9.3.6}$$

for any $\boldsymbol{x}$ inside. However, the argument is flawed by the fact that $\boldsymbol{v}(\boldsymbol{a}) = 0$ and hence (9.3.5) is no longer an appropriate approximation.

We consider, therefore, the solution of (9.3.3) within a distance of order $\sqrt{\varepsilon}$ of the origin. To assist in this, we introduce new coordinates (z, y_r) which are chosen so that z measures the distance from $\boldsymbol{a}$, while the y_r are a set of $l - 1$ tangential variables measuring the orientation around $\boldsymbol{a}$.

More precisely, choose $z(\boldsymbol{x})$ and $y_r(\boldsymbol{x})$ so that

$$\begin{aligned} &\boldsymbol{v}(\boldsymbol{x}) \cdot \nabla z(\boldsymbol{x}) = -z(\boldsymbol{x}) \\ &\boldsymbol{v}(\boldsymbol{x}) \cdot \nabla y_r(\boldsymbol{x}) = 0 \\ &z(\boldsymbol{a}) = 0 . \end{aligned} \tag{9.3.7}$$

The negative sign in the first of these equations takes account of the fact that $\boldsymbol{a}$ is assumed stable, so that $\boldsymbol{v}(\boldsymbol{x})$ points towards $\boldsymbol{a}$. Thus, $z(\boldsymbol{x})$ *increases* as $\boldsymbol{x}$ travels further from $\boldsymbol{a}$.

Thus, we find, for any function f,

$$\nabla f = \nabla z(\boldsymbol{x}) \frac{\partial f}{\partial z} + \sum_r \nabla y_r(\boldsymbol{x}) \frac{\partial f}{\partial y_r} \tag{9.3.8}$$

and hence,

$$\boldsymbol{v}(\boldsymbol{x}) \cdot \nabla \pi = -z \frac{\partial \pi}{\partial z} \tag{9.3.9}$$

and

$$\begin{aligned} \nabla^2 \pi = {} & \nabla z(\boldsymbol{x}) \cdot \nabla z(\boldsymbol{x}) \frac{\partial^2 \pi}{\partial z^2} + 2 \sum_r \nabla z(\boldsymbol{x}) \cdot \nabla y_r(\boldsymbol{x}) \frac{\partial^2 \pi}{\partial z \, \partial y_r} \\ & + \sum_{r,s} \nabla y_r(\boldsymbol{x}) \cdot \nabla y_s(\boldsymbol{x}) \frac{\partial^2 \pi}{\partial y_r \partial y_s} + \nabla^2 z(\boldsymbol{x}) \frac{\partial \pi}{\partial z} + \sum_r \nabla^2 y_r(\boldsymbol{x}) \frac{\partial \pi}{\partial y_r} . \end{aligned} \tag{9.3.10}$$

We now evaluate π asymptotically by changing to the scaled (or stretched) variable ξ defined by

$$z = \xi\sqrt{\varepsilon} \,. \tag{9.3.11}$$

Substituting (9.3.8–11) into (9.3.3) we find that, to lowest order in ε,

$$-\xi \frac{\partial \pi}{\partial \xi} + H \frac{\partial^2 \pi}{\partial \xi^2} = 0$$

where (9.3.12)

$$H = \nabla z(\boldsymbol{a}) \cdot \nabla z(\boldsymbol{a}) \,.$$

We can now solve this equation getting

$$\pi(\boldsymbol{b}, \boldsymbol{x}) = C_1 \int_0^{z/\sqrt{\varepsilon}} d\xi \exp(\xi^2/2H) + \pi(\boldsymbol{b}, \boldsymbol{a}) \,. \tag{9.3.13}$$

Because H is positive, we can only match this solution for π with the constancy of $\pi(\boldsymbol{b}, \boldsymbol{x})$ along flow lines for $\boldsymbol{x} \neq \boldsymbol{a}$ if $C_1 = 0$. Hence, for all $\boldsymbol{x}$ on the interior of R,

$$\pi(\boldsymbol{b}, \boldsymbol{x}) = \pi(\boldsymbol{b}, \boldsymbol{a}) \,. \tag{9.3.14}$$

[Notice that if $\boldsymbol{v}(\boldsymbol{x})$ points the other way, i.e., *is unstable*, we omit the negative sign in (9.3.9) and find that $\pi(\boldsymbol{b}, \boldsymbol{x})$ is given by (9.3.3) with $H \to -H$, and hence in a distance of order $\sqrt{\varepsilon}$ of $\boldsymbol{a}$, $\pi(\boldsymbol{b}, \boldsymbol{x})$ changes its value].

b) Solution Near the Boundary S

We consider, with an eye to later applications, the solution of the slightly more general equation

$$\boldsymbol{v}(\boldsymbol{x}) \cdot \nabla f(\boldsymbol{x}) + \varepsilon \nabla^2 f(\boldsymbol{x}) = 0$$

with (9.3.15)

$$f(\boldsymbol{u}) = g(\boldsymbol{u}) \qquad (\boldsymbol{u} \in \boldsymbol{S})$$

of which the boundary value problem (9.3.4) is a particular case. Two situations arise.

i) $\boldsymbol{v} \cdot \boldsymbol{v}(\boldsymbol{x}) \neq 0$ *on* $\boldsymbol{S}$ *or anywhere in* R *except* $\boldsymbol{x} = \boldsymbol{a}$ *which is stable:* clearly, in any asymptotic method, the boundary condition (9.3.15) is not compatible with a constant solution. Hence, there must be rapid changes at the boundary.

Near a point $\boldsymbol{u}$ on $\boldsymbol{S}$ we can write

$$\boldsymbol{v}(\boldsymbol{u}) \cdot \nabla f(\boldsymbol{x}) + \varepsilon \nabla^2 f(\boldsymbol{x}) = 0 \,. \tag{9.3.16}$$

Near $\boldsymbol{S}$ it is most convenient to introduce $\boldsymbol{\nu}(\boldsymbol{u})$, the normal (pointing out) at $\boldsymbol{u}$ to $\boldsymbol{S}$, and to define a variable ρ by

$$\boldsymbol{x} = \boldsymbol{u} - \varepsilon\rho\boldsymbol{\nu}(\boldsymbol{u}) \tag{9.3.17}$$

and other variables y_r parallel to $\boldsymbol{S}$.

Then to lowest order in ε, (9.3.16) reduces to (at a point near $\boldsymbol{u}$ on $\boldsymbol{S}$)

$$[\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{u})]\frac{\partial f}{\partial \rho} + H(\boldsymbol{u})\frac{\partial^2 f}{\partial \rho^2} = 0 \tag{9.3.18}$$

with $H(\boldsymbol{u}) = \boldsymbol{\nu}^2 = 1$.

The solution is then

$$f(\boldsymbol{x}) = g(\boldsymbol{u}) + C_1(\boldsymbol{u})\{1 - \exp[-\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{u})\rho]\} \ . \tag{9.3.19}$$

As $\rho \to \infty$, we approach a finite distance into the interior of R and thus

$$f(\boldsymbol{x}) \to g(\boldsymbol{u}) + C_1(\boldsymbol{u}) = C_0 \tag{9.3.20}$$

from the analysis in (a), so

$$C_1(\boldsymbol{u}) = C_0 - g(\boldsymbol{u}) \ . \tag{9.3.22}$$

One must now fix C_0, which is the principal quantity actually sought.

This can be done by means of Green's theorem. For let $p_s(\boldsymbol{x})$ be the usual stationary solution of the forward Fokker-Planck equation. We know it can be written

$$p_s(\boldsymbol{x}) = \exp\left\{-\frac{1}{\varepsilon}[\phi(\boldsymbol{x}) + O(\varepsilon)\right\} \tag{9.3.23}$$

as has been shown in Sect. 6.3.3. We take (9.3.16), multiply by $p_s(\boldsymbol{x})$ and integrate over R. Using the fact that $p_s(\boldsymbol{x})$ satisfies the forward Fokker-Planck equation, this can be reduced to a surface integral

$$0 = \int_R d\boldsymbol{x}\, p_s(\boldsymbol{x})[v(\boldsymbol{x})\cdot\nabla f(\boldsymbol{x}) + \varepsilon\nabla^2 f(\boldsymbol{x})] \tag{9.3.24}$$

$$= \int_S dS\ \{p_s(\boldsymbol{x})\boldsymbol{\nu}\cdot v(x) f(\boldsymbol{x}) + \varepsilon[p_s(\boldsymbol{x})\boldsymbol{\nu}\cdot\nabla f(\boldsymbol{x}) - f(\boldsymbol{x})\boldsymbol{\nu}\cdot\nabla p_s(\boldsymbol{x})]\} \ . \tag{9.3.25}$$

Noting that, to lowest order in ε.

$$\boldsymbol{\nu}\cdot\nabla f(\boldsymbol{x}) = -\frac{1}{\varepsilon}\frac{\partial f}{\partial \rho} = -\boldsymbol{\nu}\cdot v(x)[C_0 - g(\boldsymbol{x})] \tag{9.3.26}$$

and

$$\boldsymbol{\nu}\cdot\nabla p_s(\boldsymbol{x}) = -\frac{1}{\varepsilon}\boldsymbol{\nu}\cdot\nabla\phi(\boldsymbol{x})\exp[-\phi(\boldsymbol{x})/\varepsilon] \ , \tag{9.3.27}$$

we deduce that

$$C_0 = \frac{\int_S dS\, e^{-\phi(x)/\varepsilon}\,[2\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x}) + \boldsymbol{\nu}\cdot\nabla\phi(\boldsymbol{x})]g(\boldsymbol{x})}{\int_S dS\, e^{-\phi(x)/\varepsilon}\,\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x})}\,. \tag{9.3.28}$$

Recalling that for this problem, $g(\boldsymbol{u}) = \delta_s(\boldsymbol{u} - \boldsymbol{b})$, we find that, if $\boldsymbol{x}$ is well in the interior of R,

$$\pi(\boldsymbol{x}, \boldsymbol{b}) = C_0 = \frac{e^{-\phi(b)/\varepsilon}\,[2\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{b}) + \boldsymbol{\nu}\cdot\nabla\phi(\boldsymbol{b})]}{\int_S dS\, e^{-\phi(x)/\varepsilon}\,\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x})}\,. \tag{9.3.29}$$

We see here that the exit distribution is essentially $\exp[-\phi(\boldsymbol{b})/\varepsilon]$, i.e., approximately the stationary distribution. If the Fokker-Planck equation has a potential solution, then

$$\boldsymbol{v}(\boldsymbol{b}) = -\nabla\phi(\boldsymbol{b}) \tag{9.3.30}$$

and

$$\pi(\boldsymbol{x}, \boldsymbol{b}) = e^{-\phi(b)/\varepsilon}\,\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{b})/[\int_S dS\, e^{-\phi(x)/\varepsilon}\,\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x})]$$

and we simply have a kind of average flow result.

ii) $\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x}) = 0$ *on* $\boldsymbol{S}$: this problem is more directly related to bistability, since midway between two stable points $\boldsymbol{a}$ and $\boldsymbol{c}$, a curve $\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x}) = 0$ which separates the two regions and is known as *a separatrix* is expected.

The method is much the same except that near $\boldsymbol{u}$ on $\boldsymbol{S}$, we expect

$$\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x}) \sim \boldsymbol{\nu}\cdot(\boldsymbol{x} - \boldsymbol{u})\kappa(\boldsymbol{u})\,. \tag{9.3.31}$$

where $\kappa(\boldsymbol{u})$ is a coefficient which depends on $\boldsymbol{v}(\boldsymbol{x})$ and is assumed to be nonzero.

The situation is now like that at $\boldsymbol{x} = \boldsymbol{a}$ and it is appropriate to substitute

$$\boldsymbol{x} = \boldsymbol{u} - \sqrt{\varepsilon}\,\rho\boldsymbol{\nu}(\boldsymbol{u}) \tag{9.3.32}$$

and to lowest order in ε (9.3.16) reduces to (at a point near $\boldsymbol{u}$ on $\boldsymbol{S}$)

$$\kappa(\boldsymbol{u})\,\rho\frac{\partial f}{\partial\rho} + \frac{\partial^2 f}{\partial\rho^2} = 0 \tag{9.3.33}$$

so that

$$f(\boldsymbol{x}) = g(\boldsymbol{u}) + C_1\int_0^\rho d\rho\, \exp[-\tfrac{1}{2}\kappa(\boldsymbol{u})\rho^2] \tag{9.3.34}$$

and letting $\rho \to \infty$, we find

$$C_1 = [C_0 - g(\boldsymbol{u})]\sqrt{\frac{2\kappa(\boldsymbol{u})}{\pi}}. \tag{9.3.35}$$

The result for $\pi(\boldsymbol{x}, \boldsymbol{b})$ is now completed as before: one gets

$$\boxed{\pi(\boldsymbol{x}, \boldsymbol{b}) = \frac{e^{-\phi(\boldsymbol{b})/\varepsilon}\left[\left(1 + \sqrt{\frac{2\kappa(\boldsymbol{b})}{\pi}}\right)\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{b}) + \boldsymbol{\nu}\cdot\nabla\phi(\boldsymbol{b})\right]}{\int dS\, e^{-\phi(\boldsymbol{x})/\varepsilon}\sqrt{\frac{2\kappa(\boldsymbol{x})}{\pi}}\,\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x})}.} \tag{9.3.36}$$

9.3.2 Asymptotic Analysis of Mean Exit Time

From our experience in one-dimensional systems, we expect the mean exit time from a point within R to be of order $\exp(K/\varepsilon)$; for some $K > 0$, as $\varepsilon \to 0$. We therefore define

$$\tau(\boldsymbol{x}) = \exp(-K/\varepsilon)T(\boldsymbol{x}) \tag{9.3.37}$$

where $T(\boldsymbol{x})$ is the mean escape time from R starting at $\boldsymbol{x}$ and $\tau(\boldsymbol{x})$ satisfies (from Sect. 5.4)

$$\begin{aligned} &\boldsymbol{v}(\boldsymbol{x})\cdot\nabla\tau(\boldsymbol{x}) + \varepsilon\nabla^2\tau(\boldsymbol{x}) = -e^{-K/\varepsilon} \\ &\tau(\boldsymbol{u}) = 0 \qquad \boldsymbol{u} \in \boldsymbol{S}. \end{aligned} \tag{9.3.38}$$

If this scaling is correct, then any expansion of $\tau(\boldsymbol{x})$ in powers of ε will not see the exponential, so the equation to lowest order in ε will be essentially (9.3.16).

As in that case, we show that $\tau(\boldsymbol{x})$ is essentially constant in the interior of R and can be written as [in the case $\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x}) \neq 0$ on $\boldsymbol{S}$]

$$\tau(\boldsymbol{x}) \sim C_0\{1 - \exp[-\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{u})\rho]\} \tag{9.3.39}$$

near $\boldsymbol{S}$.

We multiply (9.3.38) by $p_s(\boldsymbol{x}) = \exp[-\phi(\boldsymbol{x})/\varepsilon]$ and use Green's theorem to obtain [in much the same way as (9.3.25) but with $\tau(\boldsymbol{x}) = 0$ on $\boldsymbol{S}$]

$$-e^{-K/\varepsilon}\int_R d\boldsymbol{x}\, e^{-\phi(\boldsymbol{x})/\varepsilon} = -\int_S dS\, e^{-\phi(\boldsymbol{x})/\varepsilon}\,[C_0\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x})], \tag{9.3.40}$$

i.e.,

$$C_0 = \int_R d\boldsymbol{x}\, e^{-[K+\phi(\boldsymbol{x})]/\varepsilon}\Big/\int_S dS\, e^{-\phi(\boldsymbol{x})/\varepsilon}\,\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x}). \tag{9.3.41}$$

By hypothesis, C_0 does not change exponentially like $\exp(A/\varepsilon)$. In the numerator of (9.3.41) the main contribution comes from the minimum of $\phi(x)$ which occurs at the point $\boldsymbol{a}$, whereas in the denominator, it occurs at the point on $\boldsymbol{S}$ where $\phi(\boldsymbol{x})$ is a minimum, which we shall call $\boldsymbol{x}_0$. Thus, the ratio behaves like

$$\exp\{[\phi(\boldsymbol{a}) - \phi(\boldsymbol{x}_0) - K]/\varepsilon\} .$$

and hence for C_0 to be asymptotically constant,

$$K = \phi(\boldsymbol{a}) - \phi(\boldsymbol{x}_0) \tag{9.3.42}$$

and, for $\boldsymbol{x}$ well into the interior of R, we have

$$\boxed{\tau(\boldsymbol{x}) = \int_R d\boldsymbol{x}\, e^{\phi(\boldsymbol{x})/\varepsilon} / \int_S dS\, [e^{-\phi(\boldsymbol{x})/\varepsilon}\, \boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x})] .} \tag{9.3.43}$$

In the case where $\boldsymbol{\nu}\cdot\boldsymbol{v}(\boldsymbol{x}) = 0$ on all of S, we now have

$$\tau(\boldsymbol{x}) \sim C_0 \int_0^{\rho} d\rho \exp[-\tfrac{1}{2}\kappa(\boldsymbol{u})\rho^2] \tag{9.3.44}$$

and hence in the interior,

$$\tau(\boldsymbol{x}) \sim C_0 \sqrt{\frac{\pi}{2\kappa(\boldsymbol{u})}} . \tag{9.3.45}$$

The analysis proceeds similarly and we find, for $\boldsymbol{x}$ well in the interior of R,

$$\boxed{\tau(\boldsymbol{x}) \sim \sqrt{\frac{\pi}{2\kappa(\boldsymbol{u})}} \int_R d\boldsymbol{x}\, e^{\phi(\boldsymbol{x})/\varepsilon} / \int_S dS\, e^{-\phi(\boldsymbol{x})/\varepsilon} .} \tag{9.3.46}$$

9.3.3 Kramers' Method in Several Dimensions

The generalisation of Kramers' method is relatively straightforward. We consider a completely general Fokker-Planck equation in l dimensions [we use $P(\boldsymbol{x})$ for the probability density for notational ease]

$$\partial_t P = \nabla\cdot[-\boldsymbol{v}(\boldsymbol{x})P + \varepsilon \underset{\sim}{\boldsymbol{D}}(\boldsymbol{x})\cdot\nabla P] \tag{9.3.47}$$

whose stationary solution is to be called $P_s(\boldsymbol{x})$ and can only be exhibited explicitly if (9.3.47) satisfies potential conditions. We assume that $P_s(\boldsymbol{x})$ has two well-defined maxima at $\boldsymbol{a}$ and $\boldsymbol{c}$ and well-defined saddlepoint at $\boldsymbol{b}$ (Fig. 9.3). We assume that the value at the saddlepoint is very much smaller than the values at $\boldsymbol{a}$ and $\boldsymbol{c}$. We introduce a family of $(l-1)$ dimensional planes $\boldsymbol{S}(w)$, where w is a parameter which labels the planes. We choose $\boldsymbol{S}(a)$ to pass through $\boldsymbol{a}$, $\boldsymbol{S}(b)$ through $\boldsymbol{b}$ and $\boldsymbol{S}(c)$ through $\boldsymbol{c}$. The planes $\boldsymbol{S}(w)$ are assumed to be oriented in such a way that $P_s(\boldsymbol{x})$ has a unique maximum when restricted to any one of them. We define, similarly to Sect. 9.2.1

$$M[\boldsymbol{S}(w)] = \int_{L(w)} dx\, P(\boldsymbol{x}) , \tag{9.3.48}$$

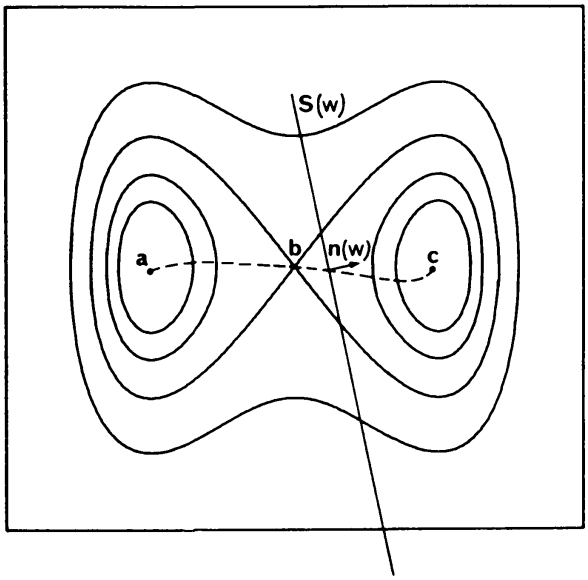


Fig. 9.3. Contours of the stationary distribution function $P_s(x)$. The plane $S(w)$ is oriented so that $P_s(x)$ has a unique maximum there, and the curve $x = u(w)$ (dashed line) is the locus of these maxima

where $L(w)$ is the region of space to the left of the plane $S(w)$; then

$$\dot{M}[S(w)] = \int_{S(w)} dS \cdot [-v(x)P + \varepsilon \underset{\sim}{D}(x) \cdot \nabla P] \,. \tag{9.3.49}$$

Current in stationary state is defined by

$$J_s = -v(x)P_s + \varepsilon \underset{\sim}{D}(x) \cdot \nabla P_s \,. \tag{9.3.50}$$

Assumption I: we exclude cases in which finite currents J_s occur where P_s is very small. Because of $\nabla \cdot J_s = 0$, we can write

$$J_s = -\varepsilon \nabla \cdot (\underset{\sim}{A} P_s) \tag{9.3.51}$$

where $\underset{\sim}{A}$ is an antisymmetric tensor. *We require that* $\underset{\sim}{A}$ *be of the some order of magnitude as* $\underset{\sim}{D}(x)$, *or smaller.*

Relaxation equations are derived in two stages. Define a quantity $\beta(x)$ by

$$\begin{aligned} \beta(x) = P(x, t)/P_s(x) &\simeq N_a(t)/n_a \qquad (x \text{ near } a) \\ &\simeq N_c(t)/n_c \qquad (x \text{ near } c) \,. \end{aligned} \tag{9.3.52}$$

This is the assumption that all relaxation within peaks has ceased. Substitute now in (9.3.49), integrate by parts discarding terms at infinity and obtain

$$\dot{M}[S(w)] = \varepsilon \int_{S(w)} dS \cdot [\mathscr{D}(x) \cdot \nabla \beta] \, P_s(x) \tag{9.3.53}$$

with

$$\mathscr{D}(x) = \underset{\sim}{D}(x) + \underset{\sim}{A}(x) \,. \tag{9.3.54}$$

Assumption II: $P_s(x)$ is sharply singly peaked on $S(w)$ so we may make the approximate evaluation

$$\dot{M}[\boldsymbol{S}(w)] = \{\varepsilon[\boldsymbol{n}(w)\cdot\mathscr{D}(\boldsymbol{x})\cdot\nabla\beta]_{\boldsymbol{u}(w)} + \delta(w)\} \left|\int_{\boldsymbol{S}(w)} d\boldsymbol{S}\, P_s(\boldsymbol{x})\right| , \tag{9.3.55}$$

where $\delta(w)$ is expected to be very much smaller than the term in square brackets. Here $\boldsymbol{u}(w)$ is the position at which $P_s(x)$ has its maximum value when restricted to $\boldsymbol{S}(w)$, and $\boldsymbol{n}(w)$ is the normal to $\boldsymbol{S}(w)$.

Assumption III: the direction of $\boldsymbol{n}(w)$ can be chosen so that $\mathscr{D}^{\mathrm{T}}(\boldsymbol{x})\cdot\boldsymbol{n}(w)$ is parallel to the tangent at w to the curve $\boldsymbol{x} = \boldsymbol{u}(w)$ — without violating the other assumptions. Hence,

$$\mathscr{D}^{\mathrm{T}}[\boldsymbol{u}(w)]\cdot\boldsymbol{n}(w) = d(w)\partial_w\boldsymbol{u}(w) . \tag{9.3.56}$$

Defining now

$$p(w) = \left|\int_{\boldsymbol{S}(w)} d\boldsymbol{S}\, P_s(\boldsymbol{x})\right| , \tag{9.3.57}$$

which is (up to a slowly varying factor) the probability density for the particle to be on the plane $\boldsymbol{S}(w)$ and is expected to have a two-peaked shape with maxima at $w = a$ and $w = c$ and a minimum at $w = b$.

Assumption IV: these are assumed to be sharp maxima and minima. Neglecting $\delta(w)$, making the choice (9.3.56) and noting

$$\partial_w\boldsymbol{u}(w)\cdot\nabla\beta[\boldsymbol{u}(w)] = \partial_w\beta[\boldsymbol{u}(w)] , \tag{9.3.58}$$

we find

$$\frac{1}{\varepsilon}\int_a^{w_0} dw\, \{\dot{M}[\boldsymbol{S}(w)]/[p(w)d(w)]\} = \beta(w_0) - \beta(a) . \tag{9.3.59}$$

Using the sharp peaked nature of $p(w)^{-1}$, (9.3.59) can now be approximated by taking the value at the peak, using (9.3.52) and

$$N(a, t) = M[\boldsymbol{S}(b), t] \tag{9.3.60}$$

as well as defining

$$\kappa(w_0) = \int_a^{w_0} [p(w)]^{-1} dw \tag{9.3.61}$$

$$\mu(w_0) = \int_{w_0}^{b} [p(w)]^{-1} dw , \tag{9.3.62}$$

to obtain the relaxation equations

$$\kappa(w_0)\dot{N}_a(t) = \varepsilon d(w_0)[N_0(t)/n_0 - N_a(t)/n_a] \tag{9.3.63}$$

$$\mu(w_0)\dot{N}_c(t) = \varepsilon d(w_0)[N_0(t)/n_0 - N_c(t)/n_c] . \tag{9.3.64}$$

These are of exactly the same form as those in the one-variable case and all the same interpretations can be made.

9.3.4 Example:Brownian Motion in a Double Potential

We consider Brownian motion in velocity and position as outlined in Sect. 5.3.6. Thus, we consider the Fokker-Planck equation

$$\frac{\partial P(x,p,t)}{\partial t} = -p\frac{\partial P}{\partial x} + U'(x)\frac{\partial P}{\partial p} + \gamma\left[\frac{\partial P}{\partial p}pP + \frac{\partial^2 P}{\partial p^2}\right]. \tag{9.3.65}$$

In the notation of the previous section we have

$$\left.\begin{aligned} &\boldsymbol{x} = (x, p)\\ &\boldsymbol{v}(\boldsymbol{x}) = (p, -U'(x) - \gamma p)\\ &\varepsilon = 1\\ &\underset{\sim}{\boldsymbol{D}}(\boldsymbol{x}) = \begin{bmatrix} 0 & 0\\ 0 & \gamma \end{bmatrix}\\ &P_s(\boldsymbol{x}) = \mathcal{N}_2 \exp\left[-\tfrac{1}{2}p^2 - U(x)\right]\\ &\mathcal{N}_2 = (2\pi)^{-1/2}\mathcal{N}_1\\ &\mathcal{N}_1 = \{\int_{-\infty}^{\infty} dx \exp\left[-U(x)\right]\}^{-1}. \end{aligned}\right\} \tag{9.3.66}$$

Hence, we can write

$$\boldsymbol{v}(\boldsymbol{x}) = \begin{bmatrix} 0 & -1\\ 1 & \gamma \end{bmatrix} \cdot \nabla(\log P_s) \tag{9.3.67}$$

and the current in the stationary state is

$$\boldsymbol{J}_s = -\boldsymbol{v}P_s + \underset{\sim}{\boldsymbol{D}} \cdot \nabla P_s = -\nabla \cdot \left\{\begin{bmatrix} 0 & -1\\ 1 & 0 \end{bmatrix} P_s\right\} \tag{9.3.68}$$

so that $\underset{\sim}{\boldsymbol{A}}$ exists, and

$$\underset{\sim}{\boldsymbol{A}} = \begin{bmatrix} 0 & -1\\ 1 & 0 \end{bmatrix}. \tag{9.3.69}$$

Thus, Assumption I is satisfied.

The plane $\boldsymbol{S}(w)$ can be written in the form

$$\lambda x + p = w. \tag{9.3.70}$$

Assumption II requires us to maximise $P_s(\boldsymbol{x})$ on this plane, i.e., to maximise $-\frac{1}{2}p^2 - U(x)$ on this plane. Using standard methods, we find that maxima must lie along the curve $\boldsymbol{u}(w)$ given by

$$\boldsymbol{u}(w) = \begin{bmatrix} x(w) \\ p(w) \end{bmatrix} = \begin{bmatrix} x(w) \\ w - \lambda x(w) \end{bmatrix} \tag{9.3.71}$$

where $x(w)$ satisfies

$$U'[x(w)] + \lambda^2 x(w) - \lambda w = 0 \; ; \tag{9.3.72}$$

whether $P_s(X)$ is sharply peaked depends on the nature of $U(x)$.

We now implement Assumption III.

The parameter λ is a function of w on the particular set of planes which satisfy (9.3.56). The tangent to $\boldsymbol{u}(w)$ is parallel to

$$\left[\frac{dx}{dw}, 1 - \lambda \frac{dx}{dw} - x \frac{d\lambda}{dw}\right] \tag{9.3.73}$$

and differentiating (9.3.72) we have

$$\frac{dx}{dw} = (U'' + \lambda^2)^{-1} \left[\lambda - \frac{d\lambda}{dw}(2\lambda x - w)\right] . \tag{9.3.74}$$

The normal to (9.3.70) is parallel to $(\lambda, 1)$. Hence,

$$\mathscr{D}^{\mathrm{T}} \boldsymbol{n} = (1 + \lambda^2)^{-1/2} \begin{bmatrix} 0 & 1 \\ -1 & \gamma \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = (1 + \lambda^2)^{-1/2} \begin{bmatrix} 1 \\ \gamma - \lambda \end{bmatrix} \tag{9.3.75}$$

and this is parallel to (9.3.73) if

$$\frac{dx}{dw} / 1 = \left[1 - \lambda \frac{dx}{dw} - x \frac{d\lambda}{dw}\right] / (\gamma - \lambda) . \tag{9.3.76}$$

We can now solve (9.3.74, 76) simultaneously, to get

$$\frac{dx}{dw} = \frac{1}{\gamma} - \frac{x}{\gamma^2} \left[\frac{U'' - \gamma\lambda + \lambda^2}{x(U'' + \lambda^2) - (2\lambda x - w)}\right] \tag{9.3.77}$$

$$\frac{d\lambda}{dw} = \frac{1}{\gamma} \left[\frac{U'' - \lambda\gamma + \lambda^2}{x(U'' + \lambda^2) - (2\lambda x - w)}\right] , \tag{9.3.78}$$

The saddle point is at $(x, p) = (0, 0)$ and thus $w = 0 \Longleftrightarrow x = 0$. Using this in (9.3.77) we see that we must have

$$x \simeq w/\gamma \qquad \text{as } w \simeq 0 . \tag{9.3.79}$$

Near $x = 0$, we write approximately

$$U[x] \simeq -\tfrac{1}{2} U_2 x^2 \tag{9.3.80}$$

and substituting (9.3.79, 80) in (9.3.72), we see that

$$\lambda^2 - \gamma\lambda + U''(0) = 0 \tag{9.3.81}$$

which determines

$$\lambda(0) = \frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} + U_2}\,. \tag{9.3.82}$$

We now see that (9.3.78) tells us that $d\lambda/dw = 0$ when $w = 0$. Thus, λ will not change significantly from (9.3.82) around the saddle point, and we shall from now on approximate λ by (9.3.82).

Only one of the roots is acceptable and physically, this should be $\lambda \to \infty$ in the high friction limit which would give Kramers' result and requires the positive sign. The other root corresponds to taking a plane such that we get a minimum of $P_s(\boldsymbol{x})$ on it.

We now integrate (9.3.57) and determine $d(w)$. Notice that $d(w)$ must be defined with $\boldsymbol{n}(w)$ a unit vector. Direct substitution in (9.3.75) and using (9.3.79)

$$(1 + \lambda^2)^{-1/2} = \frac{dx}{dw}(w = 0)d(0) = \frac{1}{\gamma}\, d(0) \tag{9.3.83}$$

so that

$$d(0) = \gamma(1 + \lambda^2)^{-1/2}\,. \tag{9.3.84}$$

Further,

$$\begin{aligned} p(w) &= \int_{S(w)} |dS\, P_s(\boldsymbol{x})| = \int_{S(w)} \sqrt{dx^2 + dp^2}\, P_s(x, p) \\ &= \mathcal{N}_2 \frac{[1 + \lambda^2]^{1/2}}{\lambda} \int dp \exp\left[-\frac{p^2}{2} - U\left(\frac{w - p}{\lambda}\right)\right]. \end{aligned} \tag{9.3.85}$$

An exact evaluation depends on the choice of $U(x)$. Approximately, we use

$$U(x) \simeq U_0 - \tfrac{1}{2}U_2 x^2 \tag{9.3.86}$$

and evaluate the result as a Gaussian: we get

$$p(w) = \frac{(1 + \lambda^2)^{1/2}}{\lambda} \mathcal{N}_2\, \mathrm{e}^{-U_0} \exp\left[\frac{U_2 w^2}{2(\lambda^2 - U_2)}\right] \tag{9.3.87}$$

and thus

$$\kappa(0) = \int_{-\infty}^{0} p(w)^{-1} dw = \tfrac{1}{2}\mathcal{N}_2^{-1} \frac{\lambda\gamma}{(1 + \lambda^2)^{1/2}} \frac{\mathrm{e}^{U_0}}{\sqrt{U_2}} = \mu(0)\,. \tag{9.3.88}$$

Thus, from (9.2.19) adapted to the many dimensional theory, we have for the mean first passage time from one well to the point $x = 0$,

$$\tau_0 = \kappa(0)d(0)^{-1} = \frac{\lambda}{2}\, e^{U_0} \mathscr{N}_1^{-1} \sqrt{\frac{2\pi}{U_2}}, \tag{9.3.89}$$

i.e.,

$$\boxed{\tau_0 = \frac{1}{2}\left(\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + U_2}\right) e^{U_0} \mathscr{N}_1^{-1} \sqrt{\frac{2\pi}{U_2}}.} \tag{9.3.90}$$

Comparisons with Other Results

a) Exact One-Dimensional Mean First Passage Time (Smoluchowski's Equation)
One reduces Kramers equation in the large friction limit to the Smoluchowski equation for

$$\hat{P}(x, t) = \int dv\, P(x, v, t), \tag{9.3.91}$$

i.e.,

$$\frac{\partial \hat{P}(x, t)}{\partial t} = \frac{1}{\gamma} \frac{\partial}{\partial x}\left[U'(x)\, \hat{P} + \frac{\partial \hat{P}}{\partial x}\right] \tag{9.3.92}$$

and the *exact* result for the mean first passage time from $x = a$ to $x = 0$ for this approximate equation is

$$\tau_1 = \gamma \int_a^0 dx \exp[U(x)] \int_{-\infty}^{x} dz \exp[-U(z)]. \tag{9.3.93}$$

This result can be evaluated numerically.

b) Kramers' Result
This is obtained by applying our method to the one-dimensional Smoluchowski equation (9.3.92) and making Gaussian approximations to all integrals. The result is

$$\tau_2 = \tfrac{1}{2}\gamma\, e^{U_0} \mathscr{N}_1^{-1} \sqrt{\frac{2\pi}{U_2}} \tag{9.3.94}$$

which differs from (9.3.90) for τ_0 by the replacement $\lambda \to \gamma$, which is clearly valid in a large γ limit. In this limit,

$$\tau_0 \simeq (1 + U_2\gamma^{-2})\tau_2. \tag{9.3.95}$$

c) Corrected Smoluchowski
A more accurate equation than the Smoluchowski equation (9.3.1) is the *corrected Smoluchowski equation* (6.4.108);

$$\frac{\partial \hat{P}}{\partial t} = \frac{1}{\gamma} \frac{\partial}{\partial x}\left\{[1 + \gamma^{-2}U''(x)]\left[U'(x)\, \hat{P} + \frac{\partial \hat{P}}{\partial x}\right]\right\}. \tag{9.3.96}$$

One now calculates the exact mean first passage time for this equation using standard theory; it is

$$\tau_3 = \gamma \int_a^0 dx\,[1 + \gamma^{-2}U''(x)] \exp[U(x)] \int_{-\infty}^{x} dz \exp[-U(z)]\,. \tag{9.3.97}$$

Note however, that the principal contribution to the x integral comes from near $x = 0$ so that the small correction term $\gamma^{-2}U''(x)$ should be sufficiently accurately evaluated by setting

$$U''(x) \simeq U''(0) = -U_2 \tag{9.3.98}$$

in (9.3.97). We then find the *corrected Smoluchowski result,*

$$\tau_3 = (1 - \gamma^{-2}U_2)^{-1}\tau_1 \simeq (1 + \gamma^{-2}U_2)\tau_1\,. \tag{9.3.99}$$

Notice that in this limit,

$$\frac{\tau_3}{\tau_1} = \frac{\tau_0}{\tau_2} \tag{9.3.100}$$

which means that in the limit that all integrals may be evaluated as sharply peaked Gaussians, our result is in agreement with the corrected Smoluchowski.

d) Simulations

By computer simulation of the equivalent stochastic differential equations

$$dx = p\,dt \tag{9.3.101}$$

$$dp = -[\gamma p + U'(x)]dt + \sqrt{2\gamma}\,dW(t)\,, \tag{9.3.102}$$

we can estimate the mean first passage time *to the plane* S_0, i.e., to the line

$$p = -\lambda x\,. \tag{9.3.103}$$

The results have to be computed for a given set of potentials. In order to assess the effect of the sharpness of peaking, we consider different temperatures T, i.e., we consider

$$dx = p\,dt \tag{9.3.104}$$

$$dp = -[\gamma p + U'(x)]dt + \sqrt{2\gamma T}\,dW(t)\,. \tag{9.3.105}$$

By the substitutions

$$\begin{aligned} p &\to pT^{1/2} \\ x &\to x\,T^{1/2}\,, \end{aligned} \tag{9.3.106}$$

we obtain

$$\left.\begin{aligned} dx &= p\,dt \\ dp &= -[\gamma p + V'(x, T)] + \sqrt{2\gamma}\,dW(t) \end{aligned}\right\} \tag{9.3.107}$$

where

$$V(x, T) = U(xT^{1/2}) . \tag{9.3.108}$$

The simulations were performed with

$$U(x) = \tfrac{1}{4}(x^2 - 1)^2 \tag{9.3.109}$$

and the results are shown in Fig. 9.4. They separate naturally into two sets: curved, or straight lines. The best answer is the corrected Smoluchowski which agrees with the simulations at all temperatures, and at low temperatures, agrees with our method. Thus, we confirm the validity of the method in the region of validity expected, since low temperature corresponds to sharply peaked distributions.

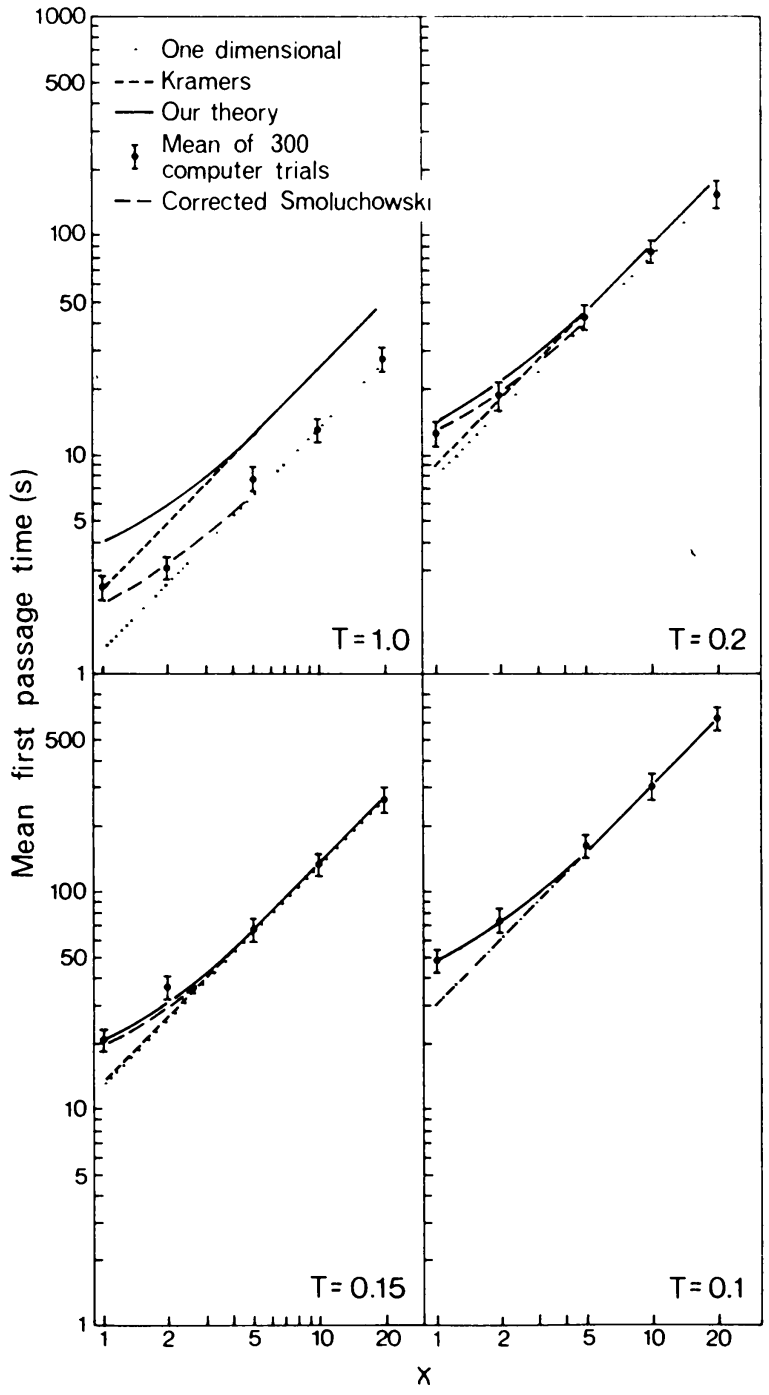


Fig. 9.4. Comparison of various estimates of the mean exit time from the double well potential of Sect. 9.3.4

Notice also that the choice of the plane S_0 as the separatrix is appropriate on another ground. For, near to $x = 0, p = 0$, we can write

$$dx = p\,dt \tag{9.3.110}$$

$$dp = (-\gamma p + U_2 x)dt + \sqrt{2\gamma T}\,dW(t)\,. \tag{9.3.111}$$

The condition that the deterministic part of (dx, dp), namely, $(p, -\gamma p + U_2 x)$ is in the direction connecting the point (x, p) to the origin is

$$\frac{p}{-\gamma p + U_2 x} = \frac{x}{p}\,. \tag{9.3.112}$$

Putting $p = -\lambda x$, we find

$$\lambda^2 - \lambda\gamma - U_2 = 0 \tag{9.3.113}$$

which is the same as (9.3.81) near $x = 0$. The two solutions correspond to the deterministic motion pointing towards the origin (+ve root) or pointing away from the origin (−ve root).

Thus, when the particle is on the separatrix, in the next time interval dt, only the random term $dW(t)$ will move it off this separatrix and it will move it right or left with equal probability, i.e., this means that the splitting probability, to left or right, should be 1:1 on this plane.

This separatrix definition also agrees with that of Sects. 9.1, 2 where the $v(x)$ should be perpendicular to the normal to S.

10. Quantum Mechanical Markov Processes

Quantum mechanics, since the very early times in the 1920's, has been recognised as a description of the world which contains an essentially statistical aspect. Hence, all quantum mechanics must be regarded as being some kind of stochastic process. However, what is essentially unique to quantum mechanics is the description in terms of *complex probability amplitudes*, the square of whose modulus gives the actual probability of occurrence of an event.

The formulation of a proper quantum mechanical probability theory, or of quantum mechanics in terms of appropriately defined stochastic processes in this generalised probability theory, is not the aim of this chapter. What is of interest is the introduction of the reader to the rather fascinating world which straddles the boundaries of quantum and classical probability theory. This world is the realm of quantum optics and quantum electronics, where there are statistical aspects arising from the intrinsic quantum nature of the system, as well as fluctuations arising from thermal effects. We shall show how the quantum mechanics of optical systems can be related closely to Markov jump processes in a suitably generalised form, which can themselves very frequently be related by means of what are known as *P-representations* or otherwise, as *phase-space methods*, to diffusion processes in the complex plane. These diffusion processes can describe quasiprobabilities which may be negative or complex, or they may define genuine positive probabilities. The situation is very similar to that of the Poisson representation of Sect. 7.7 which is itself, in fact, a restricted form of P-representation.

We will formulate this chapter as follows. We first outline the quantum mechanics of the harmonic oscillator and introduce the concept of coherent states, which are central to the task. We then define a quantum Markov process and show how generalised Master equations can be derived for these, in a manner similar to that of the adiabatic elimination methods of Chap. 6. From these generalised Master equations we can sometimes develop ordinary birth-death Master equations, and sometimes, by using P-representations, we can develop Fokker-Planck equations. Both methods allow us to apply all the apparatus of classical stochastic processes to these quantum mechanical systems.

10.1 Quantum Mechanics of the Harmonic Oscillator

We describe the Harmonic oscillator in terms of creation and destruction operators a^+ and a which satisfy the commutation relations

$$[a, a^+] \equiv aa^+ - a^+a = 1 \tag{10.1.1}$$

References

Chapter 1

1.1 Lord Rayleigh: 1891 [Phil. Mag. **32**, Scientific Papers III, 473, (1902)]
1.2 A. Einstein: Ann. Phys. (Leipzig) **17**, 549 (1905)
1.3 M. von Smoluchowski: Ann. Phys. (Leipzig) **21**, 756 (1906)
1.4 P. Langevin: Comptes. Rendues **146**, 530 (1908)
1.5 M. von Smoluchowski: Phys. Z. **16**, 321 (1915); **17**, 557, 585 (1916)
1.6 B.J. Berne, R. Pecora: *Dynamic Light Scattering* (Wiley, New York 1976)
1.7 A.J. Lotka: *Elements of Physical Biology* (Williams and Wilkins, Baltimore 1925) [Reissued as "Elements of Mathematical Biology", Dover, 1956]
1.8 V. Volterra: Mem. Acad. Lincei **2**, 31–113 (1926)
1.9 S.O. Rice: Bell Syst. Tech. J. **23**, 282 (1944); **24**, 46 (1945)
1.10 W. Schottky: Ann. Phys. (Leipzig) **57**, 541 (1918)
1.11 G.I. Taylor: Proc. London Math. Soc. **22**, 196 (1920)
1.12 N. Wiener: Acta Math. **55**, 117 (1930)
1.13 A. Khinchin: Math. Annalen **109**, 604 (1934)
1.14 J.B. Johnson: Phys. Rev. **32**, 97 (1928)
1.15 H. Nyquist: Phys. Rev. **32**, 110 (1928)
1.16 C.G. Huffaker: Hilgardia **27**, 343–383 (1958)
1.17 M.P. Hassell: *The Dynamics of Arthropod Predator-Prey Systems,* Monographs in Population Biology, No. 13 (Princeton University Press, Princeton 1978)

Chapter 2

2.1 W. Feller: *An Introduction to Probability Theory and Its Applications,* 2nd ed., Vol. II (Wiley, New York 1974)
2.2 A. Papoulis: *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York 1965)
2.3 A.N. Kolmogorov: *Foundations of the Theory of Probability* (Chelsea, New York 1950) [The German original appeared in 1933]
2.4 R.L. Stratonovich: *Introduction to the Theory of Random Noise* (Gordon and Breach, New York, 1963)

2.5 B.V. Gnedenko: *The Theory of Probability*, 2nd ed. (Chelsea, New York 1962)
2.6 E. Meeron: J. Chem. Phys. **27**, 1238 (1957)
2.7 N.G. van Kampen: Physica **74**, 215 (1973); **74**, 239 (1973)
2.8 J. Marcienkiewicz: Math. Z. **44**, 612 (1939)
2.9 A.K. Rajagopal, E.C.G. Sudarshan: Phys. Rev. **A 10**, 1852 (1974)

Chapter 3

3.1 F. Haake: Springer Tracts in Modern Physics, Vol. 66 (Springer, Berlin, Heidelberg, New York 1973)
3.2 I.F. Gihman, A.V. Skorohod: *The Theory of Stochastic Processes* Vols. I, II, III. (Springer, Berlin, Heidelberg, New York 1975)
3.3 [Ref. 3.2, Vol. II]
3.4 W. Feller: Ann. Math. **55**, 227–246 (1952)
3.5 G.E. Uhlenbeck, L.S. Ornstein: Phys. Rev. **36**, 823 (1930)
3.6 A. Papoulis: *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York 1965)

Chapter 4

4.1 I.I. Gihman, A.V. Skorohod: *Stochastic Differential Equations* (Springer, Berlin, Heidelberg, New York 1972)
4.2 R.L. Stratonovich: *Introduction to the Theory of Random Noise* (Gordon and Breach, New York, London 1963)
4.3 L. Arnold: *Stochastic Differential Equations* (Wiley-Interscience, New York 1974)
4.4 R. Kubo: In *Stochastic Processes in Chemical Physics*, ed. by K.E. Shuler (Wiley-Interscience, New York 1969)
4.5 N.G. van Kampen: Phys. Rept. **24C**, 171–228 (1976)

Chapter 5

5.1 A.D. Fokker: Ann. Phys. (Leipzig) **43**, 310 (1915)
5.2 M. Planck: Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Kl., **325** (1917)
5.3 A.N. Kolmogorov. Math. Ann. **104**, 415–418 (1931)
5.4 W. Feller: Ann. Math. **55**, 227–246 (1952)
5.5 A.T. Bharucha-Reid: *Elements of the Theory of Markov Processes and Their Application* (McGraw-Hill, New York 1960)
5.6 A. Erdelyi: *Bateman Manuscript Project*, Vols. 1–3 (McGraw-Hill, New York 1953)
5.7 N.G. van Kampen: Physica **23**, 707 (1957); **23**, 816 (1957)

5.8 U. Uhlhorn: Ark. Fys. 17, 361 (1960)
5.9 R. Graham, H. Haken: Z. Phys. **243**, 289–302 (1971)
5.10 H.A. Kramers: Physica 7, 284 (1940)
5.11 L. Onsager: Phys. Rev. 37, 405 (1931)
5.12 H.B.G. Casimir: Rev. Mod. Phys. 17, 343 (1945)

Chapter 6

6.1 H. Haken: *Synergetics* An Introduction, 2nd ed. (Springer, Berlin, Heidelberg, New York 1978)
6.2 H.C. Brinkman: Physica 22, 29 (1956)
6.3 R.L. Stratonovich: *Introduction to the Theory of Random Noise* (Gordon and Breach, New York, London 1963)
6.4 G. Wilemski: J. Stat. Phys. **14**, 153–170 (1976)
6.5 U.M. Titulaer: Physica **91A**, 321–344 (1978)
6.6 G.C. Papanicolauou: In *Modern Modelling of Continuum Phenomena,* Lecture Notes in Applied Maths, Vol. 16 (Am. Math. Soc., Providence RI 1977), p. 109
6.7 G.C. Papanicolauou, W. Kohler: Comm. Pure Appl. Math. **27**, 641–668 (1974)
6.8 R. Graham, T. Tel: J. Stat. Phys. **35**, 729 (1984)
6.9 R. Graham, T. Tel: Phys. Rev. **A31**, 1109 (1985)
6.10 R. Beals, V. Protopescu: J. Stat. Phys. **32**, 565 (1983)
6.11 J.V. Selinger, U.M. Titulaer: J. Stat. Phys. **36**, 293 (1984)
6.12 U.M. Titulaer: J. Stat. Phys. **37**, 589 (1985)
See also U. M. Titulaer: Phys. Lett. **A 108**, 19 (1985) for a treatment of partial absorption, and T. W. Marshall, E. J. Watson: J. Phys. **A 18**, 3531 (1985); **A 20**, 1345 (1987) for an exact solution.

Chapter 7

7.1 I. Matheson, D.F. Walls, C.W. Gardiner: J. Stat. Phys. **12**, 21 (1975)
H.K. Janssen: Z. Phys. **260**, 67 (1974)
7.2 N.G. van Kampen: Can. J. Phys. **39**, 551–567 (1961)
7.3 H.A. Kramers: Physica 7, 284 (1940)
7.4 J.E. Moyal: J. R. Stat. Soc. **11**, 151–210 (1949)
7.5 A. Einstein: Ann. Phys. (Leipzig) 549 (1905)
7.6 T.G. Kurtz: J. Chem. Phys. **50**, 460 (1969)
7.7 S. Karlin, H.M. Taylor: *A First Course in Stochastic Processes,* 2nd ed. (Academic Press, New York 1975)
7.8 H. Haken: *Synergetics* An Introduction, 2nd ed. (Springer, Berlin, Heidelberg, New York 1978)

7.9 J. Schnakenberg: Rev. Mod. Phys. **48**, 571–586 (1976); also *Thermodynamic Network Analysis of Biological Systems*, 2nd ed. (Springer, Berlin, Heidelberg, New York 1981)
7.10 C.W. Gardiner, S. Chaturvedi: J. Stat. Phys. **17**, 429–468 (1977); **18**, 501–522 (1978)
7.11 W. Bernard, H.B. Callen: Rev. Mod. Phys. **31**, 1017 (1959), Phys. Rev. **118**, 1466 (1960)
7.12 K. Hochberg: Ann. Probab. **3**, 433–458 (1978)
7.13 N.G. van Kampen: Adv. Chem. Phys. **34**, 245–309 (1976)

Chapter 8

8.1 R.F. Curtain: "Stochastic Partial Differential Equations", in *Stochastic Nonlinear Systems*, ed. by L. Arnold, R. Lefever (Springer, Berlin, Heidelberg, New York 1981)
8.2 J. Keizer: J. Chem. Phys. **64**, 1679–1687 (1976)
8.3 G. Nicolis: J. Stat. Phys. **6**, 195 (1972)
8.4 F. Schlögl: Z. Phys. **253**, 147 (1972)
8.5 L. Arnold: "Consistency of Models of Chemical Reactions" in *Dynamics of Synergetic Systems*, ed. by H. Haken (Springer, Berlin, Heidelberg, New York 1980)
8.6 N.G. van Kampen: "Fluctuations in Continuous Systems", in *Topics in Statistical Mechanics and Biophysics*, ed. by R.A. Piccirelli (Am. Inst. of Physics, 1976)
8.7 C. van den Broek, W. Horsthemke, M. Malek-Mansour: Physica **89A**, 339–352 (1977);
L. Brenig, C. van den Broek: Phys. Rev. **21A**, 1039 (1980)
8.8 H. Grad: "Principles of the Kinetic Theory of Gases", in *Handbuch der Physik*, Vol. 12, ed. by S. Flügge (Springer, Berlin, Göttingen, Heidelberg 1958)
8.9 C.W. Gardiner, M.L. Steyn-Ross: Phys. Rev. **29A**, 2823 (1984)

Chapter 9

9.1 M. Abramowitz, I. Stegun: *Handbook of Mathematical Functions* (Dover, New York 1964, 1965)
9.2 H.A. Kramers: Physica **7**, 284 (1940)
9.3 A summary of the situation is given in the appendix to M. Buttiker, R. Landauer: "Transport and Fluctuations in Linear Arrays of Multistable Systems" in *Nonlinear Phenomena at Phase Transitions and Instabilities*, ed. by T. Riste (Plenum Press, New York, London 1982)

9.4 H. Eyring: J. Chem. Phys. **3**, 1935 (1935)
9.5 S.A. Hawley: Biochem. **10**, 2436 (1971);
S.A. Hawley, R.M. Mitchell: Biochem. **14**, 3257 (1975)
9.6 J.S. Langer: Ann. Phys., N.Y. **54**, 258 (1969);
R. Landauer, J.A. Swanson: Phys. Rev. **121**, 1668 (1961)
9.7 Z. Schuss: S.I.A.M. Rev. **22**, 119 (1980);
B.J. Matkowsky, Z. Schuss: S.I.A.M. J. Appl. Math. **33**, 365 (1977)
9.8 A.D. Ventsel, M.I. Freidlin: Russian. Math. Surveys **25**, 1 (1970)

Chapter 10

10.1 R.J. Glauber: Phys. Rev. **130**, 2529 (1963); **131**, 2766 (1963)
10.2 E.C.G. Sudarshan: Phys. Rev. Lett. **10**, 277 (1963)
10.3 J.R. Klauder, E.C.G. Sudarshan: *Fundamentals of Quantum Optics* (Benjamin, New York 1970)
10.4 R.J. Glauber: In *Quantum Optics*, ed. by S. Kay, A. Maitland (Academic Press, New York 1970)
10.5 W. Louisell: In *Quantum Optics*, ed. by S. Kay, A. Maitland (Academic Press, New York 1970); also *Quantum Statistical Properties of Radiation* (Wiley, New York 1974)
10.6 H.J. Carmichael, D.F. Walls: J. Phys. **B9**, 1199 (1976)
10.7 A. Einstein: Phys. Z. **18**, 121 (1917)
10.8 M. Lax: Phys. Rev. **172**, 350 (1968)
10.9 P.D. Drummond, C.W. Gardiner: J. Phys. A **13**, 2353–2368 (1980)

Stratonovich, R.L.: *Topics in the Theory of Random Noise,* Vols. I, II (Gordon and Breach, New York 1963)

A compendium of a radio engineer's view of stochastic methods. Contains many beautiful results in a characteristic style which is often very illuminating. Stochastic differential equations are introduced as limits of real processes. One of the fundamental books on applied stochastic differential equations

van Kampen, N.G.: *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, New York, Oxford 1981)

A book on the application of stochastic processes to physical and chemical phenomena. Characterised by clear reasoning, and a precise style, with particular attention to evaluation of the conditions of validity of approximations used commonly in Physical and Chemical applications

Wax N.: *Selected Papers on Noise and Stochastic Processes* (Dover, New York 1954)

A fundamental historical reference, containing the original work of Uhlenbeck and Ornstein on Brownian Motion, and Chandrasekhar's classic paper on Stochastic Problems in Physics and Astronomy, as well as other classic papers

Symbol Index

The pages referred to are normally the first mention of the symbol in text. The order is alphabetical, lower case first, and ignoring typeface. There are three parts to the index; Latin, Greek, and Mathematical.

Latin Symbols

a^+, a	Creation and destruction operators	373
ac-lim	Almost certain limit	40
$A(z,t)$	Drift vector	47, 52
$A_a[J(\alpha)]$	Drift vector in Poisson representation FPE	278
$\hat{A}_a(\eta)$	Drift vector in Poisson representation	279
$\underset{\sim}{B}(z,t)$	Diffusion matrix	47, 52
$B_{ab}[J(\alpha)]$	Diffusion matrix in Poisson representation FPE	278
$\hat{B}_{ab}(\eta)$	Diffusion matrix in Poisson representation, after removing a factor $\varepsilon^2 = V^{-1}$	280
d_{ij}	Diffusion jump probability	307
$d\tilde{W}(q,t)$	Fourier transform of local Wiener process	326
$dW(r,t)$	Spatially dependent, local, Wiener process	314, 325
$\mathscr{D}$	Domain of integration in Poisson representations	282
$\underset{\sim}{D}(r)$	Non-diagonal spatially dependent diffusion tensor	310
$(\mathscr{D}\eta)(r,t)$	Nonlocal diffusion	325
$\operatorname{erf}(x) = \sqrt{\frac{\pi}{2}} \int_0^x dt\, e^{-t^2}$	Error function	346
$f(\alpha,t)$	Poisson representation quasiprobability	277
$f(\alpha,t \mid \alpha',t')$	Poisson representation conditional quasiprobability	289
$G(s,t)$	Generating function	12, 38, 238
$G(\tau)$	Autocorrelation function	16
H	Hamiltonian operator	374
$H_n(x)$	Hermite polynomial	134
$j(r,t)$	Diffusion current	303
$J(x)$	Probability current in a jump process	236
$J(z,t)$	Probability current	119
$J_A(\alpha)$	Probability current in Poisson representation	278

$\hat{J}_A[\eta]$	Probability current in Poisson representation	279
$k_A^{\pm}$	Chemical master equation rate constant	263
$L_n(x)$	Laguerre polynomial	136
ms-lim	Mean square limit	40
M^A	Final state parameter in chemical master equation	263
$\mathcal{N}$	Normalisation constant	124
N^A	Initial state parameter in chemical master equation	263
$p(x_1, t_1; x_2, t_2; x_3, t_3; \ldots)$	Joint probability density	42
$p(x_1, t_1 \mid x_2, t_2)$	Conditional probability density	43
$P(A \mid B)$	Conditional probability	25
P	Projection operator in adiabatic elimination	198, 212
$P(x, t \mid x', t')$	Conditional probability in a jump process	52, 236
$P(\alpha, \alpha^*)$	Glauber Sudarshan P-function	382
$P(\alpha, \beta)$	Generalized P-function	410
$P_\lambda(x)$	Eigenfunction of forward FP operator	130
q	Fourier transform variable	326
$Q_\lambda(x)$	Eigenfunction of backward FP operator	130
$r^A = M^A - N^A$		263
$R(\alpha^*, \beta)$	Glauber R-function	409
st-lim	Stochastic limit	41
$S(\omega)$	Spectrum	16
S^+, S^-, S_z	Pauli matrices	399
$t^+(x), t^-(x)$	Transition probability per unit time for the jumps $x \to x + 1$, $x \to x - 1$	236
$t_A^{\pm}(\boldsymbol{x})$	Transition probability per unit time in chemical master equations	263
$T(x)$	Mean first passage time	137, 259
var$\{X\}$	Variance of X	30
V, Ω	System volume	277
$W(t)$	Wiener process	66, 67
$\boldsymbol{W}(t)$	Multivariate Wiener process	67
$W(\boldsymbol{x} \mid z, t)$	Jump probability	47
$x[V, t] = \int_V d^3\boldsymbol{r}\rho(\boldsymbol{r}, t)$	Number of molecules in a volume V	320

Greek Symbols

$\alpha, \boldsymbol{\alpha}, \alpha(t), \boldsymbol{\alpha}(t)$	Poisson representation variable	277
$\underset{\sim}{\alpha} = (\alpha, \beta)$		411

$\alpha_n(x)$	nth derivate moment	249
Γ_i	Bath operator	388
$\delta(x)$	Dirac delta function	
$\delta(m,n) \equiv \delta_{m,n}$	Kronecker delta	34
$\dfrac{\delta F(\rho)}{\delta \rho(\mathbf{r})}$	Functional derivative	305
η	Poisson representation concentration variable	279
$\kappa_A^{\pm}$	Chemical rate constant in terms of concentrations	266
$\Lambda(\alpha,\beta) = \lvert\alpha\rangle\langle\beta^*\rvert/\langle\beta^*\rvert\alpha\rangle$		410
$\mu(\alpha)$	Measure used in Poisson representations	282
$\xi(t)$	Fluctuating force	80
$\pi_a(x)$	Probability of exit through a	142
ρ	Density matrix (or operator)	380
$\rho(T)$	Density operator for canonical ensemble	388
$\rho(\mathbf{r},t)$	Concentration variable	304
$\sigma[X]$	Standard deviation of X	30
τ_C	Correlation time	20
$\Phi(s)$	Cumulant generating function	33
$\phi(s)$	Moment generating function	32
$\chi(\lambda,\lambda^*), \chi(\lambda)$	Quantum characteristic function	386
Ω, V	System size parameter	250

Mathematical Symbols

$[A,B] = AB - BA$	Commutator of operators A, B	230
$\langle X_i, X_j\rangle \equiv \langle X_iX_j\rangle - \langle X_i\rangle\langle X_j\rangle$	Covariance matrix of X_i and X_j	30
$\langle R\rangle$	Mean value of R	28
$\langle X^r\rangle_f$	Factorial moment of X	38
$\langle\alpha(t)\rvert[\alpha',s]\rangle$	Mean of $\alpha(t)$ under initial condition $\alpha(s) = \alpha'$	289
$\langle\langle X^r\rangle\rangle_f$	Factorial cumulant of X	38
$\langle\langle X_1^{m_1}X_2^{m_2}\ldots X_n^{m_n}\rangle\rangle$	Cumulants of variables, $X_1, X_2, \ldots, X_n$	33
$\lvert n\rangle$	Number state	374
$\lvert\alpha\rangle$	Coherent state	376
$\lVert\alpha\rangle$	Bargmann state	377
$\int_{t_0}^{t} G(t')\,dW(t')$	Ito stochastic integral	84

$S\int_{t_0}^{t} G(t')\,dW(t')$	Stratonovich stochastic integral	86
$\tilde{f}(s) = \int_{0}^{\infty} \mathrm{e}^{-st} f(t)$	Laplace transform of $f(t)$	201

Author Index

Beals, R. 423
Bernard, W. 292, 423
Boltzmann, L. 3
Brinkman, H. C. 197, 423
Brown, R. 2

Callen, H. B. 292, 423
Carmichael, H. J. 400, 408, 425
Casimir, H. B. G. 161, 423

Drummond, P. D. 409, 425

Einstein, A. 2, 249, 402, 421, 423, 425

Feller, W. 21, 421, 422
Fokker, A. D. 117, 422
Freidlin, M. I. 357, 424

Gardiner, C. W. 410, 423, 425
Glauber, R. J. 150, 375, 382, 424, 425
Graham, R. 154, 194, 423

Haken, H. 150, 154, 223, 266, 423
Hochberg, K. 300, 423

Johnson, J. B. 18, 421

Keizer, J. 314, 424
Khinchin, A. 17, 421
Klauder, J. R. 383, 409, 425
Kohler, W. 215
Kolmogorov, A. N. 23, 421, 422
Kramers, H. A. 155, 249, 349, 423, 424
Kurtz, T. G. 254, 423

Landauer, R. 357, 424
Langer, J. S. 357, 424
Langevin, P. 5, 6, 421
Lax, M. 404, 425
Lotka, A. J. 9, 421
Louisell, W. 394, 425

Marcienkiewicz, J. 36, 422
Matkowsky, B. J. 357, 358, 424
Meeron, E. 34, 422
Moyal, J. E. 249, 250, 423

Nyquist, H. 18, 421

Onsager, L. 161, 423
Ornstein, L. S. 75, 422

Papanicolaou, G. C. 197, 215, 423
Papoulis A. 21, 421, 422
Planck, M. 117, 422
Protopescu, V. 423

Rayleigh, Lord 3, 421
Rice, S. O. 13, 421

Schnakenberg, J. 266, 423
Schottky, W. 13, 421
Schuss, Z. 357, 358, 424
Selinger, J. V. 423
Smoluchowski, M. von 3, 7, 421
Stratonovich, R. L. 34, 86, 197, 216, 421, 422, 423
Sudarshan, E. C. G. 382, 383, 409, 424, 425

Taylor, G. I. 16, 421
Tel, T. 209, 423
Titulaer, U. M. 197, 209, 423

Uhlenbeck, G. E. 75, 422
Uhlhorn, U. 150, 423

Van den Broek, C. 331, 424
Van Kampen, N. G. 34, 150, 154, 161, 235, 246, 331, 422, 423, 424
Ventsel, A. D. 357, 424
Volterra, V. 9, 421

Walls, D. F. 400, 408, 423, 425
Wiener, N. 66, 421
Wilemski, G. 197, 423

Subject Index

Absorbing boundary condition for
  backward FPE 129
Additive invariants 336
Adiabatic elimination
  in quantum systems 391
  of fast variables 195 – 234
  of inhomogeneous modes in reaction diffusion systems 328 – 331
Almost certain 29
Arrhenius formula 141
Autocorrelation function 16, 58
  and Poisson rep. 289 – 294
  for Markov process 64 – 66
  in terms of eigenfunctions 131
  with detailed balance 149
Autocorrelation matrix
  and eigenfunctions 168
  stationary 65

Backward differential Chapman-Kolmogorov equation 55
Backward FPE, boundary conditions for 128
Backward master equation
  and first passage times 259
Baker-Hausdorff formula 387
Bargmann states, defined 377
Bath correlation functions 389
Bernoulli distribution 72
Bernoulli trials 42
Birth-death master equation 10, 13
  and bistability 354 – 357
  in quantum harmonic oscillator 395
  one-variable 236 – 262
Birth-death process 8
  for flow 332 – 336
Birth-death processes, boundary conditions for 257 – 259
Birth-death systems, many-variable 262 – 277
Bistability 342
  in multivariable systems 357 – 371
  with birth-death master equation 354 – 357
Bistable system, chemical 241 – 245
Boltzmann equation from Boltzmann master equation 337
Boltzmann master equation 303, 336 – 339
  in Poisson representation 338, 341

Boundary conditions for
  backward FPE 128
  birth-death processes 257 – 259
  Kramers' equation 205
  many-variable FPE 146
  Smoluchowski equation 205
Boundary conditions for FPE 118 – 125
  absorbing 121
  at a discontinuity 121
  at infinity 123
  reflecting 121
Brownian motion 2
  and adiabatic elimination 195
  continuous sample path 46
  corrections to Smoluchowski equation 206 – 210
  in double potential 365 – 371
  Kramers' equation for 155
Brownian particle 6

Cauchy process 46 – 47
Central limit theorem 37
Chaos 1
Chapman-Kolmogorov equation 5, 43 – 44
  differential 47 – 51
Characteristic function 32
  of a Gaussian 36
  quantum 386 – 388
Chemical FPE 266
Chemical intermediates, elimination of 218
Chemical reaction
  $A + 2X \rightleftharpoons 3X$, $A \rightleftharpoons X$ 300 – 301
  $A + 2X \rightleftharpoons 3X$, $A \rightleftharpoons X$, master equation 241
  $A + 2X \rightleftharpoons 3X$, in Poisson rep. 294
  $A + X \rightarrow 2X + D$, $B + X \rightleftharpoons C$, solution of ME 274 – 276
  $A + X \rightleftharpoons 2X$, $B + X \rightleftharpoons C$, positive Poisson rep. 288
  $A + X \rightleftharpoons 2X$, $B + X \rightleftharpoons C$, Poisson rep. solution 279 – 281
  $A + X \rightleftharpoons 2X$, $B + X \rightleftharpoons C$, complex Poisson rep. 283 – 284
  $A + X \rightleftharpoons 2X$, $B + X \rightleftharpoons C$, spatially distributed 324 – 328

Chemical reaction
$A + Y \rightleftharpoons X + Y$, $Y \rightleftharpoons 2X$, elimination of intermediate 295 – 299
$B + X \rightleftharpoons C$, $A + X \to 2X$, spatially distributed 319 – 324
$B \to X$, $2X \to A$, and positive Poisson rep. 288 – 289
$B \to X$, $2X \to A$, complex Poisson rep. 284 – 285
$B \rightleftharpoons X$, $A + X \to 2X$, analogy to harmonic oscillator 395
$X + A \rightleftharpoons 2X$, FPE 127
$X + A \rightleftharpoons 2X$, chemical FPE 267 – 273
$X \rightleftharpoons A$, and system size expansion 252
$X \rightleftharpoons A$, master equation 238
$X \rightleftharpoons Y$, as reaction diffusion system 315 – 318
$X \rightleftharpoons Y \rightleftharpoons A$ 218
$X \rightleftharpoons Y \rightleftharpoons A$, chemical FPE 267 – 273
$X_1 \rightleftharpoons 2X_2$, reaction diffusion equations 314
$X_1 \rightleftharpoons X_2$, solution of master equation 276 – 277
nonlocal 328
prey-predator system, FPE 268 – 273
unimolecular 279
Chemical reactions, transition state theory 351
Chymotrypsinogen, reversible denaturation of 352
Coherent states
defined 375
expansion of an operator in 378
expansion of arbitrary states in terms of 377
Poissonian number distribution of 379
properties of 376 – 380
Collisions
and flow 339 – 341
Boltzmann master equation for 336 – 339
Combinatorial kinetics; defined 262
Commutation relations 373
Complex P-representation
defined 410
FPE from 416
operator identities for 414
Conditional probability 25
and eigenfunctions 131, 168
Continuity in stochastic processes 45 – 46
Corrected Smoluchowski equation and escape problem 369
Correlation 30
Correlation functions
defined 33
of heat bath 389
Correlation length 320
Correlation time 20, 58
Correlations
behaviour at instability point 323
of smooth functions of bath operators 389
space-time 318
Covariance 30
Covariance matrix 30
Creation operator 373
Critical fluctuations 255
Critical slowing down 257
Cubic process 185
Cumulant 33 – 36
Cumulant, factorial 39
Cumulant generating function 33
factorial 39

Density matrix 380 – 382
properties of 381 – 382
Destruction operator 373
Detailed balance 148 – 165
conditions for 151
defined 148
in birth-death master equation 237 – 238
in diffusion 307
Deterministic process 53
Differential Chapman-Kolmogorov equation 47 – 51
backward 55
Diffusion
and detailed balance 307
coefficient 5
current, defined 303
equation 5
fluctuating partial differential equation 312
in a gravitational field 126
in multivariate master equation 307 – 308
inhomogeneous anisotropic 309
Diffusion master equation
continuum form of 308 – 313
system size expansion of 308
Diffusion matrix, defined 52
Diffusion process 52
approximation by jump process 246 – 249
Divergence problems in nonlinear reaction diffusion systems 327
Double well potential 140
diffusion in 342 – 348
Drift vector; defined 52

Eigenfunctions
and autocorrelation matrix 168
and conditional probability 131, 168
and exit time 172
and spectrum matrix 169
for many-variable FPE 165
for one-variable FPE 129 – 136

Eigenfunctions for FPE, variational principle for 168
Einstein equations and two-level atom 402
Ensemble average 17
Entrance boundary for FPE 123
Equilibration of populations in bistable system 348 – 357
Ergodic 17 – 18
Ergodic properties and stationary processes 57
Escape 342
Escape probability of particle in a double well 351
Escape time 141
Event 21
Existence theorems for P-representations 411 – 415
Exit boundary for FPE 123
Exit points, distribution of by asymptotic method 357 – 362
Exit time
see also first passage time
and eigenfunctions 172
asymptotic analysis of 362 – 363
from a region, many variable FPE 170 – 171
in a double well potential 345
through a particular end of an interval 142

Factorial correlation in reaction diffusion system 315
Factorial cumulant 39
Factorial moment 38
Fast variable 197
Fick's law 303
with fluctuations 304
First passage time
for one dimensional FPE 136 – 142
of particle in a double well 351
Flow
and collisions 339 – 341
as a birth death process 332 – 336
in position space 334
in velocity space 334
Fluctuating force 6, 15
Fluctuation dissipation theorem 162
Flux, phenomenological 162
Fokker-Planck equation (FPE) 5, 8, 117 – 176
backward, boundary conditions 128
boundary conditions 118 – 125
chemical 266
connection with SDE 96
defined 52 – 53
from complex P-representation 416
from Poisson representation 278 – 285
from positive P-representation 416 – 418
from P-representation 396, 415
functional 305
many dimensional 143 – 170
many variable, boundary conditions 146
many variables, eigenfunctions 165
one dimensional 117 – 143
one variable, eigenfunctions 129 – 136
small noise expansion for 187 – 194
stationary solution for one variable 124
Stratonovich form 100
Force, phenomenological 161
Fourier analysis of fluctuating functions 17
FPE, *see* Fokker Planck equation
Functional derivative, defined 305
Functional FPE, for reaction diffusion systems 305 – 307

Gaussian distribution 36 – 38
Gaussian random variable 36
Generalised P-representations 408 – 419
and time development equations 415 – 419
defined 409
relation to Poisson representation 413
Generating function 13
for birth-death master equation 273 – 274
for Poisson distribution 38
Glauber-Sudarshan P-representation 382 – 384, 410

Haken's model 223
Harmonic oscillator
defined 374
driven by fluctuating field 398
driven damped 397 – 398
driven; in P-representation 384
interaction with external field 375
quantum master equation 395 – 399
quantum mechanics of 373
Heat bath; defined 388
Hermite polynomials 134
in adiabatic elimination 202
Homogeneous Markov process 60
defined 56

Independence 7, 27
Independent random variables 27
and characteristic function 32
Interaction picture 394
and driven atom 399
Ito stochastic integral; defined 84
properties of 88 – 92
Ito's formula; derivation of 95

Johnson noise 18 – 20
Joint probability 24
Jump process 52

Kramers' equation 155
boundary conditions 205
Kramers' method
for escape problems 349 – 352
in several dimensions 363 – 371
Kramers-Moyal expansion 5
and system size expansion 251
defined 249
in birth-death master equation 266
in Boltzmann master equation 337, 340
Kurtz's theorem 254

Laguerre polynomial 136
Langevin equation 11, 14, 80 – 83
Laplace transform and adiabatic elimination 200
Laser light scattering 7
Law of larse numbers 30
Limit of sequence of random variables 39
Limit
almost certain 40
in distribution 41
in probability 41
mean square 40
stochastic 40
Lindeberg condition 37, 46
Linear SDE
multivariable 113
single-variable 112
Liouville equation 53
quantum 382
Liouville operators; defined 391
Local and global descriptions of chemical reaction, connection between 328 – 331
Local and global fluctuations 320

Markov assumption 13, 43
Markov postulate 5, 10
Markov process
autocorrelation for 64 – 66
continuous; defined 46
defined 43
homogeneous; defined 56
quantum mechanical 373
stationary; defined 56
Master equation 51, 235 – 301
approximation by Fokker-Planck equation 246 – 257
many-variable 262 – 277
many-variable; Kramers-Moyal expansion for 266
mean first passage times 259 – 262
one-variable; stationary solution of 236 – 238
phase space 331 – 341
quantum; defined 394
Master equation (cont.)
quantum; derivation 390 – 394
stationary solutions without detailed balance 266
ME, *see* Master equation
Mean first passage time
in one-variable FPE 137
for master equations 259 – 262
Mean square limit in definition of stochastic integral 84
Mean value 28, 29
Metastability 342
Microscopic reversibility 159
Moment 30
Moment generating function 32
Moments and characteristic function 32
Multitime averages in quantum systems 394
Multivariate master equation 303
for diffusing system 307 – 308

Natural boundary for FPE 122
Noise 11
Nonanticipating function; defined 86
Nonlinear absorber 415
Nyquist's formula 163
Nyquist's theorem 18 – 20

Onsager relations 160 – 165
Operator correspondences in Glauber-Sudarshan *P*-representation 383
Operator identities
for complex *P*-representation 414
for positive *P*-representation 414 – 415
Ornstein-Uhlenbeck process 74 – 77
and Onsager relations 160
and quantum harmonic oscillator 396
eigenfunctions 134
exit time from a circle 171
in small noise expansion of SDE 181
spectrum matrix 111
stationary distribution 126
stationary variance 110
stationary variance in two dimensions 110
time correlation function 76
time correlation matrix in stationary state 111
time dependent 115
Ornstein-Uhlenbeck process (SDE) 106 – 107
multivariate 109 – 111
Oscillator, noisy 105

P-representation
complex; defined 410
existence theorems for 411 – 415
FPE from 396, 415
generalised 408 – 419
generalised; defined 409

Glauber-Sudarshan 382 – 384, 410
positive; defined 410
Pauli matrices, defined 399
Phase space master equation 331 – 341
Phase transition, second order; in reaction diffusion system 324
Phenomenological flux 162
Phenomenological force 161
Poisson distribution 13, 14, 38 – 39
and coherent states 379
Poisson distribution relaxation in reaction diffusion system 316 – 318
Poisson process 13, 72 – 75
Poisson representation 235, 277 – 301
and Boltzmann master equation 338, 341
and time correlation functions 289 – 294
and unimolecular reactions 279
complex; defined 282
in reaction diffusion system 314 – 315
positive; defined 285 – 289
real; defined 282
relation to generalised *P*-representation 413
Poissonian time dependent solutions of certain master equations 241
Positive *P*-representation
defined 410
FPE from 416 – 418
operator identities for 414 – 415
Potential barrier, escape over 138 – 141
Potential conditions
for FPE in complex *P*-representation 418 – 419
for many-variable FPE 146
Potential solution of one-variable FPE 124
Prey-predator system 8
chemical FPE for 268 – 273
Probability 22
a priori 23
conditional 25
joint 24
Probability axioms 22
Probability current; defined 119
Probability density 28
Projection operator in adiabatic elimination 198

Q-representation 409
Quantum characteristic function 386 – 388
Quantum Liouville equation 382
Quantum Markov process, derivation of 388 – 394
Quantum Markov process, time correlation functions in 402 – 408
Quantum master equation
defined 394
derivation of 390 – 394
for harmonic oscillator 395 – 399
Quantum mechanics of harmonic oscillator 373
Quantum regression theorem 404 – 405

R-representation 409
Random telegraph process 77 – 79
and driven atom 402
Random variable 24
Gaussian 36
Random variables
independent 27
independent (and central limit theorem) 37
Random walk 70 – 73
approximation by FPE 247
Rayleigh process 144, 147
eigenfunctions 135
Reaction diffusion chemical equation 305
Reaction diffusion master equation; defined 313
Reaction diffusion system 303
fluctuating partial differential equations 313
in Poisson representation 314
Reaction diffusion systems, divergence problems 327
Reflecting boundary condition for backward FPE 129
Regression theorem 65
quantum 404 – 405
Relaxation time, dependence on peak populations 352

Sample path, continuous 5, 45 – 47
Scaling assumption in approximation of master equation 246
Schrödinger's equation 374
SDE *see* Stochastic differential equation
Separatrix 361, 372
Set of events 21
Sets of probability zero 29
Shot noise 12
Sigma algebra 22
Slave
noisy 226
silent 225
Slaving principle 197
Small noise expansion
for SDE 180 – 187
of FPE 187 – 194
Smoluchowski equation 197
and escape problem 369
boundary conditions 205
corrected 209
corrections to 206 – 210
defined 197, 203
for diffusion in a double well 349

Spatial correlations 315 – 328
Spectrum 16
  flat 19
  of Ornstein-Uhlenbeck process 111
Spectrum matrix
  and eigenfunctions 169
  with detailed balance 150
Splitting probability in double well potential 345
Standard deviation 30
Stationary 18
Stationary distribution of FPE, asymptotic method for 194
Stationary Markov process; defined 56
Stationary process, approach to 61 – 63
Stationary solutions, many-variable FPE 146 – 148
Stationary systems 17
Statistical mechanics 6
Stochastic differential equation 5, 7, 14, 15, 80 – 115
  and positive Poisson rep. 287 – 289
  as limit of nonwhite process 210 – 218
  connection with FPE 96
  definition and properties of 92 – 102
  dependence of solutions on initial conditions 101
  dependence of solutions on parameters 101
  partial 305
  small noise expansion for 180 – 187
Stochastic integral
  defined 83
  Ito 84
  Stratonovich 86
Stochastic partial differential equation 305
Stochastic process
  defined 42
  separable 42
Stratonovich SDE
  as limit of nonwhite process 210 – 218
  definition and connection with Ito SDE 98 – 101
Stratonovich stochastic integral 86
System size 235
System size expansion
  and critical fluctuations 255
  and Kramers-Moyal expansion 251
  and Kurtz's theorem 254
  in birth-death master equation 266
  in Boltzmann master equation 337, 340
  in reaction diffusion master equation 313
  of master equation 250 – 257
Third-order noise 295
  defined 299 – 300
Three-state interpretation of bistable system 350
Time correlation function, *see* Autocorrelation function
Time correlation functions
  and P-representation 405 – 408
  and quantum harmonic oscillator 405 – 408
  for two-level atom 408
  in quantum Markov processes 402 – 408
Time reversal invariance 150
Transition probability 10
Transition state theory of chemical reactions 351
Trimolecular reaction as a limit of bimolecular reaction 294 – 299
Two-level atom
  driven 399 – 402
  time correlations 408

Unimolecular reactions in Poisson representation 279
Unstable state, decay from 347

Van Kampen's system size expansion 250
Variance 30
Variational principle for eigenfunctions of FPE 168
Viscosity 6
Volterra-Lotka system, *see also* Prey-predator system
Von Neumann's equation 382

White noise 19 – 20, 80 – 83
  as limit of nonwhite process 210 – 218
Wiener process 46, 47, 67 – 70
  as limit of random walk 72
  autocorrelation function 70
  eigenfunctions 132
  independence of increments 69
  irregularity of sample paths 68
  nondifferentiability of sample paths 68
Wiener-Khinchin theorem 17
Wigner representation 409

Printing: Druckhaus Beltz, Hemsbach
Binding: Buchbinderei Schäffer, Grunstadt

Springer Series in Synergetics

Editor: Hermann Haken

An ever increasing number of scientific disciplines deal with complex systems. These are systems that are composed of many parts which interact with one another in a more or less complicated manner. One of the most striking features of many such systems is their ability to spontaneously form spatial or temporal structures. A great variety of these structures are found, in both the inanimate and the living world. In the inanimate world of physics and chemistry, examples include the growth of crystals, coherent oscillations of laser light, and the spiral structures formed in fluids and chemical reactions. In biology we encounter the growth of plants and animals (morphogenesis) and the evolution of species. In medicine we observe, for instance, the electromagnetic activity of the brain with its pronounced spatio-temporal structures. Psychology deals with characteristic features of human behavior ranging from simple pattern recognition tasks to complex patterns of social behavior. Examples from sociology include the formation of public opinion and cooperation or competition between social groups.

In recent decades, it has become increasingly evident that all these seemingly quite different kinds of structure formation have a number of important features in common. The task of studying analogies as well as differences between structure formation in these different fields has proved to be an ambitious but highly rewarding endeavor. The Springer Series in Synergetics provides a forum for interdisciplinary research and discussions on this fascinating new scientific challenge. It deals with both experimental and theoretical aspects. The scientific community and the interested layman are becoming ever more conscious of concepts such as self-organization, instabilities, deterministic chaos, nonlinearity, dynamical systems, stochastic processes, and complexity. All of these concepts are facets of a field that tackles complex systems, namely synergetics. Students, research workers, university teachers, and interested laymen can find the details and latest developments in the Springer Series in Synergetics, which publishes textbooks, monographs and, occasionally, proceedings. As witnessed by the previously published volumes, this series has always been at the forefront of modern research in the above mentioned fields. It includes textbooks on all aspects of this rapidly growing field, books which provide a sound basis for the study of complex systems.

A selection of volumes in the Springer Series in Synergetics:

1 **Synergetics** An Introduction
3rd Edition By H. Haken

8 **Stochastic Nonlinear Systems**
in Physics, Chemistry, and Biology
Editors: L. Arnold, R. Lefever

13 **Handbook of Stochastic Methods**
2nd Edition By C. W. Gardiner

15 **Noise-Induced Transitions**
Theory and Applications in Physics, Chemistry, and Biology
By W. Horsthemke, R. Lefever

18 **The Fokker-Planck Equation**
2nd Edition By H. Risken

19 **Chemical Oscillations, Waves, and Turbulence** By Y. Kuramoto

20 **Advanced Synergetics** 2nd Edition
By H. Haken

21 **Stochastic Phenomena and Chaotic Behaviour in Complex Systems**
Editor: P. Schuster

24 **Chaos and Statistical Methods**
Editor: Y. Kuramoto

32 **Dimensions and Entropies in Chaotic Systems** Editor: G. Mayer-Kress

40 **Information and Self-Organization**
A Macroscopic Approach to Complex Systems By H. Haken

51 **Foundations of Synergetics I**
2nd Edition By A. S. Mikhailov

52 **Foundations of Synergetics II**
2nd Edition By A. S. Mikhailov, A. Yu. Loskutov

54 **Quantum Signatures of Chaos**
By F. Haake

56 **Quantum Noise** By C. W. Gardiner

57 **Nonlinear Nonequilibrium Thermodynamics I** Linear and Nonlinear Fluctuation-Dissipation Theorems
By R. Stratonovich

59 **Nonlinear Nonequilibrium Thermodynamics II** Advanced Theory
By R. Stratonovich

60 **Limits of Predictability**
Editor: Yu. A. Kravtsov

68 **Intelligent Measurements, Prediction and Control** By I. Grabec, W. Sachse

69 **Predictability of Complex Dynamical Systems** By Yu. A. Kravtsov, J. B. Kadtke

目　录

1. 历史介绍 …… 1
1.1 动机 …… 1
1.2 一些历史实例 …… 2
1.2.1 布朗运动 …… 2
1.2.2 郎之万方程 …… 6
1.3 生灭过程 …… 9
1.4 电子系统中的噪声 …… 12
1.4.1 散粒噪声 …… 12
1.4.2 自相关函数与谱 …… 17
1.4.3 涨落函数的傅里叶分析：平稳系统 …… 19
1.4.4 约翰逊噪声与尼奎斯定理 …… 20

2. 概率概念 …… 23
2.1 事件和事件集 …… 23
2.2 概率 …… 24
2.2.1 概率公理 …… 24
2.2.2 $P(A)$的意义 …… 25
2.2.3 公理的意义 …… 26
2.2.4 随机变量 …… 26
2.3 联合概率和条件概率，独立性 …… 27
2.3.1 联合概率 …… 27
2.3.2 条件概率 …… 28
2.3.3 不同次序的联合概率间的关系 …… 29
2.3.4 独立性 …… 29
2.4 均值和概率密度 …… 30

2.4.1 由任意函数的均值确定概率密度……31
2.4.2 零概率集……32
2.5 均值……32
2.5.1 矩、相关和协方差……33
2.5.2 大数法则……34
2.6 特征函数……35
2.7 累积母函数；相关函数和累积量……36
2.7.1 例：4阶累积量：$\langle X_1X_2X_3X_4\rangle$……38
2.7.2 累积量的意义……39
2.8 高斯和泊松概率分布……39
2.8.1 高斯分布……39
2.8.2 中心极限定理……41
2.8.3 泊松分布……42
2.9 随机变量序列的极限……43
2.9.1 几乎必然极限……44
2.9.2 均方极限(依均方收敛)……44
2.9.3 随机极限(依概率收敛)……44
2.9.4 依分布收敛……45
2.9.5 各极限之间的关系……45

3. 马尔可夫过程……46
3.1 随机过程……46
3.2 马尔可夫过程……47
3.2.1 相容性——查普曼-柯尔莫哥洛夫方程……48
3.2.2 离散状态空间……49
3.2.3 更一般的测度……49
3.3 随机过程连续性……49
3.3.1 连续马尔可夫过程的数学定义……51
3.4 微分查普曼-柯尔莫哥洛夫方程……52
3.4.1 微分查普曼-柯尔莫哥洛夫方程的推导……53
3.4.2 微分查普曼-柯尔莫哥洛夫方程的地位……57
3.5 条件和结论的解释……58

3.5.1 跳过程：主方程……58
3.5.2 扩散过程——福克-普朗克方程……59
3.5.3 确定性过程——刘维方程……60
3.5.4 一般过程……61
3.6 初始时刻的时间发展方程——后向方程……62
3.7 平稳齐次马尔可夫过程……64
3.7.1 遍历性质……65
3.7.2 齐次过程……67
3.7.3 趋向于平稳过程……69
3.7.4 马尔可夫过程的自相关函数……73
3.8 马尔可夫过程的例子……75
3.8.1 维纳过程……76
3.8.2 一维随机游动……80
3.8.3 泊松过程……83
3.8.4 奥恩斯坦-乌伦贝克过程……85
3.8.5 随机电报过程……88

4. 伊藤分析和随机微分方程……91
4.1 动机……91
4.2 随机积分……94
4.2.1 随机积分的定义……94
4.2.2 例子 $\int_{t_0}^{t} W(t')\mathrm{d}W(t')$……96
4.2.3 斯特拉托诺维奇积分……97
4.2.4 非可料函数……98
4.2.5 $\mathrm{d}W(t)^2=\mathrm{d}t$ 和 $\mathrm{d}W(t)^{2+N}=0$ 的证明……99
4.2.6 伊藤随机积分的性质……101
4.3 随机微分方程(SDE)……105
4.3.1 伊藤随机微分方程：定义……106
4.3.2 伊藤随机微分方程的解的马尔可夫性质……108
4.3.3 变量代换：伊藤公式……108
4.3.4 福克-普朗克方程和随机微分方程之间的联系……110
4.3.5 多变量系统……111

4.3.6 斯特拉托诺维奇随机微分方程 ……112
4.3.7 对初始条件和参数的依赖性 ……117
4.4 一些例子和解 ……118
4.4.1 不依赖于 x 的系数 ……118
4.4.2 乘积线性白噪声过程 ……119
4.4.3 具有噪声频率的复振子 ……121
4.4.4 奥恩斯坦-乌伦贝克过程 ……122
4.4.5 从直角坐标到极坐标的转换 ……124
4.4.6 多变量奥恩斯坦-乌伦贝克过程 ……126
4.4.7 一般的单变量线性方程 ……129
4.4.8 多变量线性方程 ……131
4.4.9 与时间有关的奥恩斯坦-乌伦贝克过程 ……133

5. 福克-普朗克方程 ……135
5.1 背景 ……135
5.2 一维福克-普朗克方程 ……136
5.2.1 边界条件 ……137
5.2.2 齐次福克-普朗克方程的平稳解 ……143
5.2.3 一些平稳解的例子 ……145
5.2.4 后向福克-普朗克方程的边界条件 ……147
5.2.5 本征函数方法(齐次过程) ……149
5.2.6 一些例子 ……151
5.2.7 齐次过程的首次通过时间 ……156
5.2.8 经由区间的一特定端点出口的概率 ……162
5.3 多维福克-普朗克方程 ……164
5.3.1 变量的代换 ……164
5.3.2 边界条件 ……167
5.3.3 平稳解：位势条件 ……168
5.3.4 细致平衡 ……169
5.3.5 细致平衡的推论 ……172
5.3.6 福克-普朗克方程中的若干细致平衡例子 ……178
5.3.7 多变量齐次过程情况下的本征函数方法 ……188
5.4 区域的首次出口时间(齐次过程) ……194

5.4.1 平均出口时间问题的解 ……195
5.4.2 出口点的分布 ……198

6. 扩散过程的近似方法……201
6.1 小噪声摄动理论……201
6.2 随机微分方程的小噪声展开……204
6.2.1 展开的有效性 ……207
6.2.2 (齐次过程的)平稳解 ……208
6.2.3 均值、方差与时间相关函数 ……209
6.2.4 小噪声摄动理论的失效 ……210
6.3 福克-普朗克方程的小噪声展开 ……213
6.3.1 矩方程及自相关函数 ……215
6.3.2 例子 ……218
6.3.3 平稳分布的渐近方法 ……220
6.4 快变变量的绝热消去……221
6.4.1 利用算子及投影算子的抽象表述 ……224
6.4.2 使用拉普拉斯变换的解法 ……227
6.4.3 短期行为 ……229
6.4.4 边界条件 ……231
6.4.5 系统的摄动分析 ……233
6.5 白噪声过程作为非白噪声过程的极限……237
6.5.1 结果的一般性 ……242
6.5.2 更一般的涨落方程 ……242
6.5.3 非齐次系统 ……243
6.5.4 算子 L_1 与时间相依的后果……244
6.6 快变变量的绝热消去：一般情况……245
6.6.1 例子：短寿命中间化合物的消去 ……245
6.6.2 哈肯模型中的绝热消去法 ……251
6.6.3 快变变量的绝热消去：非线性情况 ……255
6.6.4 任意非线性耦合的例子 ……260

7. 主方程与跳过程……263
7.1 生灭主方程——单变量情况……264

7.1.1 平稳解 ……264
7.1.2 例子：化学反应 $X \rightleftarrows A$ ……266
7.1.3 一个化学双稳系统 ……270
7.2 用福克-普朗克方程逼近主方程 ……274
7.2.1 扩散过程的跳过程近似 ……274
7.2.2 克雷默斯-莫耶尔展开 ……278
7.2.3 范坎彭的系统大小展开 ……279
7.2.4 库尔茨定理 ……284
7.2.5 临界涨落 ……285
7.3 生灭过程的边界条件 ……287
7.4 平均首次通过时间 ……289
7.4.1 吸收概率 ……291
7.4.2 与福克-普朗克方程的比较 ……291
7.5 多变量生灭系统 ……292
7.5.1 细致平衡成立时的平稳解 ……294
7.5.2 没有细致平衡时的平稳解 ……296
7.5.3 系统大小展开与有关的展开 ……296
7.6 某些例子 ……298
7.6.1 $X + A \rightleftarrows 2X$ ……298
7.6.2 $X \underset{k}{\overset{\gamma}{\rightleftarrows}} Y \overset{k}{\underset{\gamma}{\leftrightarrows}} A$(*) ……298
7.6.3 被捕食-捕食系统 ……299
7.6.4 母函数方程 ……305
7.7 泊松表示 ……309
7.7.1 泊松表示的种类 ……314
7.7.2 实泊松表示 ……314
7.7.3 复泊松表示 ……315
7.7.4 正泊松表示 ……319
7.7.5 时间相关函数 ……323
7.7.6 三分子反应 ……330
7.7.7 三阶噪声 ……335

(*) 原文目录中把反应的方向印反了。——译注

8. 空间分布系统……340
8.1 背景……340
8.1.1 泛函福克-普朗克方程……343
8.2 多变量主方程的描述……344
8.2.1 扩散……344
8.2.2 扩散主方程的连续形式……346
8.2.3 反应与扩散的结合……351
8.2.4 泊松表示法……353
8.3 空间和时间相关结构……353
8.3.1 反应 $X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} Y$……354
8.3.2 反应 $B + X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C$, $A + X \xrightarrow[k_3]{} 2X$……358
8.3.3 两级相变的非线性模型……364
8.4 局部和整体描述之间的关系……369
8.4.1 非均匀模式的明显绝热消去法……369
8.5 相空间的主方程……372
8.5.1 流动的处理方法……372
8.5.2 作为生灭过程的流……373
8.5.3 包含碰撞在内的情况——玻耳兹曼主方程……377
8.5.4 碰撞和流同时存在……386

9. 双稳定性、亚稳定性和逃逸问题……385
9.1 在双井势中的扩散(单变量)……385
9.1.1 $D=0$ 时的行为……386
9.1.2 D 非常小时的行为……387
9.1.3 出口时间……388
9.1.4 分裂概率……389
9.1.5 来自非稳态的衰变……391
9.2 在每个势井中的粒子数平衡……392
9.2.1 克雷默斯方法……393
9.2.2 例子：糜蛋白酶的可逆质变……397
9.2.3 生灭主方程(单变量)的双稳定性……399

9.3 多变量系统中的双稳定性……402
9.3.1 出口点的分布……403
9.3.2 平均出口时间的渐近分析……408
9.3.3 多维空间中的克雷默斯方法……409
9.3.4 例子：双势井中的布朗运动……412

10. 量子力学马尔可夫过程……419
10.1 谐振子的量子力学……420
10.1.1 与外场的相互作用……421
10.1.2 相干态的性质……422
10.2 密度矩阵和概率……426
10.2.1 冯·诺伊曼方程……428
10.2.2 格劳伯-苏达香 P 表示……429
10.2.3 算子对应……430
10.2.4 对受迫谐振子的应用……431
10.2.5 量子特征函数……433
10.3 量子马尔可夫过程……435
10.3.1 热槽……435
10.3.2 槽算子平滑函数的相关……437
10.3.3 与热槽发生相互作用的系统的量子主方程……438
10.4 量子马尔可夫过程的例子及应用……442
10.4.1 谐振子……442
10.4.2 被驱动的双能级原子……447
10.5 量子马尔可夫过程中的时间相关函数……452
10.5.1 量子回归定理……454
10.5.2 P 表示对谐振子的应用……455
10.5.3 双能级原子的时间相关……458
10.6 广义的 P 表示……459
10.6.1 广义 P 表示的定义……460
10.6.2 存在性定理……462
10.6.3 关于泊松表示……464
10.6.4 算子恒等式……465

10.7 广义 P 表示对时间演化方程的应用 ……466
10.7.1 复 P 表示……467
10.7.2 正 P 表示……468
10.7.3 例子……470

参考文献……472
参考书目……478
符号索引……482
作者索引……486
论题索引……489

1. 历史介绍

1.1 动机

直至十九世纪末，理论科学一直可被视为微分方程求解的研究以及用这些微分方程的确定性的解作为自然现象的模型。当时的普遍想法是只要收集到了所有的初始数据，人们就能确凿地预测未来的状态。

现在我们知道，至少在两种情况下事实并非如此。首先，量子力学的出现在四分之一世纪内导致了一个新的物理学，因而产生了整个科学的新的理论基础，作为主要的基础，它具有一个纯粹的统计原理；其次，最近已经出现了混沌的概念，其中甚至十分简单的微分方程系统亦颇为令人惊讶地具有那种导致本质上不可预测的性质。无疑，给出了这种系统的初始条件，它的未来值是可以预测的，但初始条件中的任何误差被迅速放大以至有实际价值的可预测性荡然无存。事实上，混沌的存在是不足为怪的，因为它比起纯粹的可预测性来更为符合我们的日常经验。也许令人奇怪的是这一论点花了如此长的时间才得以确立。

混沌和量子力学并非本章的主题。这里我希望给出一个"半历史的"略述，说明涨落唯象理论的兴起及其精髓之所在。可预测模型的实用性恰恰指明了现实生活不都是混沌的，但可预测性有其限度，本书中我们所关心的便是有限度的可预测性的模型。仔细测量的科学实验常常给予我们象图 1.1 那样的数据，它表示由 $X \rightleftharpoons A$ 形式的化学反应所产生的物质 X 的分子个数增长状况。图中明白地显示了确定性的增长动向，它是可再现的，围绕这一动向的涨落却不一样，是不可再现的。

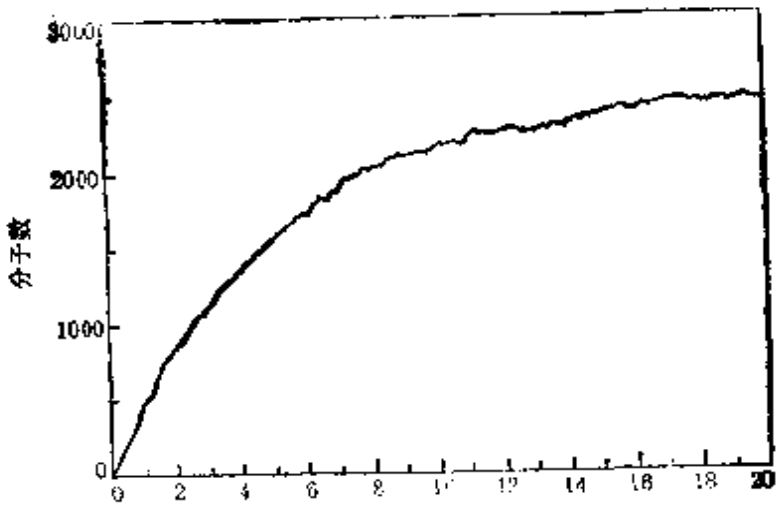


图 1.1　一个异构化反应 $X \rightleftarrows A$ 的随机模拟

1.2　一些历史实例

1.2.1　布朗运动

罗伯特·布朗(Robert Brown)在 1827 年第一个研究了这样的事实：人们发现悬浮在水中的花粉微粒处于一种非常活跃的、不规则的运动状态中；因为他的十分重要的先驱工作，所观测到的现象被命名为布朗运动. 布朗是一位植物学家——确是一位很著名的植物学家——当然他检验了这种运动是否是某种方式的生命现象. 他证实了这一现象对任何悬浮粒子总是存在的，如玻璃微粒、矿物微粒甚至埃及狮身人面像碎片的微粒，据此他排除了这种运动的任何特殊生物有机体的起因. 图 1.2 中显示了布朗运动.

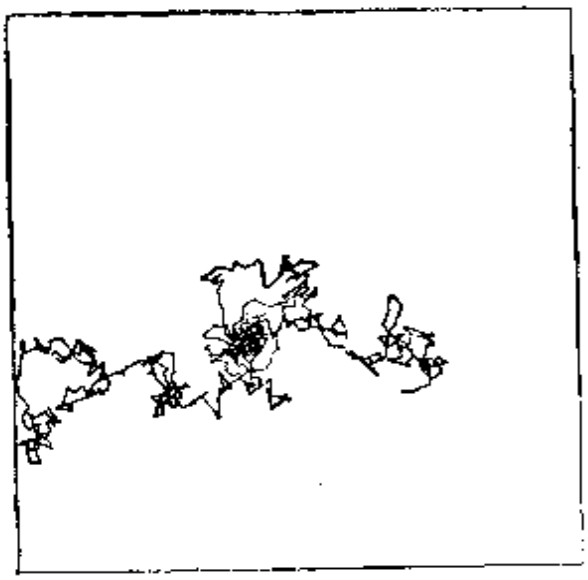


图 1.2　一个小点的布朗运动

布朗运动之谜并未被迅速解开，直至 1905 年才有了令人满意的解释，当时爱因斯坦(Einstein)在相当谨慎的标题

下发表了一个解释："热的分子运动论所要求的静止液体中悬浮粒子的运动"[1.2]。斯莫路乔夫斯基(Smoluchowski)独立地发现了相同的解释[1.3]，后来，他对布朗运动理论的许多系统的发展和大量的实验论证负起了重任。

爱因斯坦对布朗运动理论的解释有两个要点：

(i) 布朗运动是由悬浮液体中的持续不断运动着的分子极其频繁地撞击花粉微粒所引起的。

(ii) 这些分子的运动是如此之复杂，以至其对花粉微粒的影响只能依据极其频繁地统计独立的撞击来作出概率上的描述。

与此相似的涨落的存在引出了一个对这类现象的统计解释。统计学虽然早已由马克斯韦尔(Maxwell)和玻耳兹曼(Boltzman)用于他们的著名的气体理论，但这仅作为各个可能的状态以及这些状态完成的可能性的描述，而不是作为系统时间演变的一个内在的部分。其实是瑞利(Rayleigh)首先考虑了这些现象的统计解释，但因某种原因很少在他的著作中出现。实际上必须把爱因斯坦对布朗运动本性的解释视为自然现象随机模型的开端。

爱因斯坦的推理是非常清晰而精致的，它包含了所有将构成本书主题的基本概念。与其对经典篇章作释义，我倒不如直接给出爱因斯坦的论文的大段摘录(作者自己的英译)：

"显然应该假设每个单个粒子进行的运动与所有其他粒子的运动无关；也应认为同一粒子在不同的时间间隔中的运动是独立的过程，只要时间间隔取得不是太小。"

"我们考虑一个时间间隔 τ，虽然它与可观测的时间间隔相比非常之小，但仍然大到足以使粒子在两个相继的时间间隔内所进行的运动可被认为是相互独立的事件。"

"现在假设 n 个粒子全体悬浮在液体中，在时间间隔 τ 内，单个粒子的 X-坐标将增加一个量 Δ，对各个粒子 Δ 取不同的(正的或负的)值。Δ 将遵循某个频率法则：位移在 Δ 到 $\Delta+d\Delta$ 之间的粒子数 dn 可表示为方程

$$dn = n\phi(\Delta)d\Delta \qquad (1.2.1)$$

这里

$$\int_{-\infty}^{\infty}\phi(\Delta)\mathrm{d}\Delta=1, \tag{1 2.2}$$

且 ϕ 仅对很小的 Δ 值异于零，又满足条件

$$\phi(\Delta)=\phi(-\Delta)。" \tag{1.2.3}$$

"现在我们研究扩散系数是如何依赖于 ϕ 的。我们将再次限于单位体积内粒子数 ν 仅与 x 和 t 有关的场合。"

"令 $\nu=f(x,t)$ 是单位体积内的粒子数，我们从时刻 t 的粒子分布来计算时刻 $t+\tau$ 的粒子分布。由函数 $\phi(\Delta)$ 的定义，不难求出在时刻 $t+\tau$ 位于通过点 x 和点 $x+\mathrm{d}x$ 的，垂直于 x 轴的两个平面之间的粒子数为

$$f(x,t+\tau)\mathrm{d}x=\mathrm{d}x\int_{-\infty}^{\infty}f(x+\Delta,t)\phi(\Delta)\mathrm{d}\Delta \tag{1.2.4}$$

因 τ 很小，可以置

$$f(x,t+\tau)\approx f(x,t)+\tau\frac{\partial f}{\partial t}。 \tag{1.2.5}$$

进而，将 $f(x+\Delta,t)$ 按 Δ 的幂展开，

$$f(x+\Delta,t)=f(x,t)+\Delta\frac{\partial f(x,t)}{\partial x}+\frac{\Delta^2}{2!}\frac{\partial^2 f(x,t)}{\partial x^2}+\cdots \tag{1.2.6}$$

在积分号下应用此级数，因式中的 Δ 取值微小，可得

$$f+\frac{\partial f}{\partial t}\tau\approx f\int_{-\infty}^{\infty}\phi(\Delta)\mathrm{d}\Delta+\frac{\partial f}{\partial x}\int_{-\infty}^{\infty}\Delta\phi(\Delta)\mathrm{d}\Delta+\frac{\partial^2 f}{\partial x^2}\int_{-\infty}^{\infty}\frac{\Delta^2}{2}\phi(\Delta)\mathrm{d}\Delta+\cdots\text{(*)} \tag{1.2.7}$$

因为 $\phi(x)=\phi(-x)$，故等式右方第 2、第 4 等项为零，同时第 1、第 3、第 5 等项中每一项同前一项相比都很小。由这个等式，考虑到

(*) 原文中 $\frac{\partial f}{\partial t}$ 误为 $\frac{\partial f}{\partial \tau}$，又遗漏"+…"部分，现已改正，此类明显的错误后文中也已改正，但不一定作注。——译注

$$\int_{-\infty}^{\infty}\phi(\Delta)\mathrm{d}\Delta=1 \tag{1.2.8}$$

并令

$$\frac{1}{\tau}\int_{-\infty}^{\infty}\frac{\Delta^2}{2}\phi(\Delta)\mathrm{d}\Delta=D \tag{1.2.9}$$

且只保留第1项和第3项，可得

$$\frac{\partial f}{\partial t}=D\frac{\partial^2 f}{\partial x^2} \tag{1.2.10}$$

……这是早已知道的扩散微分方程，可以看出D是扩散系数。……”

“该问题对应于单个点的扩散问题(忽略在扩散粒子之间的相互作用)，它在数学上已经完满地定解，其解为

$$f(x,t)=\frac{n}{\sqrt{4\pi D}}\frac{\mathrm{e}^{-x^2/4Dt}}{\sqrt{t}} \tag{1.2.11}$$

……”

“借助于这个方程我们来计算粒子沿x轴方向的平均位移，或者更确切地说，沿x轴方向位移平方的算术平均的平方根；它是

$$\lambda_x=\sqrt{x^2}=\sqrt{2Dt}\ .\ ” \tag{1.2.12}$$

爱因斯坦的推导实际上基于离散时间假设，撞击仅在时刻0、τ、2τ、3τ、…发生。他所得到的分布函数$f(x,t)$的方程(1.2.10)及其解(1.2.11)应视为一种近似，其中τ是小到足以把t认为是连续的。然而他的叙述还是包含了此后所发展的越来越一般而又严格的大多数主要概念，这些概念正是本书的核心内容，例如：

i) 查普曼-柯尔莫哥洛夫方程以类似于爱因斯坦方程(1.2.4)的形式出现。该方程描述粒子在时刻$t+\tau$位于点x处的概率，它由所有可能的从位置$x+\Delta$“推移”Δ的概率乘上在时刻t位于$x+\Delta$处的概率后求和而得。这一假定基于推移Δ与所有先前的运动经历的独立性：必须知道的仅是粒子在时刻t——而不是任何先前的时刻——的初始位置。这就是马尔可夫公设和作为所有马尔可夫过程的核心动力学方程的查普曼-柯尔莫哥洛夫方程，(1.2.4)是它的特殊形式。这些内容将在第3章中细致讨

论。

ii) 福克-普朗克方程。扩散方程(1.2.10)是福克-普朗克方程的特殊情况，福克-普朗克方程刻划了一大类非常令人感兴趣的随机过程，在这种过程中系统具有连续样本路径。在扩散方程的场合有着这样的含义：如果把花粉微粒的位置认为是服从由连续时间的(不是爱因斯坦所假设的离散时间的)扩散方程(1.2.10)解得的概率法则的话，该位置可写成为一个时间的连续函数，不过它是一个随机函数。这一事实导致我们用某种更直接的概率方法去考虑描述系统动力学的概率，由此才得出了路径的随机微分方程。这种程序是由郎之万(Langevin)所创始的，他所导出的著名的方程现今已冠以他的名字。

iii) 克雷默斯-莫耶尔以及相仿的展开实质上与爱因斯坦所用的方法相同，即前文中从查普曼-柯尔莫哥洛夫方程(1.2.4)导出扩散方程(1.2.10)的方法。采用这种类型的逼近可将一个样本路径未必是连续的过程用一个路径是连续的过程取代，这是过去十年间的热门话题。第7章中将讨论其应用及效能。

1.2.2 郎之万方程

在爱因斯坦最早的推导之后不久，郎之万[1.4]提出了一个与爱因斯坦的方法完全不同的新方法，按他自己所说，是"无比简单"的方法。他的推理过程如下。

根据统计力学，处在平衡状态下的布朗粒子的平均动能具有值

$$\left\langle \frac{1}{2} mv^2 \right\rangle = \frac{1}{2} kT \tag{1.2.13}$$

式中T为绝对温度，k为玻耳兹曼常数。爱因斯坦和斯莫路乔夫斯基两人都用过这一论据。作用于质量为m的粒子的力应该有这样两个：

i) 一个是粘滞阻力，假设这一阻力由宏观流体力学的相同公式给出，即是$-6\pi\eta a\mathrm{d}x/\mathrm{d}t$，其中$\eta$是粘滞度，$a$是设想粒子为球

形时的半径。

ii) 另一个是涨落力 X，表示液体分子对布朗粒子的持续撞击。对此我们全部所知道的只是它是真实的力，以及它应该以相等的概率取正值和负值。这样一来，按牛顿定律可以给出粒子位置移动的方程为

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2}=-6\pi\eta a\frac{\mathrm{d}x}{\mathrm{d}t}+X \tag{1.2.14}$$

乘上 x，可改写成

$$\frac{m}{2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}(x^2)-mv^2=-3\pi\eta a\frac{\mathrm{d}(x^2)}{\mathrm{d}t}+Xx \tag{1.2.15}$$

这里 $v=\mathrm{d}x/\mathrm{d}t$. 再就巨大数量的不同粒子取平均，并应用(1.2.13)得到关于$\langle x^2\rangle$的方程为

$$\frac{m}{2}\frac{\mathrm{d}^2\langle x^2\rangle}{\mathrm{d}t^2}+3\pi\eta a\frac{\mathrm{d}\langle x^2\rangle}{\mathrm{d}t}=kT \tag{1.2.16}$$

其中$\langle xX\rangle$项已被取为零，原因是(按郎之万的说法)“量X的不规则性”。进而可求得通解为

$$\frac{\mathrm{d}\langle x^2\rangle}{\mathrm{d}t}=kT/(3\pi\eta a)+C\exp(-6\pi\eta at/m) \tag{1.2.17}$$

式中C是任意常数。郎之万估计衰减指数项以10^{-8}秒数量级的时间常数趋向于零，这对于当时的实际观测来讲事实上是即刻消失的。这样，我们可以略去这一项，并再次积分得到

$$\langle x^2\rangle-\langle x_0^2\rangle=[kT/(3\pi\eta a)]t \tag{1.2.18}$$

此式对应于爱因斯坦所推出的(1.2.12)式，只要认定

$$D=kT/(6\pi\eta a) \tag{1.2.19}$$

即可。爱因斯坦在上述同一篇论文中用他自己的方法也导出了这个结果。

郎之万方程是随机微分方程的第一个实例。这是一个带有随机项X的微分方程，因此在某种意义上它的解是一个随机函数。郎之万方程的各个解表示不同的随机轨道，只要应用X(涨落力)的相当简单的性质，就可以导出其可测量的结果。

在此有一个问题：爱因斯坦明确要求(在一个充分大的时间尺

度上）增量 Δ 与 Δ 的先前的值完全独立，郎之万虽然并未明显地提及这个观念，但在令 $\langle Xx\rangle$ 等于零时，它却隐约地出现了。X 是极其不规则的，又是与作为不规则时间函数的 x 相互独立的（这一点郎之万没有明说）；这两点实际上等价于爱因斯坦的独立性假设，有了它们，就不会莫明其妙地凑巧使 x 和 X 老是处于同样的方向，致使两者乘积的平均也许不能取为零。郎之万的方法显然是直捷了当的，至少初看起来是这样，该方法给出了一个将动力学方程推广到概率性方程的非常自然的手段。然而对郎之万的方法给以适当的数学基础却在四十年后伊藤提出了系统的随机微分方程的概念才得以实现。在这一套理论中，X 和 x 的独立性的精确陈述导致了现在取名为伊藤的随机微分运算，这些内容将在第 4 章中充分地展开。

作为一个物理学论题，在本世纪最初二十年是布朗运动的全盛时期。许多学者，特别是斯莫路乔夫斯基进行了广泛的理论和实验研究，这些研究显现了同斯莫路乔夫斯基自己和爱因斯坦（见[1.5]）所创立的原始论述的完全一致性。近年来，随着激光散射光谱学的发展，布朗运动的数量上的可测性大有进展。这种技术用强烈的相干激光照射少量包含布朗粒子的液体，然后研究直接联系于布朗粒子运动的散射光的涨落。借助于这一手段，使得观测比传统的花粉远为细小的粒子的布朗运动成为可能，并导出了关于病毒和高分子的大小的数据。制备了更浓缩的悬浮，粒子间的相互作用看来引出了关于高分子悬浮和碰撞的有趣而相当复杂的问题[1.6]。

这些方程描述的涨落力的一般概念已经有了非常广泛的发展。连续性的描述的好处是显著的，因为仅需极少几个参量，实际上就是(1.2.7)式中导数的系数：

$$\int_{-\infty}^{\infty}\Delta\phi(\Delta)\mathrm{d}\Delta \text{ 与 } \int_{-\infty}^{\infty}\Delta^2\phi(\Delta)\mathrm{d}\Delta \tag{1.2.20}$$

难得找到一个问题不能用这种系统来说明，至少在某种程度的近似意义下。对于问题的简明定性分析而言，在正常状况下考虑一

个适当形式的福克-普朗克方程已颇为足够了，其中(1.2.20)的两个系数可以是与 x 有关的，而 x 则是在一个适当维数的空间中。

1.3 生灭过程

多种现象可以用一类特殊的过程——生灭过程作为其模型。这个名称显然出自人类或动物的群体模型，其中的个体有生有死。最有趣的模型之一是由两种动物构成的被捕食-捕食系统，其中的一种捕食另一种，后者自己则有取之不尽的食物供应。令 X 表示被捕食者，Y 表示捕食者，A 是被捕食者的食物，则考虑的过程可以是

$$X + A \to 2X \tag{1.3.1a}$$

$$X + Y \to 2Y \tag{1.3.1b}$$

$$Y \to B \tag{1.3.1c}$$

其中有着朴素而又有趣的解释：第一个方程用符号表示被捕食者吃下一个单位的食物，就立即繁殖一个被捕食者；第二个方程表示一个捕食者吞食掉一个被捕食者，也就立即繁殖，而被吞食导致的死亡则是被捕食者的仅有的死亡机制；最后一个方程表示捕食者的自然死亡。不难猜测 X、Y 的数量 x、y 的微分方程模型的形式。可以设想第一个符号方程表示 X 的繁殖率与 x 跟食物总量的乘积成正比；第二个方程表示 Y 的繁殖率(等于 X 的消耗率)正比于 xy；最后的方程则表示 Y 的死亡率与 y 成正比。这样我们可以写出

$$\frac{dx}{dt} = k_1 ax - k_2 xy \tag{1.3.2a}$$

$$\frac{dy}{dt} = k_2 xy - k_3 y \tag{1.3.2b}$$

洛特卡(Lotka)[1.7]和伏尔特拉(Volterra) [1.8]各自独立地求出了这方程的解，它有着很有趣的振荡解，如图 1.3a 所示。这种振荡易于定性地解释。在捕食者数量稀缺时，被捕食者群体迅速增长直至有充裕的被捕食者供捕食而刺激捕食者迅速繁殖，与此同时

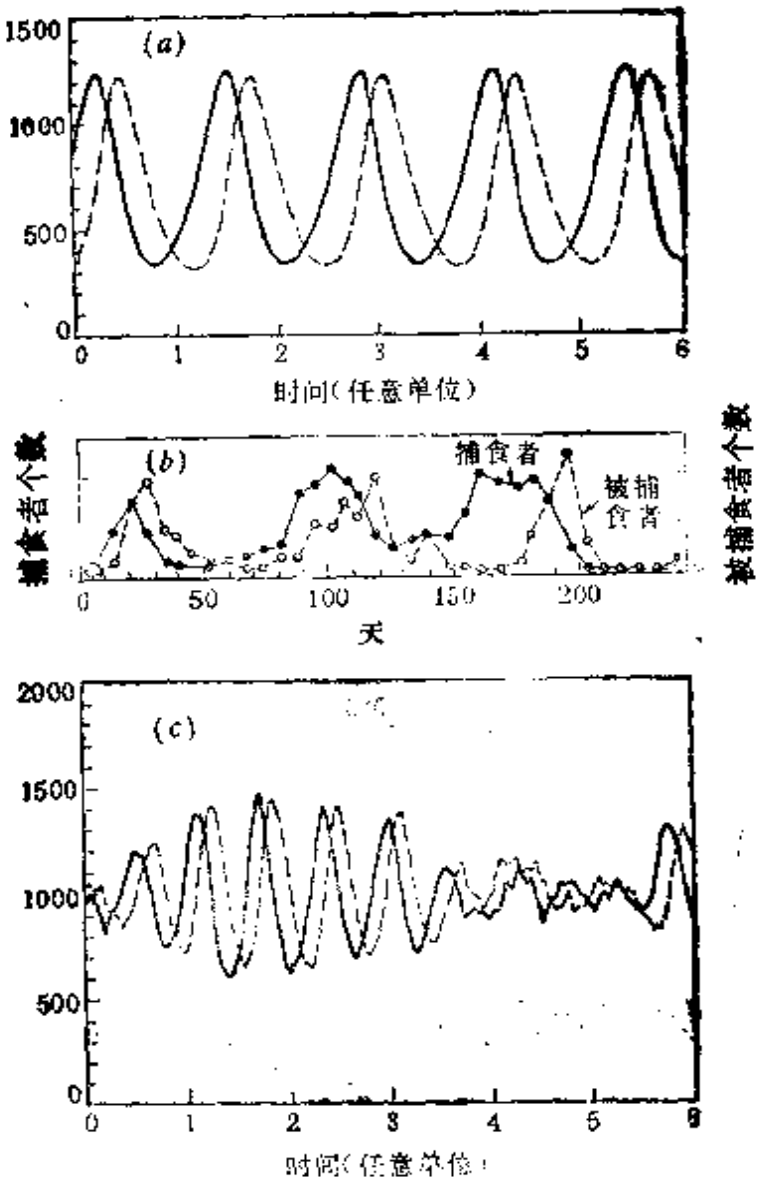


图 1.3 被捕食者-捕食者系统的时间展开

(a) 确定性方程(1.3.2)的解的图形(x=实线，y=虚线). (b) 一个真实的被捕食者-捕食者系统的数据，这里被捕食者是一种以柑桔为食的螨(六点始叶螨——实线)，捕食者是另一种螨(西方盲走螨). 数据来自[1.16, 17]. (c) 随机方程(1.3.3)的模拟

被捕食者数量有所缩减。因为大量的被捕食者被吃掉而不再足以维持捕食者群体，随着它们的死亡而回到我们所考虑的初始状况。这种周期无限地重复，确实是许多真实的被捕食者-捕食者系统的特性，至少在定性意义上是这样。图 1.3b 给出了一个例子(*)。

(*) 原图说明中将被捕食者与捕食者搞颠倒了，现已更正。——译注

自然，真实的系统并不确切地遵循微分方程的解，它们在这种曲线附近涨落。人们必须把这些涨落包括在内，最简单的方法是借助于生灭主方程。假设在给定时刻 t 的个体数量的概率分布为 $P(x, y, t)$ 并寻求对应于 (1.3.2) 的概率法则。这只要在无穷小时间 Δt 内下列转移概率法则成立就行了：

$$\mathrm{Prob}(x \to x+1; y \to y) = k_1 ax\Delta t \tag{1.3.3a}$$

$$\mathrm{Prob}(x \to x-1; y \to y+1) = k_2 xy\Delta t \tag{1.3.3b}$$

$$\mathrm{Prob}(x \to x; y \to y-1) = k_3 y\Delta t \tag{1.3.3c}$$

$$\mathrm{Prob}(x \to x; y \to y) = 1-(k_1 ax + k_2 xy + k_3 y)\Delta t \tag{1.3.3d}$$

这样，我们用概率法则取代了简易的法则。然后象爱因斯坦等人在查普曼-柯尔莫哥洛夫方程中所用的方法那样将这些量用于同样的方程，即将 $t+\Delta t$ 时的概率写成若干项之和，其中每一项表示一个先前的状态乘上转移到状态 (x, y) 的概率。所以我们得到

$$\begin{aligned}&\frac{P(x, y, t+\Delta t)-P(x, y, t)}{\Delta t}\\ &= k_1 a(x-1)P(x-1, y, t) + k_2(x+1)(y-1)\\ &\quad \times P(x+1, y-1, t) + k_3(y+1)P(x, y+1, t)\\ &\quad -(k_1 ax + k_2 xy + k_3 y)P(x, y, t)\end{aligned} \tag{1.3.4}$$

再令 $\Delta t \to 0$，上式等于 $\partial P(x, y, t)/\partial t$。在写下设想的概率法则 (1.3.3) 时，我们假定各个事件的出现概率由 x 和 y 的值就可确定，这又是 1.2.1 节中提及的马尔可夫公设。在布朗运动的场合，可作出一些非常令人信服的论点支持这个马尔可夫假设。这里可没那么明确，遗传的概念就是后裔的性状与祖先的性状有关，显然同马尔可夫假设相矛盾。如何把遗传包括进去是另一个问题，世上绝无“万应灵丹”。

在此，当同一品种的不同个体相似时，马尔可夫公设这个假定是有效的，但当可觉察的遗传差异确实存在时，它又是站不住脚的。

这类模型有着广泛的应用，事实上可用于任何一个可以归结为由个体的总和组成的系统，例如不同化合物分子的系统，电子、

光子和其他物理粒子的系统，以及生物系统。转移概率的特定选择取决于对有关的生成和灭亡的细节的了解程度。(1.3.3) 中的单纯的乘法法则是最基本的选择，其中略去了几乎所有与过程有关的细节。在某些物理过程中我们能更加细致和精确地推导出这些转移概率。

方程(1.3.4)一般没有简单的解，但有一个主要特征可把它那样的方程与郎之万型的方程区别开来，在郎之万方程中，涨落项直接加到了微分方程上。(1.3.4) 的解同时确定总体上的确定性运动与涨落：该涨落与所含个体的个数的平方根有相同的数量级。不难对过程作出时间采样展开的模拟，如图 1.3c 所示，这个图确实显示了正确的一般特征。但模型显然已大为简化以致不可能预期会有精确的相符。于是，同布朗运动的情况相比，这里对唯象理论的论述远不如一类数学模型那么详尽，这种数学模型充分地作了简化，所以在非常广泛的领域内是近似有效的。在第7 章中我们将会看到，可以发展一种理论来阐明很多这类模型，还会看到这种理论和随机微分方程理论确有着紧密的联系。

1.4 电子系统中的噪声

早期的无线电传输功率很低，接收机粗糙，但已令每个耳朵明显地听到了大量统称为“噪声”的极其不规则的电子信号，这种信号遍布于大气中，接收机中，或者无线电发射机中，听起来就象无线电一样。两个主要的噪声源是散粒噪声和约翰逊噪声。

1.4.1 散粒噪声

在一个真空管中(以及在固态器件中)所得到的电流是不稳定的，因为它由许多单个电子产生，它们被加速穿过一段距离后将其电荷一个一个地积淀于阳极上。这一过程所产生的电流可写成

$$I(t)=\sum_{t_k} F(t-t_k)。\tag{1.4.1}$$

这里 $F(t-t_k)$ 表示在时刻 t_k 到达的一个电子对电流的贡献，设想每一电子生成形状相同但有一适当时延的脉冲，如图 1.4 所示。

一旦我们考虑对 t_k 应作怎样的选择，立即就出现了统计方面的问题。最简单的选择是每一电子与先前的电子相独立地到达，就是说，各时刻 t_k 以每一个单位时间内的确定的平均数随机分布于范围 $(-\infty, \infty)$，或者所考虑的别的什么时间范围。



图 1.4 散粒噪声示意图
相同的电脉冲在随机时刻到达

这种噪声的分析是在本世纪二十和三十年代发展起来的，主要由赖斯 (Rice) [1.9] 所总结并完成，但早在 1918 年肖特基 (Schottky) 就首先考虑了这一问题[1.10]。

我们将会发现在散粒噪声和生灭主方程所描述的过程之间有着紧密的联系。如果直至时刻 t 到达的电子数为 n，它是一个由概率 $P(n, t)$ 刻划的统计量，那末电子独立到达的假设显然就是马尔可夫假设。因而假定一个电子在 t 到 $t+\Delta t$ 的时间间隔内会到达的概率与 t 和 n 完全无关，只可能依赖于 Δt。选定一个适当的常数 λ，我们可以写出

$$\text{Prob}(n \to n+1, \text{在时间 } \Delta t \text{ 内}) = \lambda \Delta t \tag{1.4.2}$$

所以

$$P(n, t+\Delta t) = P(n, t)(1-\lambda\Delta t) + P(n-1, t)\lambda\Delta t \tag{1.4.3}$$

取极限 $\Delta t \to 0$

$$\frac{\partial P(n, t)}{\partial t} = \lambda[P(n-1, t) - P(n, t)] \tag{1.4.4}$$

这是一个纯粹的增殖过程。若记

$$G(s, t) = \Sigma s^n P(n, t) \tag{1.4.5}$$

[其中 $G(s, t)$称为是 $P(n, t)$的母函数，这个求出解(1.4.8) 的特殊方法用途甚广]，我们有

$$\frac{\partial G(s, t)}{\partial t} = \lambda(s-1)G(s, t) \tag{1.4.6}$$

因而

$$G(s, t) = \exp[\lambda(s-1)t]G(s, 0) \tag{1.4.7}$$

现要求在时刻 $t=0$ 没有电子到达，显然 $P(0, 0)$ 等于 1 而对所有 $n \geqslant 1$ 有 $P(n, 0)$为零，所以 $G(s, 0) = 1$。将(1.4.7)依 s 幂展开，可得

$$P(n, t) = \exp(-\lambda t)(\lambda t)^n/n! \tag{1.4.8}$$

这就是众所周知的泊松(Poisson) 分布(2.8.3 节)。引入变量 $N(t)$ 表示直至时刻 t 所到达的电子数，它是一个随机量，因此

$$P(n, t) = \mathrm{Prob}\{N(t) = n\} \tag{1.4.9}$$

并且称 $N(t)$为泊松过程变量。由此显然可知，形式地由

$$\mu(t) = \mathrm{d}N(t)/\mathrm{d}t \tag{1.4.10}$$

定义的量 $\mu(t)$除了在 $N(t)$ 增长 1 的地方以外都是零，在这些增长处是一个狄拉克 δ 函数，即

$$\mu(t) = \sum_k \delta(t - t_k) \tag{1.4.11}$$

其中 t_k 是各个电子到达的时刻。我们可以写出

$$I(t) = \int_{-\infty}^{\infty} \mathrm{d}t' F(t-t')\mu(t') \tag{1.4.12}$$

对于 $F(t-t')$的合乎情理的限制是当 $t<t'$ 时 $F(t-t')$为零，且当 $t \to \infty$ 时，它也应成为零。其意义就是在电子到达之前它不产生电流，而它到达后的效果最终要消失。为简单起见，我们采用甚为常见的形式

$$\begin{aligned} F(t) &= q\mathrm{e}^{-\alpha t} \quad (t>0) \\ &= 0 \qquad\quad (t<0) \end{aligned} \tag{1.4.13}$$

于是(1.4.12)可重写为

$$I(t)=\int_{-\infty}^{t}\mathrm{d}t' q\mathrm{e}^{-\alpha(t-t')}\frac{\mathrm{d}N(t')}{\mathrm{d}t'} \tag{1.4.14}$$

我们可以推出一个简单的微分方程，对 $I(t)$ 求微分可得

$$\frac{\mathrm{d}I(t)}{\mathrm{d}t}=\left[q\mathrm{e}^{-\alpha(t-t')}\frac{\mathrm{d}N(t')}{\mathrm{d}t'}\right]_{t'=t}+\int_{-\infty}^{t}\mathrm{d}t'(-\alpha q)\mathrm{e}^{-\alpha(t-t')}\frac{\mathrm{d}N(t')}{\mathrm{d}t'} \tag{1.4.15}$$

所以有

$$\frac{\mathrm{d}I(t)}{\mathrm{d}t}=-\alpha I(t)+q\mu(t) \tag{1.4.16}$$

这是一种与郎之万方程相仿的随机微分方程，但其中的涨落力由 $q\mu(t)$ 给出，$\mu(t)$ 是由(1.4.11)给出的泊松过程的导数。但是 $\mu(t)$ 的均值并不是零，事实上，按泊松分布的均值等于方差的性质，由(1.4.10)可得

$$\langle\mu(t)\mathrm{d}t\rangle=\langle\mathrm{d}N(t)\rangle=\lambda\mathrm{d}t \tag{1.4.17}$$

$$\langle[\mathrm{d}N(t)-\lambda\mathrm{d}t]^2\rangle=\lambda\mathrm{d}t \tag{1.4.18}$$

再将涨落定义为 $\mathrm{d}N(t)$ 及其均值之间的差，我们记

$$\mathrm{d}\eta(t)=\mathrm{d}N(t)-\lambda\mathrm{d}t \tag{1.4.19}$$

则随机微分方程(1.4.16)具有形式

$$\mathrm{d}I(t)=[\lambda q-\alpha I(t)]\mathrm{d}t+q\mathrm{d}\eta(t) \tag{1.4.20}$$

现在怎样去解这个方程呢？在此场合这是一个空谈的问题，因为毕竟解是已知的，但是人们还是希望有一个求解技巧。我们就用郎之万所用的方法试试吧，会得到什么解答呢？这个问题的答案是干脆的，毫无意义。比方说，用通常的微积分运算并假定

$$\langle I(t)\mathrm{d}\eta(t)\rangle=0$$

我们可导出

$$\frac{\mathrm{d}\langle I(t)\rangle}{\mathrm{d}t}=\lambda q-\alpha\langle I(t)\rangle \tag{1.4.21}$$

与

$$\frac{1}{2}\frac{\mathrm{d}\langle I^2(t)\rangle}{\mathrm{d}t}=\lambda q\langle I(t)\rangle-\alpha\langle I^2(t)\rangle \tag{1.4.22}$$

当 $t\to\infty$ 时，有理由预期均值成为常数，求解可得

$$\langle I(\infty)\rangle = \lambda q/\alpha \tag{1.4.23}$$

以及

$$\langle I^2(\infty)\rangle = (\lambda q/\alpha)^2 \tag{1.4.24}$$

第一个解是合理的，它仅仅以适当的方程给出了通过系统的平均电流；但是第二个解意味着均方电流与均值的平方相同，就是说，当 $t\to\infty$ 时没有涨落！这是颇为不合理的。这个困惑问题的解决将会证实下面的事实：与我们至今所介绍过的方法相比，随机微分方程是远为精细而巧妙的。

首先，(1.4.17～20) 中所用的关于微分的记号是被审慎地选择的，在推导(1.4.22)时用了通常的微积分运算，即是先写出

$$\mathrm{d}(I^2) \equiv (I-\mathrm{d}I)^2 - I^2 = 2I\,\mathrm{d}I + (\mathrm{d}I)^2 \tag{1.4.25}$$

再去掉 $(\mathrm{d}I)^2$，因为它是关于 $\mathrm{d}I$ 的两次项。但是现在来看(1.4.18)，它等价于

$$\langle \mathrm{d}\eta(t)^2\rangle = \lambda\,\mathrm{d}t \tag{1.4.26}$$

所以关于 $\mathrm{d}\eta$ 的二次量实际上是 $\mathrm{d}t$ 的一阶量。不难找出其理由，显然

$$\mathrm{d}\eta(t) = \mathrm{d}N(t) - \lambda\,\mathrm{d}t \tag{1.4.27}$$

但是 $N(t)$ 的曲线是一个阶梯函数，它在单个电子到达的时刻是不连续的，而且肯定是不可微的。在通常意义下，这些微积分处理没有一个是允许的。但我们可以用下面的方法使它们有意义。让我们用(1.4.20, 25, 26)直接计算 $\langle \mathrm{d}(I^2)\rangle$：

$$\begin{aligned}\langle \mathrm{d}(I^2)\rangle &= 2\langle I\{[\lambda q - \alpha I]\mathrm{d}t + q\,\mathrm{d}\eta(t)\}\rangle \\ &\quad + \langle\{[\lambda q - \alpha I]\mathrm{d}t + q\,\mathrm{d}\eta(t)\}^2\rangle\end{aligned} \tag{1.4.28}$$

现在我们再次假设 $\langle I(t)\,\mathrm{d}\eta(t)\rangle = 0$，并利用 $\langle \mathrm{d}\eta(t)^2\rangle = \lambda\,\mathrm{d}t$ 的事实取平均，然后展开到关于 $\mathrm{d}t$ 的一阶量，得到

$$\frac{1}{2}\,\mathrm{d}\langle I^2\rangle = \left[\lambda q\langle I\rangle - \alpha\langle I^2\rangle + \frac{q^2\lambda}{2}\right]\mathrm{d}t \tag{1.4.29}$$

此式给出

$$\langle I^2(\infty)\rangle - \langle I(\infty)\rangle^2 = \frac{q^2\lambda}{2\alpha} \tag{1.4.30}$$

因而从这样的观点来考察，当 $t \to \infty$ 时就确有涨落存在。(1.4.29)与(1.4.22)相比较多出了一项，它直接来自 $N(t)$ 是不连续随机函数这一事实所蕴涵的统计考虑。

这样我们就发现了考察郎之万方程的一个更深入的方法，现在由这一观点来看，郎之万方程的论述是极为朴素的。在郎之万的方法中涨落力 X 没有明确说明，但在本书中将会看清楚，就这个主题而言象刚才所考虑的那种问题是非常广泛的。其寓意在于：随机函数通常不能按普通微积分的法则进行微分，必须发展一类特殊的法则，而且微分含义的精确说明是至关重要的。第4章中将就高斯涨落的情况具体说明这一问题及其解答。

1.4.2 自相关函数与谱

对于电路之类的涨落系统，可以进行实际测量的种类并非没有限制。我们已经考虑过分布函数，它告诉我们在任何时刻一个随机量取值的概率分布是什么。如果我们考虑一个随时间涨落的可测量的量 $x(t)$，实践中有时我们能确定 x 的值的分布，然而更常见的情况是只能同时确定均值 $\bar{x}(t)$ 和方差 $\mathrm{var}\{x(t)\}$。

均值和方差并未告诉我们很多关于系统正在发生的基本动力学特性。还有一个量也是重要的，它是在时刻 t 的 x 值对时刻 $t+\tau$ 处的值有多大影响的一个量度，这个量就是由泰勒 (Taylor) 所首先明确引进的自相关函数：

$$G(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}t x(t)x(t+\tau) \qquad (1.4.31)$$

这是对双时刻乘积在任意大的时间 T 内的时间平均，然后让 T 趋向于无穷大。

现在已有特制的自相关器，它们采集样本数据并直接构成所需过程的自相关函数。这种过程包括了从激光散射信号乃至细菌计数的广泛范围。另一个办法是为实验用高速联机计算机编制自相关程序。更有甚者，对十分快变的系统有着限幅自相关器，它定义变量 $c(t)$ 为

$$c(t)=0 \qquad x(t)<l$$
$$=1 \qquad x(t)>l \tag{1.4.32}$$

并计算其自相关函数，用这个办法给出 $x(t)$ 的自相关函数的近似。

一个较为传统的方法是计算 $x(t)$ 的谱，这要分两步定义。首先定义

$$y(\omega)=\int_0^T \mathrm{d}t\ \mathrm{e}^{-i\omega t}x(t) \tag{1.4.33}$$

然后把谱定义为

$$S(\omega)=\lim_{T\to\infty}\frac{1}{2\pi T}|y(\omega)|^2 \tag{1.4.34}$$

自相关函数与谱是紧密相联系的。作少许运算可得

$$S(\omega)=\lim_{T\to\infty}\left[\frac{1}{\pi}\int_0^T \cos(\omega\tau)\mathrm{d}\tau\int_0^{T-\tau}x(t)x(t+\tau)\mathrm{d}t\right] \tag{1.4.35}$$

再求 $T\to\infty$ 时的极限（在适当的假设下可保证某些次序交换的合理性）又得到

$$S(\omega)=\frac{1}{\pi}\int_0^\infty \cos(\omega\tau)G(\tau)\mathrm{d}\tau \tag{1.4.36}$$

这是将自相关函数的傅里叶变换同谱相联系的基本结果。注意到

$$G(-\tau)=\lim_{T\to\infty}\frac{1}{T}\int_{-\tau}^{T-\tau}\mathrm{d}t\,x(t-\tau)x(t)=G(\tau) \tag{1.4.37}$$

(1.4.36) 可以取稍为不同的形式：

$$\boxed{S(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-i\omega\tau}G(\tau)\mathrm{d}\tau} \tag{1.4.38}$$

及其相应的逆变换式

$$\boxed{G(\tau)=\int_{-\infty}^{\infty}\mathrm{e}^{i\omega\tau}S(\omega)\mathrm{d}\omega} \tag{1.4.39}$$

这个结果就是著名的、有广泛应用的维纳－辛钦定理 [1.12, 13]。

它的意义在于人们既可以直接测量一个信号的自相关函数，

也可以测量谱，再作变换，藉助于快速傅里叶变换和计算机这是颇为直截了当的。

1.4.3 涨落函数的傅里叶分析：平稳系统

至今为止自相关函数定义为一个信号的时间平均，但我们也可以考虑系综平均，其中多次重复同样的测量，并计算平均值，用记号〈 〉来表示。后文中将会证实对很多系统而言，时间平均就等于系综平均，这种系统称为是各态历经的(3.7.1 节)。

如果有一个这样的涨落量 $x(t)$，我们就可以认定平均值

$$\langle x(t)x(t+\tau)\rangle = G(\tau) \tag{1.4.40}$$

这个结论来自各态历经的假设。

现在很自然地可把随机量 $x(t)$ 写成一个傅里叶变换

$$x(t) = \int \mathrm{d}\omega c(\omega)\mathrm{e}^{\mathrm{i}\omega t} \tag{1.4.41}$$

并因此有

$$c(\omega) = \frac{1}{2\pi}\int \mathrm{d}t\, x(t)\mathrm{e}^{-\mathrm{i}\omega t} \tag{1.4.42}$$

注意 $x(t)$ 取实值隐含着

$$c(\omega) = c^*(-\omega) \tag{1.4.43}$$

如果系统是各态历经的，则 $\langle x(t)\rangle$ 必定是常数，因为时间平均显然是不变的。这时过程是平稳的，意思是指所有与时间有关的平均值仅仅是时间差的函数，即 $x(t_1)$、$x(t_2)$、…、$x(t_n)$ 的函数的平均值与 $x(t_1+\Delta t)$、$x(t_2+\Delta)$、…、$x(t_n+\Delta)$ 的相等。

为方便起见，后文中我们假定 $\langle x\rangle = 0$，因此

$$\langle c(\omega)\rangle = \frac{1}{2\pi}\int \mathrm{d}t\langle x\rangle\mathrm{e}^{-\mathrm{i}\omega t} = 0 \tag{1.4.44}$$

$$\begin{aligned}\langle C(\omega)C^*(\omega')\rangle &= \frac{1}{(2\pi)^2}\iint \mathrm{d}t\mathrm{d}t'\mathrm{e}^{-\mathrm{i}\omega t+\mathrm{i}\omega' t'}\langle x(t)x(t')\rangle \\ &= \frac{1}{(2\pi)}\ \delta(\omega-\omega')\int \mathrm{d}\tau\mathrm{e}^{\mathrm{i}\omega\tau}G(\tau) \\ &= \delta(\omega-\omega')S(\omega)\end{aligned} \tag{1.4.45}$$

这里我们不仅找到了平方平均值 $\langle |c(\omega)|^2 \rangle$ 与谱之间的关系，也得出结论单是平稳性就蕴涵了 $c(\omega)$ 与 $c^*(\omega')$ 为不相关，因为 $\delta(\omega-\omega')$ 项出自于 $\langle x(t)x(t') \rangle$ 仅是 $t-t'$ 的函数。

1.4.4 约翰逊噪声与尼奎斯定理

1928 年发表了两篇简短而出色的论文，一篇是约翰逊(Johnson)的[1.14]，用实验证明电阻会自动地产生电压涨落，另一篇是尼奎斯(Nyquist)的 [1.15]，给出了同约翰逊的实验完全一致的理论推导，所涉及的原理在此之前已为肖特基所知，与爱因斯坦和郎之万所用的原理相同。该原理就是热平衡原理。如果一个电阻 R 产生电涨落，它们会形成电流并发热。电阻中产生的热量必须精确地与来自涨落的能量平衡。这个原理的详细论述不是本节的主题，但我们将会发现物理学和化学的随机过程中，这样的结论是共同的，统计力学的基础本质上不是随机的，这里引入统计力学的原理是作为随机过程原理的一种补充。已发现的实验结果如下：有一阻值为 R 的电阻器，其绝对温度为 T，假设用适当的滤波器测得电阻两端角频率在 $(\omega, \omega+\mathrm{d}\omega)$ 范围内的电压为 $E(\omega)\mathrm{d}\omega$，则当 k 是玻耳兹曼常数时

$$\langle E^2(\omega) \rangle = 2RkT \tag{1.4.46}$$

这个结果现在称为尼奎斯定理。约翰逊说过："这个效应是真空管放大器中的称为'电子管噪声'的起因之一，它常是一个良好放大器'噪声'的绝大部分"。

约翰逊噪声易于用上一小节的形式来叙述。电阻两端的平均噪声电压为零，且系统被安排成处于稳定状态中，设想可以很好地用平稳过程来表示。实际上，约翰逊的量是(1.4.34)类型的极限，总之可以说电压谱 $S(\omega)$ 由

$$S(\omega) = 2RkT \tag{1.4.47}$$

所给出，就是说谱是平直的，是 ω 的常数函数。在光的场合，频率对应于光的不同颜色。如果看到的是白光，则发现实际上所有的颜色都有相同的份额——所以白光的光谱是平直的——至少在视

觉范围内。相仿地，术语白噪声被用于具有平直谱的噪声电压(或任何其他涨落量)。

白噪声实际上是不可能存在的。最简单的证明是注意电阻中在频率范围(ω_1, ω_2)中的平均耗散功率由

$$\int_{\omega_1}^{\omega_2} \mathrm{d}\omega S(\omega)/R = 2kT(\omega_2 - \omega_1) \tag{1.4.48}$$

所给出，以致在全部频率范围内总的耗散功率为无穷大！尼奎斯领悟到了这一点，并注意到事实上有着量子修正的方法可使它的谱在室温下仅在7×10^{13} Hz 以下为平直的，这个频率在无线电中实际上是测不到的。电阻中的真正耗散功率无论如何得比无穷大要小，事实上是10^{-10} W，实际上还有着另外的限制因素，如系统的电感之类，会使谱限于更低的频率。

由自相关函数(1.4 节)所表示的谱的定义，我们有

$$\langle E(t+\tau)E(t)\rangle = G(\tau) \tag{1.4.49}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}\omega\tau} 2RkT \tag{1.4.50}$$

$$= 2RkT\delta(\tau) \tag{1.4.51}$$

它意味着不管时间差 τ 怎样小，$E(t+\tau)$与$E(\tau)$是不相关的。这当然是谱的平直性的结果。一个几乎是平直的 $S(\omega)$ 的典型模型是

$$S(\omega) = 2RkT/(\omega^2\tau_c^2 + 1) \tag{1.4.52}$$

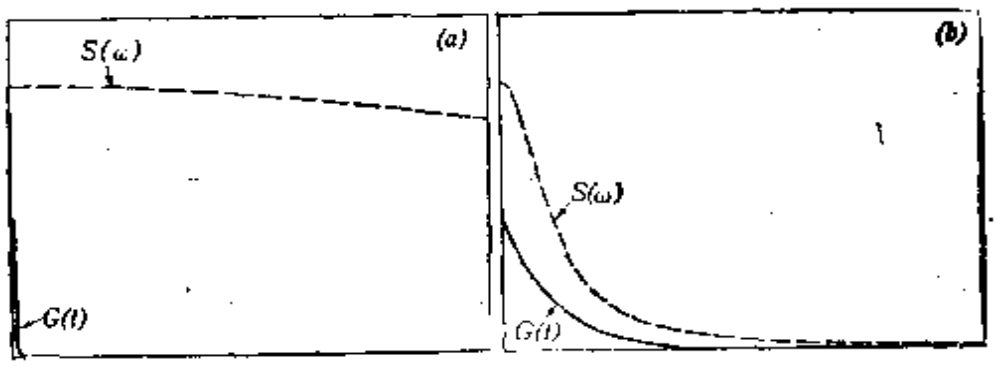


图 1.5　相关函数(实线)与其对应的谱(虚线)

(a) 短的相关时间对应于一个几乎是平直的谱；

(b) 长的相关时间给出一个相当急速地减小的谱

在 $\omega \ll \tau_c^{-1}$ 的条件下它是平直的。在此场合直接算出傅里叶变换可给出

$$\langle E(t+\tau)E(t)\rangle = (RkT/\tau_c)\exp(-\tau/\tau_c) \qquad (1.4.53)$$

所以只要 $\tau \gg \tau_c$ 自相关函数便减小为零，这里 τ_c 称为涨落电压的相关时间。这样一来，δ 函数形式的相关函数看来是理想化的，仅对充分长的时间尺度才站得住脚。

这正使人联想起关于布朗运动的爱因斯坦假设和郎之万涨落力的特性。理想的白噪声在本书中将扮演一个极其重要的角色，但就象随机微分方程中出现的涨落项不同于通常的微分那样，我们将会发现，包含有白噪声作为驱动项的微分方程必须非常小心地处置。任何涨落系统中很自然地产生这种方程，有可能借助于斯特拉托诺维奇法则作出安排使得普通微积分的法则可用，但要以数学定义的不精确和随机操作中有某些难点为代价。人们已经弄清楚，好得多的方法是摒弃普通微积分而使用伊藤分析 (Ito calculus)，它并不是截然不同的(事实上，它同散粒噪声中出现的运算很相仿)，还保持着易于处理的统计特性。所有这些将在第4章中作详尽的讨论。

如前所述，作为一个物理上的现实过程，白噪声是不存在的，它所展示的颇为奇异的特性在任何现实条件下也不会出现。然而，从数学上说它是基础性的，它又是许多真实出现的过程的理想化，在这个物理意义上，它也是基础性的。在关于白噪声的运算中，我们将展现一些稍稍令人觉得奇怪的法则，它们实质上并不很困难，比起处理现实噪声的常用方法容易得多。而且，在白噪声不是一个良好的近似的情况下，常常可以用白噪声相当简单地间接表示出来。在这个意义上，白噪声是一个导出范围广泛的随机描述的一个起点，因而它对本书的主题是重要的基础。

2. 概率概念

在上一章，我们引入了概率概念而未有任何定义。为了把基本概念更精确地用公式表示，必须对这些概念作出某种更精确的表达。本章致力于提供一些背景，并给出若干基本结论。这并不是数学化概率的一个全面的概述，读者可参阅标准数学教科书，如费勒(Feller)[2.1]和帕普里斯(Papoulis)[2.2]的著作。

2.1 事件和事件集

为了方便起见，我们采用尽可能通用的记号来描述那些或许要指定其发生概率的事项。例如，我们可能要讨论在空间的一个确定区域中有着 6.4×10^{14} 个分子的有关情况，或者是关于一个布朗粒子位于空间中一个确定点的情况；也可能是在森林中的某个区域中有着 10 只老鼠和 3 只猫头鹰的情况。

这些情况的发生全是事件的具体实现。抽象地说，一个事件就是某个空间中的一员，在大多数实际场合，它可以用整数向量

$$\boldsymbol{n}=(n_1, n_2, n_3, \cdots) \tag{2.1.1}$$

或实数向量

$$\boldsymbol{x}=(x_1, x_2, x_3, \cdots) \tag{2.1.2}$$

来表示。向量的维数是任意的。

采用集合论的语言是适宜的，引入事件集的概念，用记号

$$\omega\in A \tag{2.1.3}$$

表示 ω 是 A 中所包含的事件之一。例如，可以把集 $A(25)$ 认为是生态群体中存活有不多于 25 只动物的事件集；显然，有 3 只老鼠、1 只老虎而没有其他动物存活的事件 ω 满足

$\omega \in A(25)$ (2.1.4)

更有实际意义地，让我们定义事件集 $A(r, \Delta V)$ 为分子位于以点 r 为中心的体积单元 ΔV 中．在此场合，有关事件集的运算的实际意义便可看清楚了，因为通常我们是能够确定一个分子是否在 r 的邻域 ΔV 中的，但要确定粒子是否确切地位于 r 处是不可能的．这样，如果我们定义事件 $\omega(y)$ 为分子位于点 y，那么询问是否有

$\omega(y) \in A(r, \Delta V)$ (2.1.5)

以及给集 $A(r, \Delta V)$ 指定某个概率都是有意义的，这个概率被视为 (2.1.5) 发生的概率。

2.2 概率

大多数人都有着基于自己经验的概率的直观概念。然而，直观概念的精确的系统表述却充满着困难，人们已经发现最适当的方法是把概率论公理化，成为一个本质上抽象的科学，其中概率测度 $P(A)$ 被指定给事件空间中的每个集 A，包括

全部事件的集：Ω (2.2.1)

没有事件的集：$\emptyset$ (2.2.2)

为了定义概率，我们要求这些事件集在集合论运算并和交之下构成一个封闭系统（数学家通称为 σ-代数）。

2.2.1 概率公理

我们引入 A 的概率 $P(A)$ 作为 A 的一个函数，满足下列概率公理：

(i) $P(A) \geqslant 0$ 对所有 A 成立， (2.2.3)

(ii) $P(\Omega) = 1$， (2.2.4)

(iii) 如果 $A_i (i = 1, 2, 3, \cdots)$ 是一组可数的（但可能是无限的）不相交集，即

$A_i \cap A_j = \emptyset$ （对所有 $i \neq j$ 成立） (2.2.5)

则

$$P(\bigcup_i A_i) = \sum_i P(A_i) \tag{2.2.6}$$

这些就是全部所需的公理。但作为它们的推论，我们还有：

(iv) 如果 $\bar{A}$ 是 A 的补集，即所有不在 A 内的事件的集，则

$$P(\bar{A}) = 1 - P(A), \tag{2.2.7}$$

(v) $$P(\varnothing) = 0 \tag{2.2.8}$$

2.2.2 $P(A)$的意义

为了使概率理论与现实世界相符合，某种程度的直观性是必需的。公理化的概率 $P(A)$ 就是“任意的”事件 ω（即“随机选择”的事件 ω）满足 $\omega \in A$ 的直观概率。或者讲得更明白些，如果我们从 Ω“随机”选择一个事件 N 次，则当次数 N 趋向于无穷大时，所选定的事件满足 $\omega \in A$ 的相对频率接近于 $P(A)$。选择次数 N 既可视为一个接一个作出（一个骰子的多次“独立”投掷），也可视为同时作出（N 个骰子同时“独立地”投掷）。正如我们从这些不确定的用词（“任意”、“随机”、“独立”）的不断明朗化的过程中可以看到的那样，这类说法的定义都必定是直观的。然而柯尔莫哥洛夫 (Kolmogorov) 在他的公理化概率论中排除了我们现在视为直观观念的部分，因而为数学概率论的严格发展扫清了道路。可是企求直观的理解仍然会招致循环定义的困惑。对公理化概率的最简单的看法是把它视为一种用公理处置概率的形式化方法。为了应用这一理论必须确定概率空间并指定概率测度，这无非是设想的先验概率。这种先验概率充斥于多种应用学科。例如，在平衡态统计力学中人们给相空间中相等的体积指定相等的概率；爱因斯坦在关于布朗运动的推理中，对时刻 t 时从位置 x 处“推移”Δ 的概率指定为概率值 $\phi(\Delta)$。

应用概率的任务有两个：(i) 假定某个似乎合理的先验概率的集合，并由此按概率空间的结构推演出结论；(ii) 用某种仪器测量实验结果，这种仪器是为测量对应于那些先验概率的量而制作的。

概率空间的结构是非常重要的，特别是在事件空间复合着附

加的时间概念的场合，这一扩充使有效的概率空间为无限维，因为我们可以构造出"粒子在时间t_n位于点x_n，$n=0、1、2、\cdots、\infty$"这种事件。

2.2.3 公理的意义

概率的任何直观概念都导致概率的非负性，而且任意事件包含在所有事件的集合中的概率必定是1，不管我们怎样去定义"任意"这个词。因此，公理(i)和(ii)是可以理解的。问题的核心在于公理(iii)。假设我们仅处理两个集合A和B，$A\cap B=\emptyset$，这意味着没有一个事件同时包含在A和B之中。所以$\omega\in A\cup B$的概率就是"要末$\omega\in A$，要末$\omega\in B$"的概率。直观的考虑告诉我们这个概率是各概率之和，即

$$P(A\cup B)=P\{(\omega\in A)\text{或}(\omega\in B)\}=P(A)+P(B) \quad (2.2.9)$$

(注意这不是一个证明，只是一个说明)。

现在显然可以推广到任何有限个数的不相交集的情形，但只能再推广到可数个数的不相交集的情形，这须要作些说明。

这个推广必须作出限制是因为存在着用连续足标标记的集，这种足标的一个例子是空间中的位置x。一个分子在仅有一个元素x的集中的概率是零；但它在一个具有有限体积的区域R中的概率则不是零。该区域R是$\{x\}$形集的并，但不是一个可数的并。公理(iii)不能应用，在R中的概率并不等于在$\{x\}$中的概率之和。

2.2.4 随机变量

随机变量的概念是一个记号上的便利，这对于本书是极为重要的。假设有一抽象概率空间，其事件可写为x，然后可以引进随机变量$F(x)$作为x的一个函数，对每一x取一确定值。特别地，x的恒等函数$X(x)$是值得注意的，它由

$$X(x)=x \quad (2.2.10)$$

所给出。本书中我们规定用大写字母表示随机变量，而在必须作出区别的场合用小写字母表示它们的值。

通常，我们有多种多样的、具有值 ω 的基本概率空间 Ω，并讨论 ω 的某个函数 $X(\omega)$，然后略去 ω，这是因为下面两个理由：

(i) 我们用任意的 x 值指定事件，即是将 x 等同于 ω；

(ii) 基本事件实在太复杂以至无法描述，有时甚至无从了解。

例如，对液体中分子位置的情况，实际上我们应把各个 ω 理解为有能力指明液体中各个分子的全部位置、矩和方向，这简直难以表达，通常也是不必要的。

引进随机变量的一大好处是我们可以用它们简单地处理随机变量的函数，例如 X^2、$\sin(\alpha \cdot X)$ 等等，并计算它们的均值与分布。进一步定义了随机微分方程后，也可以相当简洁地讨论随机变量的时间导数，其方法同非概率性系统的微分方程经典方法相当类似。

2.3 联合概率和条件概率：独立性

2.3.1 联合概率

在 2.2.3 节中我们解释了互斥事件的出现是如何同不相交集的概念相联系的。现在我们考虑 $P(A \cap B)$ 是什么意思，这里 $A \cap B$ 为非空。一个满足 $\omega \in A$ 的事件 ω 如果还满足 $\omega \in B$ 才能满足 $\omega \in A \cap B$。这样就有

$$P(A \cap B) = P\{(\omega \in A) \text{与} (\omega \in B)\} \tag{2.3.1}$$

并把 $P(A \cap B)$ 称为联合概率，即事件同时包含在两个集中，或者换一个说法，事件 $\omega \in A$ 和 $\omega \in B$ 两者同时发生。联合概率以两种方式自然地出现于本书内容中：

(i) 当事件是由一向量指定，比如说 m 只老鼠与 n 只老虎，这个事件的概率是“m 只老鼠(和任意个数的老虎)”与“n 只老虎(和任意个数的老鼠)”两者的联合概率。在这种意义上向量都隐含着联合概率。

ii) 在考虑多个时刻时，若问“在时刻 t_1 有 m_1 只老虎和 n_1 只老鼠”与“在时刻 t_2 有 m_2 只老虎和 n_2 只老鼠”的概率是什么，我

们就已经由时刻 t_1 的事件和时刻 t_2 的事件有效地生成了联合事件，它在每个时刻涉及一个事件。上述两种情况之间除了时间的基本动态作用以外实质上并无差别。

2.3.2 条件概率

可以给事件指定我们所关心的条件，并只考虑这种事件，比如“已知有 100 只狮子，这时有 21 只野牛”的概率，这是什么意思呢？显然，我们仅关心那些包含在集合 $B=$〈恰有 100 只狮子出现的所有事件〉中的那些事件，这意味着应将条件概率仅定义在所有包含在 B 中的集之上，我们把条件概率定义为

$$P(A|B)=P(A\cap B)/P(B) \tag{2.3.2}$$

这就同我们的直观观念相符：在已知 $\omega\in B$ 时 $\omega\in A$ 的条件概率是联合出现的概率除以 $\omega\in B$ 的概率。

我们可以按两种方向的考虑来作出定义，即有

$$P(A\cap B)=P(A|B)P(B)=P(B|A)P(A) \tag{2.3.3}$$

在两者之间，比方说〈已知有 100 只狮子时有 21 只野牛〉的概率与〈已知有 21 只野牛时有 100 只狮子〉的概率之间，并无特殊的概念差异。但在涉及两个时刻的时候，我们确实看到了差异。例如，已知粒子在先前的时刻 t_2 位于 x_2 考虑它在时刻 t_1 位于 x_1 的概率是很自然的事；其实，读者会明白这是本书的一个中心概念。反过来则听上去是奇怪的，即是已知粒子在后来的时刻 t_2 位于 x_2 而要考虑一个粒子在时刻 t_1 位于 x_1 的概率，这可有点超自然的味道。我们不能想象有任何自然的方式来考虑它，虽然它在原理上是跟“自然的”，即条件先于所考虑的事件的条件概率极为相似的量。

这个自然而然的定义在本书中早已出现，例如，爱因斯坦的 $\phi(\Delta)\,\mathrm{d}\Delta$(1.2.1 节)是时刻 t 位于 x 的粒子将在时刻 $t+\tau$ 落在范围 $[x+\Delta,\ x+\Delta+\mathrm{d}\Delta]$ 中的概率，另一些例子中也是相似的。我们的直觉告诉我们就象它当年告诉爱因斯坦一样（阅读他的论文摘录就可明白），这种条件概率是直接联系于一个概率系统的时间展开的。

2.3.3 不同次序的联合概率间的关系

假定我们有一组集 B_i，满足

$$B_i \cap B_j = \phi \tag{2.3.4}$$

$$\bigcup_i B_i = \Omega \tag{2.3.5}$$

所以这些集将空间 Ω 分割成互不相交的子集．故有

$$\bigcup_i (A \cap B_i) = A \cap (\bigcup_i B_i) = A \cap \Omega = A \tag{2.3.6}$$

现在应用概率公理(iii)，由于 $A \cap B_i$ 满足那里用 A_i 表示的条件，所以

$$\sum_i P(A \cap B_i) = P[\bigcup_i (A \cup B_i)] \tag{2.3.7}$$

$$= P(A) \tag{2.3.8}$$

这样一来

$$\sum_i P(A|B_i)P(B_i) = P(A) \tag{2.3.9}$$

于是，在所有可能的互斥的 B 上对联合概率求和就消去了那个变元。

因此，一般地有

$$\sum_i P(A_i \cap B_j \cap C_k \cap \cdots) = P(B_j \cap C_k \cap \cdots) \tag{2.3.10}$$

在深为依赖于联合概率的随机过程理论的发展中，结论(2.3.9)有许多意味深长的推论。

2.3.4 独立性

我们需要一个概率方法去规定独立事件的意义。如果指定一个特定事件包含在 B 中却对它属于 A 的概率没有影响，则事件集 A 和 B 应是独立的集合。这时条件概率应独立于 B，故有

$$P(A \cap B) = P(A)P(B) \tag{2.3.11}$$

在多个事件的场合，我们需要一个更强一些的规定。如有事件 $(\omega \in A_i)(i = 1, 2, \cdots, n)$，对任何集 $(1, 2, \cdots, n)$ 的子集 $(i_1, i_2, \cdots, i_k)$ 成立

$$P(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}) \quad (2.3.12)$$

则认为它们是独立的。式(2.3.12)所要求的对所有可能组合的因子分解是很重要的。例如对三个集 A_i，的确可以设想对所有不同的 i、j 成立

$$P(A_i\cap A_j)=P(A_i)P(A_j) \quad (2.3.13)$$

但又有

$$A_1\cap A_2=A_2\cap A_3=A_3\cap A_1 \quad (\text{见图 } 2.1)$$

这就要求

$$\begin{aligned}P(A_1\cap A_2\cap A_3)&=P(A_2\cap A_3\cap A_3)=P(A_2\cap A_3)\\&=P(A_2)P(A_3)=P(A_1)P(A_2)P(A_3)\end{aligned} \quad (2.3.14)$$

我们可以看出，$\omega\in A_2$ 和 $\omega\in A_3$ 的成立必定导致成立 $\omega\in A_1$，在这意义上这些事件显然不会是独立的。

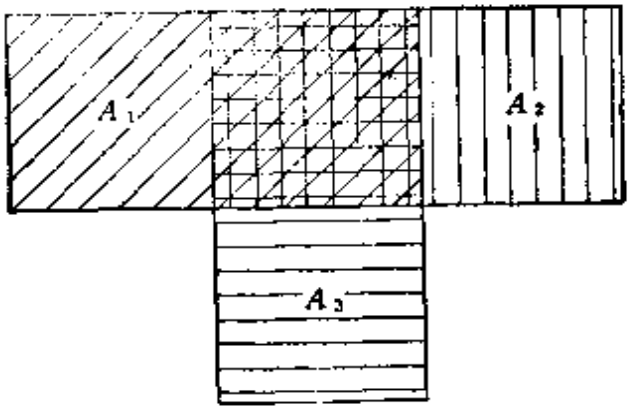


图 2.1 三个集合并不独立，但两两成对统计独立的图示。在三个集合中 $A_i\cap A_j$ 在所有情况下都是中间的区域。适当选择概率可使

$$P(A_i\cap A_j)=P(A_i)P(A_j)$$

至于随机变量 X_1、X_2、X_3、…，如果对 $A_i=(x:a_i\leqslant x\leqslant b_i)$ 形式的所有集，事件 $X_1\in A_1$、$X_2\in A_2$、$X_3\in A_3$、…是独立事件，则这些随机变量称为是独立随机变量。

2.4 均值和概率密度

在基本事件 ω 能够可数地列出时，随机变量 $R(\omega)$ 的均值由

$$\langle R\rangle = \sum_{\omega} P(\omega) R(\omega) \tag{2.4.1}$$

所给出，这里 $P(\omega)$ 是仅含有单个事件 ω 的集合的概率．在连续变量的场合，前述的概率公理使我们能定义一个概率密度函数，如果 $A(\omega_0, \mathrm{d}\omega_0)$ 是集

$$(\omega_0 \leqslant \omega < \omega_0 + \mathrm{d}\omega_0) \tag{2.4.2}$$

则有

$$p(\omega_0)\mathrm{d}\omega_0 = P[A(\omega_0, \mathrm{d}\omega_0)] \tag{2.4.3}$$

$$\equiv p(\omega_0, \mathrm{d}\omega_0) \tag{2.4.4}$$

最后一式的记号是数学家们常用的．有关细节已由费勒 (Feller) [2.1]作了很好的说明．在此情况下

$$\langle R\rangle = \int_{\omega\in\Omega} \mathrm{d}\omega\, R(\omega) p(\omega) \tag{2.4.5}$$

通常可以用 R 自己来表示事件(如 2.2.4 节所指出的那样)，所以我们常常写成

$$\langle R\rangle = \int \mathrm{d}R\, R_p(R) \tag{2.4.6}$$

显然，$p(R)$ 不是象 ω 的函数 $p(\omega)$ 那样是 R 的函数，更精确地有

$$p(R_0)\mathrm{d}R_0 = P[(R_0 < R > R_0 + \mathrm{d}R_0)] \tag{2.4.7}$$

2.4.1 由任意函数的均值确定概率密度

假定对每个函数 $f(R)$ 我们知道

$$\langle f(R)\rangle = \int \mathrm{d}R\, f(R) p(R) \tag{2.4.8}$$

那末我们就知道了 $p(R)$．其证明可由选用

$$f(R) = 1 \qquad R_0 \leqslant R < R_0 + \mathrm{d}R_0$$

$$= 0 \qquad \text{其他}$$

得到．因为任意函数的期望值有时比密度使用起来方便，本书中将偶而使用这一关系．

2.4.2 零概率集

如果有密度 $p(R)$，则 R 在区间 $(R_0, R_0+\mathrm{d}R)$ 中的概率随 $\mathrm{d}R$ 而趋向于零。因此，R 恰有值 R_0 的概率为零；对任何其他值也是相似的。

在这种情况下，就有了许多各自只含有一个点 R_i 的零概率集 $S(R_i)$。由概率公理 (iii)，这种集合的可数的并也有零概率，例如任何仅包含有可数个点(比方说所有有理数)的集合。一般，概率论中所有的相等概念最好也不过是"几乎必然正确"，即是它们可以在一个零概率集上不正确。换一个方式举例说，

$$X = Y \quad (\text{以概率 } 1) \tag{2.4.9}$$

跟

$$X(R) = Y(R) \quad \text{对所有 } R \tag{2.4.10}$$

是截然不同的。如果理论同现实情形有任何联系的话，当然具有零概率的事件并不出现。

特别地，如果仔细核查一下可以注意到：对所有 $f(R)$ 已知 $\langle f(R)\rangle$ 仅仅意味着以概率 1 知道了 $p(R)$。

2.5 均值

在一个概率系统中去测量什么样的值不是一个无足轻重的问题。在实践中，人们或者测量随机变量的一组个别值(某种动物在某时某地的数量，大量同样电路的每一个中通过一个给定电路单元的电流，等等)，或者换一种方式，测量过程可以隐含地构成某种平均值。例如，为了测量电流，我们可以测量传输的电荷并除以所花的时间——这给出了单位时间内传输的电子平均数的测量值。必须重视这种情况的实质性不同：常规情况下无法测量除少数选定的平均值以外的任何值，例如高阶矩就无从获取。

对比之下，在测量个别事件(如动物计数)时，我们能用显而易见的方法构成观测值的平均

$$\bar{X}_N = \frac{1}{N}\sum_{n=1}^{N} X(n) \tag{2.5.1}$$

量 $X(n)$ 是量 X 的各个观测值。我们预期当样本数 N 很大时，$\bar{X}_N$ 接近于均值 $\langle X\rangle$，事实上有

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} f[X(n)] = \lim_{N\to\infty}\overline{f(X)}_N = \langle f(X)\rangle \tag{2.5.2}$$

如果我们对所有函数 f 执行这一程序，它将确定 X 的概率密度函数。该程序的有效性取决于顺次测量之间的独立性程度，在2.5.2节中将作出论述。

在仅有平均值本身由测量方法所直接确定的情况，通常无法测出 $X(n)$，因此，要确定 $\overline{f(X)}_N$ 一般是不可能的。可资应用的就是 $f(\bar{X}_N)$——除非 f 是线性函数，它是相当不同的东西。我们常会发现可测量的量(按某种理论)与一些确定的函数的均值相联系的情况，但若希望去测量它，例如导线中电子数的一个任意函数的均值，则是没有希望的。我们谈论平均数，甚至平方平均数，但是可用的测量方法都不是直接的。我们决不会去清点不同时刻的电子数，因此也得不到任意的函数。

2.5.1 矩、相关和协方差

由矩 $\langle X^n\rangle$ 所给出的量是值得注意的，因为它们通常不难算出。但是当 $x\to\pm\infty$ 时概率密度总得消失，所以较高阶的矩仅告诉我们 X 不大会取巨大数值的特性。实用上我们发现最重要的量是跟一阶和二阶矩有关的。特别地，对于单个变量 X，方差定义为

$$\mathrm{var}\{X\} \equiv \{\sigma[X]\}^2 \equiv \langle[X-\langle X\rangle]^2\rangle \tag{2.5.3}$$

众所周知，方差 $\mathrm{var}\{X\}$ 或它的平方根标准差 $\sigma[X]$ 是 X 的值偏离其均值 $\langle X\rangle$ 的程度的一个度量。

在多变量的场合，我们定义协方差矩阵为

$$\begin{aligned}\langle X_i, X_j\rangle &\equiv \langle (X_i-\langle X_i\rangle)(X_j-\langle X_j\rangle)\rangle \\ &\equiv \langle X_iX_j\rangle - \langle X_i\rangle\langle X_j\rangle\end{aligned} \tag{2.5.4}$$

显然有

$$\langle X_i, X_i\rangle = \mathrm{var}\{X_i\} \tag{2.5.5}$$

如果变量是两两独立的，则协方差矩阵是对角阵。

2.5.2 大数法则

作为前面一些概念的应用，让我们研究下面的测量模型。假定我们对同一量测量N次，得到随机变量的样本值 $X(n)$ $(n=1, 2, \cdots, N)$。因为这全是同一个量在相续时刻的测量，我们假设对每一n，$X(n)$ 具有相同的概率分布。但并不假定 $X(n)$ 是相互独立的。然而，只要协方差$\langle X(n), X(m)\rangle$在 $|n-m|\to\infty$ 时足够迅速地消失为零，并定义

$$\overline{X}_N = \frac{1}{N}\sum_{n=1}^{N} X(n) \tag{2.5.6}$$

我们将证明

$$\lim_{N\to\infty} \overline{X}_N = \langle X\rangle \tag{2.5.7}$$

显然有

$$\langle \overline{X}_N\rangle = \langle X\rangle \tag{2.5.8}$$

现在计算 $\overline{X}_N$ 的方差并证明当 $N\to\infty$ 时它在某些条件下趋向于零。

$$\langle \overline{X}_N \overline{X}_N\rangle - \langle \overline{X}_N\rangle^2 = \frac{1}{N^2}\sum_{n,m=1}^{N}\langle X_n, X_m\rangle \tag{2.5.9}$$

在$\langle X_n, X_m\rangle$当 $|n-m|\to\infty$ 时减小得足够快的条件下，我们得到

$$\lim_{N\to\infty}(\mathrm{var}\{\overline{X}_N\}) = 0 \tag{2.5.10}$$

所以 $\lim\limits_{N\to\infty}\overline{X}_N$ 是等于$\langle X\rangle$的确定性变量。

两种$\langle X_n, X_m\rangle$的模型可被选用。

a) $\langle X_n, X_m\rangle \sim K\lambda^{|m-n|}$ $(\lambda<1)$ (2.5.11)

对此可得

$$\mathrm{var}\{\overline{X}_N\} = \frac{2K}{N^2}\left(\frac{\lambda^{N+1} - N(\lambda-1) - \lambda}{(\lambda-1)^2}\right) - \frac{K}{N} \to 0^{(*)} \tag{2.5.12}$$

(*) 原文中 λ^{N+2} 项误为 λ^{N+1}，现已作更正。——译注

b) $\langle X_n, X_m\rangle \sim |n-m|^{-1} \quad (n \neq m)$ (2.5.13)

渐近地可得

$$\mathrm{var}\{\bar{X}_N\} \sim \frac{2}{N}\log N - \frac{1}{N} \to 0 \tag{2.5.14}$$

两种情况下都有 $\mathrm{var}\{\bar{X}_N\} \to 0$，但收敛速率是很不相同的。把 n m 解释为进行测量的时刻，人们领悟到甚至衰减极慢的相关也是许可的。大数法则以多种形式出现，帕普里斯[2.2]作了很好的总结。中心极限定理是一个更加精密的结果，其中确定了 $\bar{X}_N - \langle X\rangle$ 的极限分布函数(见2.8.2节)。

2.6 特征函数

人们希望有一种条件使变量是独立的，不只是成对地独立，为此(以及其他)目标我们来定义特征函数。

如果 s 是向量$(s_1, s_2, \cdots, s_n)$，X是随机变量的向量$(X_1, X_2, \cdots, X_n)$，则特征函数(或称矩母函数)定义为

$$\phi(s) = \langle \exp(is\cdot X)\rangle = \int dx p(x)\exp(is\cdot x) \tag{2.6.1}$$

特征函数有下列性质[参考文献2.1，第15章]

i) $\phi(o) = 1$

ii) $|\phi(s)| \leqslant 1$

iii) $\phi(s)$ 对所有有限实值的自变量 s 是一致连续函数[2.5]。

iv) 如果矩$\langle \prod_i X_i^{m_i}\rangle$存在，则

$$\langle \prod_i X_i^{m_i}\rangle = \left[\prod_i \left(-i\frac{\partial}{\partial s_i}\right)^{m_i} \phi(s)\right]_{s=0} \tag{2.6.2}$$

v) 概率密度序列收敛于极限概率密度，当且仅当其相应的特征函数序列收敛于极限概率密度相应的特征函数。

vi) 傅里叶逆转公式

$$p(x) = (2\pi)^{-n}\int ds\, \phi(s)\exp(-ix\cdot s) \tag{2.6.3}$$

由这个逆转公式，$\phi(s)$以概率1确定$p(x)$。因此，特征函数确实表示了概率密度的特征。

vii) 独立随机变量：从2.3.4节中独立随机变量的定义得知，当且仅当

$$p(x_1, x_2, \cdots, x_n) = p_1(x_1)p_2(x_2)\cdots p_n(x_n) \tag{2.6.4}$$

时随机变量X_1，X_2，…是独立的，在此情况下

$$\phi(s_1, s_2, \cdots, s_n) = \phi_1(s_1)\phi_2(s_2)\cdots\phi_n(s_n) \tag{2.6.5}$$

viii) 独立随机变量之和：如果X_1，X_2，…是独立随机变量，又若

$$Y = \sum_{i=1}^{n} X_i \tag{2.6.6}$$

而Y的特征函数是

$$\phi_y(s) = \langle \exp(isY) \rangle \tag{2.6.7}$$

则有

$$\phi_y(s) = \prod_{i=1}^{n} \phi_i(s) \tag{2.6.8}$$

由于收敛性质(v)，特征函数在本书中起着重要的作用，我们可以对特征函数施行极限过程而不是概率分布本身，而且常常使证明更为容易。特征函数是真正的“特征”的事实，即逆转公式(vi)，证实了不同的特征函数来自不同的分布函数。此外，(2.6.2)直接推导矩的公式使一些确定特征函数的方法与可测量的量联系了起来。

2.7 累积母函数：相关函数与累积量

特征函数的更深一层的重要性质出自考虑它的对数

$$\Phi(s) = \log \phi(s) \tag{2.7.1}$$

我们称它为累积母函数。假定所有的矩都存在，则$\phi(s)$可展开为幂级数，因此$\Phi(s)$也可展开为

$$\Phi(s) = \sum_{r=1}^{\infty} \frac{i^r}{r!} \sum_{(m)} \langle\langle X_1^{m_1} X_2^{m_2} \cdots X_n^{m_n} \rangle\rangle s_1^{m_1} s_2^{m_2} \cdots s_n^{m_n} \delta(r, \sum_{i=1}^{n} m_i) \tag{2.7.2}$$

其中量$\langle\langle X_1^{m_1}X^{m_2}\cdots X_n^{m_n}\rangle\rangle$称为变量 $\boldsymbol{X}$ 的累积量。不应想当然地把累积量所选用的记号解释为 X 的幂的特定乘积的函数；确切地说，它指明了用矩来表达累积量时所出现的最高阶矩。斯特拉托诺维奇(Stratonovich)[2.4]也采用了术语相关函数，我们将把该术语专用于涉及多于一个 X_i 的累积量。因若 $\boldsymbol{X}$ 全为独立的话，因式分解性质(2.6.5)蕴涵着 $\Phi(s)$(累积母函数)是 n 项之和，每一项仅是一个变量 s_i 的函数。因此混合项的系数，即相关函数(按上述命名法)全为零，反之亦然。所以，相关函数的幅度是相关程度的一种度量。

累积量和相关函数可以借助于矩来算出，只要把特征函数展开为幂级数

$$\phi(s)=\sum_{r=0}^{\infty}\frac{i^r}{r!}\sum_{(m)}\langle X_1^{m_1}X_2^{m_2}\cdots X_n^{m_n}\rangle \times\frac{r!}{m_1!m_2!\cdots m_n!}\delta\left(r,\sum_{i=1}^{n}m_i\right)s_1^{m_1}s_2^{m_2}\cdots s_n^{m_n} \tag{2.7.3}$$

再把它的对数展开为幂级数，并与(2.7.2)的 $\Phi(s)$ 比较。这样做得不出简单的公式，但开头几个累积量可以列出：我们有

$$\langle\langle X_i\rangle\rangle=\langle X_i\rangle \tag{2.7.4}$$

$$\langle\langle X_iX_j\rangle\rangle=\langle X_iX_j\rangle-\langle X_i\rangle\langle X_j\rangle \tag{2.7.5}$$

$$\langle\langle X_iX_jX_k\rangle\rangle=\langle X_iX_jX_k\rangle-\langle X_iX_j\rangle\langle X_k\rangle-\langle X_i\rangle\langle X_jX_k\rangle -\langle X_iX_k\rangle\langle X_j\rangle+\langle X_i\rangle\langle X_j\rangle\langle X_k\rangle \tag{2.7.6}$$

(这里，对任意个数的同等地位的 i、j、k、l 的所有公式全都保持正确)。下面来给出一般的显式公式。假定我们要计算累积量 $\langle\langle X_1X_2X_3\cdots X_n\rangle\rangle$，其程序如下：

i) 写出 n 个点的序列……；

ii) 插入尖括弧将它分成 $p+1$ 个子集

$\langle\cdots\rangle\langle\cdots\rangle\langle\cdots\cdots\rangle\cdots\langle\cdots\rangle$；

iii) 将符号 $X_1\cdots X_n$ 分配在各个点的位置，使所有那些不同的表达式全都出现，例如

$\langle X_1\rangle\langle X_2X_3\rangle = \langle X_1\rangle\langle X_3X_2\rangle \neq \langle X_3\rangle\langle X_1X_2\rangle$;

iv) 给定 p 值，对所有这种项求和，记为 $C_p(X_1, X_2, \cdots, X_n)$;

v) $$《X_1\ X_2\cdots X_n》=\sum_{p=0}^{n-1}(-1)^p p!\,C_p(X_1, X_2, \cdots, X_n) \tag{2.7.7}$$

米隆(Meeron)[2.6] 导出了这个公式，其特殊的程序归功于范坎彭(Van Kampen)[2.7]。

vi) 具有一个或多个重复元素时的累积量，例如《$X_1^2X_3X_2$》，只要算出《$X_1X_2X_3X_4$》并在所得式子中置 $X_4 = X_1$ 即可。

2.7.1 例：4 阶累积量：《$X_1X_2X_3X_4$》

a) $p=0$

仅有的项是 $\langle X_1X_2X_3X_4\rangle \equiv C_0(X_1, X_2, X_3, X_4)$

b) $p=1$

分划 $\langle\cdot\rangle\langle\cdots\rangle$

项 $\langle X_1\rangle\langle X_2X_3X_4\rangle + \langle X_2\rangle\langle X_3X_4X_1\rangle$
$+\langle X_3\rangle\langle X_4X_1X_2\rangle + \langle X_4\rangle\langle X_1X_2X_3\rangle \equiv D_1$

分划 $\langle\cdot\cdot\rangle\langle\cdot\cdot\rangle$

项 $\langle X_1X_2\rangle\langle X_3X_4\rangle + \langle X_1X_3\rangle\langle X_2X_4\rangle$
$+\langle X_1X_4\rangle\langle X_2X_3\rangle \equiv D_2$

因此

$$D_1 + D_2 \equiv C_1(X_1, X_2, X_3, X_4)$$

c) $p=2$

分划 $\langle\cdot\rangle\langle\cdot\rangle\langle\cdot\cdot\rangle$

项 $\langle X_1\rangle\langle X_2\rangle\langle X_3X_4\rangle + \langle X_1\rangle\langle X_3\rangle\langle X_2X_4\rangle$
$+\langle X_1\rangle\langle X_4\rangle\langle X_2X_3\rangle + \langle X_2\rangle\langle X_3\rangle\langle X_1X_4\rangle$
$+\langle X_2\rangle\langle X_4\rangle\langle X_1X_3\rangle + \langle X_3\rangle\langle X_4\rangle\langle X_1X_2\rangle$
$= C_2(X_1, X_2, X_3, X_4)$

d) $p=3$

分划 $\langle\cdot\rangle\langle\cdot\rangle\langle\cdot\rangle\langle\cdot\rangle$

项 $\langle X_1\rangle\langle X_2\rangle\langle X_3\rangle\langle X_4\rangle \cdots C_3(X_1, X_2, X_3, X_4)$

因此

$$\langle\!\langle X_1X_2X_3X_4\rangle\!\rangle = C_0 - C_1 + 2C_2 - 6C_3 \tag{2.7.8}$$

2.7.2 累积量的意义

由(2.7.4, 5)两式可见开头两个累积量是均值$\langle X_i\rangle$和协方差$\langle X_i X_j\rangle$。跟高阶矩不一样，高阶累积量包含信息的重要性会降低。我们不可能令高于某个阶的所有矩都等于零，因为$\langle X^{2n}\rangle \geqslant \langle X^n\rangle^2$，由此可知，所有的矩包含其低阶矩的信息。

但是，对累积量可以协调一致地令

$$\langle\!\langle X\rangle\!\rangle = a$$

$$\langle\!\langle X^2\rangle\!\rangle = \sigma^2$$

$$\langle\!\langle X^n\rangle\!\rangle = 0 \quad (n>2)$$

我们易于用特征函数的逆转公式推出

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp[-(x-a)^2/2\sigma^2] \tag{2.7.9}$$

这是高斯密度函数。然而，似乎不能给出比这个直观的论据更多的东西。实际上，马辛凯维奇(Marcienkiewicz) [2.8, 9] 的定理证明了累积母函数不可能是高于二阶的多项式，就是说，或者除开头两个累积量以外全为零，或者有着无限多个非零累积量。累积量的最重要的意义在于用它们表示的不同变量的相关函数的定义，这进一步导致重要的逼近方法。

2.8 高斯和泊松概率分布

2.8.1 高斯分布

最最重要的概率分布是高斯分布，或称正态分布。在此我们集中列出最重要的有关结论。

如果 $\boldsymbol{X}$ 是一个 n 元高斯随机变量的向量，相应的多变量概率

密度函数为

$$p(\boldsymbol{x}) = [(2\pi)^n \det(\sigma)]^{-1/2} \exp\left[-\frac{1}{2} (\boldsymbol{x}-\bar{\boldsymbol{x}})^T \sigma^{-1} (\boldsymbol{x}-\bar{\boldsymbol{x}}) \right] \tag{2.8.1}$$

所以

$$\langle \boldsymbol{X} \rangle = \int d\boldsymbol{x}\, \boldsymbol{x} p(\boldsymbol{x}) = \bar{\boldsymbol{x}} \tag{2.8.2}$$

$$\langle \boldsymbol{X}\boldsymbol{X}^T \rangle = \int d\boldsymbol{x}\ \boldsymbol{x}\ \boldsymbol{x}^T p(\boldsymbol{x}) = \bar{\boldsymbol{x}}\ \bar{\boldsymbol{x}}^T + \sigma \tag{2.8.3}$$

特征函数为

$$\phi(\boldsymbol{s}) = \langle \exp(i\boldsymbol{s}^T \boldsymbol{X}) \rangle = \exp\left((i\boldsymbol{s}^T \bar{\boldsymbol{x}} - \frac{1}{2} \boldsymbol{s}^T \sigma \boldsymbol{s} \right) \tag{2.8.4}$$

这个特别简单的特征函数隐含着所有高于2阶的累积量为零的事实，因此意味着所有高于2阶的矩可以用1阶和2阶矩表示出来。关系式(2.8.3)意味着 σ 是协方差矩阵(如2.5.1节所定义)，即其元素的二阶相关函数。当然，σ 是对称的。

高阶矩与协方差矩阵 σ 之间的精确关系可以用矩和特征函数之间的关系直接写出来。只有在 $\bar{\boldsymbol{x}} = 0$ 时公式才是简单的，这种情况下奇阶矩为零，偶阶矩满足

$$\langle X_i X_j X_k \cdots \rangle = \frac{(2N)!}{N!\, 2^N} \{ \sigma_{ij} \sigma_{kl} \sigma_{mn} \cdots \}_{\text{sym}} \tag{2.8.4}$$

这里足标"sym"意指 σ 的乘积的对称形式(*)，而 $2N$ 是矩的阶数。例如

$$\langle X_1 X_2 X_3 X_4 \rangle = \frac{4!}{4 \cdot 2!} \left\{ \frac{1}{3} [\sigma_{12}\sigma_{34} + \sigma_{41}\sigma_{23} + \sigma_{13}\sigma_{24}] \right\}$$

$$= \sigma_{12}\sigma_{34} + \sigma_{41}\sigma_{23} + \sigma_{13}\sigma_{24} \tag{2.8.5}$$

$$\langle X_1^4 \rangle = \frac{4!}{4 \cdot 2!} \{ \sigma_{11}^2 \} = 3\sigma_{11}^2 \tag{2.8.6}$$

(*) 更确切地说，应是"σ 的乘积的对称形式除以其总的项数"。这个项数就是 $\frac{(2N)!}{N!\, 2^N}$。——译注

2.8.2 中心极限定理

因多种理由高斯分布是很重要的。实践中，许多变量在经验上可很好地用高斯变量近似，其根据来自中心极限定理。粗略地说，它断言许多独立而任意分布的组成部分加起来得到的随机变量是高斯随机变量。更精确地说，令 X_1，X_2，X_3，…，X_n，…为独立随机变量，满足

$$\langle X_i\rangle = 0,\ \mathrm{var}\{X_i\} = b_i^2 \tag{2.8.7}$$

并令 X_i 的分布函数为 $p_i(x_i)$。定义

$$S_n = \sum_{i=1}^{n} X_i \tag{2.8.8}$$

和

$$\sigma_n^2 = \mathrm{var}\{S_n\} = \sum_{i=1}^{n} b_i^2 \tag{2.8.9}$$

我们进一步要求满足林德贝格(Lindeberg)条件：

$$\lim_{n\to\infty}\left[\frac{1}{\sigma_n^2}\sum_{i=1}^{n}\int_{|x|>t\sigma_n} \mathrm{d}x\, x^2 p_i(x)\right] = 0 \tag{2.8.10}$$

对任何固定的 $t>0$ 成立。那末，在这些条件下，规范化和式 S_n/σ_n 的分布趋向于具有零均值和单位方差的高斯分布。

该定理的证明可在[2.1]中找到。然而，在此对假设条件作些注释是值得的。我们首先注意求和的项 X_i 被要求是独立的，这并非是绝对必需的条件。例如，选择

$$X_i = \sum_{r=i}^{i+j} Y_r \tag{2.8.11}$$

其中 Y_r 是独立的，因为 X 的和可以改写为 Y 的和（带有某些有限系数），所以定理仍然是正确的。

大致地讲，只要 X_i 和 X_j 之间的相关当 $|i-j|\to\infty$ 时趋向于零足够快，中心极限定理还是可望成立。林德贝格条件(2.8.10)不是一个一目了然的条件，但它是表示 $|X_i|$ 取大值的概率非常小这种要求的最弱的条件。例如，假若所有的 b_i 或是无穷大、或是大于某个常数 C，显然当 $n\to\infty$ 时 σ_n^2 发散。(2.8.10)中积

分的和是所有$|x_i|>t\sigma_n$所给于方差的"贡献"之和，显然当$n\to\infty$时每一贡献趋向于零。林德贝格条件要求所有贡献之和不如σ_n^2发散得快。实际上它是一个相当弱的要求，只要对所有X_i有$|X_i|<C$就可满足，或者当$x\to\pm\infty$时$p_i(x)$趋向于零足够快也可满足。一个例外是

$$p_i(x)=a_i[\pi(x^2+a_i^2)]^{-1} \tag{2.8.12}$$

即柯西(Cauchy)或洛仑兹(Lorentzian)分布，该分布的方差是无穷大，而且X_i的和的分布实际上具有与(2.8.12)相同的形式，其中用$\sum_{i=1}^{n}a_i$取代a_i。显然，林德贝格条件不满足。

3.3.1节中将出现一个有关的条件，也称为林德贝格条件，在那里我们要讨论离散过程用连续步进过程来取代的问题。

2.8.3 泊松分布

泊松(Poisson)分布在取值为正整数的随机变量研究中起着核心的作用。如果X是有关变量，则泊松分布由

$$P(X=x)\equiv P(x)=e^{-a}a^x/x! \tag{2.8.13}$$

所定义，显然，由

$$\langle X^r\rangle_f=\langle X(X-1)\cdots(X-r+1)\rangle \tag{2.8.14}$$

所定义的阶乘矩由

$$\langle X^r\rangle_f=a^r \tag{2.8.15}$$

所给出，对值域为非负整数的变量，很自然地可以定义母函数为

$$G(s)=\sum_{x=0}^{\infty}s^xP(x)=\langle s^x\rangle \tag{2.8.16}$$

它跟特征函数的关系是

$$G(s)=\phi(-i\log s) \tag{2.8.17}$$

母函数有着一个有用的性质

$$\langle X^r\rangle_f=\left[\left(\frac{\partial}{\partial s}\right)^rG(s)\right]_{s=1} \tag{2.8.18}$$

对泊松分布我们有

$$G(s)=\sum_{x=0}^{\infty}\frac{e^{-\alpha}(s\alpha)^x}{x!}=\exp[\alpha(s-1)] \tag{2.8.19}$$

我们还可以用

$$g(s)=\log G(s) \tag{2.8.20}$$

定义阶乘累积母函数 $g(s)$，以及由

$$g(s)=\sum_{x=1}^{\infty}\langle\langle X^r\rangle\rangle_f\frac{(s-1)^r}{r!}$$

定义阶乘累积量 $\langle\langle X^r\rangle\rangle_f$。我们看到泊松分布的阶乘累积量除第一个外全都是零。

泊松分布在本书中多处自然而然地出现，例如，我们已在1.4.1节中遇到过，它作为一个简单的主方程之解。在整数值随机变量的研究中，泊松分布起着同高斯分布在连续随机变量研究中所具有的相似的核心作用。然而，泊松分布仅有的简单的多变量推广就是泊松分布的乘积，即有

$$P(x_1,x_2,x_3,\cdots)=\prod_{i=1}^{n}\frac{e^{-\alpha_i}(\alpha_i)^{x_i}}{x_i!} \tag{2.8.21}$$

的形式，不存在相似于相关多变量高斯分布的相关多变量泊松分布的合符逻辑的概念。

2.9 随机变量序列的极限

很多计算工作由确定对随机变量的逼近所构成，其中自然出现了随机变量序列的极限的概念。但是，并没有唯一的方式去定义这种极限。

例如，假设有概率空间 Ω 和一个定义在 Ω 上的随机变量序列，那么关于当 $n\to\infty$ 时序列的极限

$$X=\lim_{n\to\infty}X_n \tag{2.9.1}$$

我们是指随机变量序列在某种意义下趋向于随机变量 X。在考虑以 ω 为元素的，具有概率密度 $p(\omega)$ 的概率空间 Ω 时，多种可能性会出现，我们可以选用下面的几个定义。

2.9.1 几乎必然极限

如果对于除了一个零概率集以外的所有的 ω 有

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \tag{2.9.2}$$

则 X_n 几乎必然收敛于 X。这样各个 X_n 的现实收敛于 X，并写成为

$$\text{ac-}\lim_{n\to\infty} X_n = X \tag{2.9.3}$$

2.9.2 均方极限(依均方收敛)

另一个可能性是将 $X_n(\omega)$ 认为是 ω 的函数，并寻求 $X_n(\omega)$ 与 $X(\omega)$ 的均方偏差，这样，如有

$$\lim_{n\to\infty}\int d\omega\, p(\omega)\,[X_n(\omega) - X(\omega)]^2 \equiv \lim_{n\to\infty}\langle (X_n - X)^2\rangle = 0 \tag{2.9.4}$$

我们就说 X_n 依均方收敛于 X。这是一种在希尔伯特 (Hilbert) 空间理论中众所周知的极限，写成

$$\text{ms-}\lim_{n\to\infty} X_n = X \tag{2.9.5}$$

2.9.3 随机极限(依概率收敛)

我们可以考虑这样的可能性：因 $X_n(\omega)$ 与 X 的有偏差的概率趋向于零而导致 $X_n(\omega)$ 趋向于 X，精确地说，意思是指如果对任何 $\varepsilon>0$

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0 \tag{2.9.6}$$

则 X_n 的随机极限是 X。

注意这个概率可用下面的方法写出来，设 $\chi_\varepsilon(z)$ 是函数

$$\begin{aligned} \chi_\varepsilon(z) &= 1 \quad |z| > \varepsilon \\ &= 0 \quad |z| \leqslant \varepsilon \end{aligned} \tag{2.9.7}$$

则有

$$P(|X_n - X| > \varepsilon) = \int d\omega\ P(\omega)\chi_\varepsilon(|X_n - X|) \tag{2.9.8}$$

在此情况，我们写成

$$\text{st-}\lim_{n\to\infty} X_n = X \tag{2.9.9}$$

2.9.4 依分布收敛

甚至更弱的收敛形式也可出现，只要对任何连续有界函数 $f(x)$ 有

$$\lim_{n\to\infty}\langle f(X_n)\rangle = \langle f(X)\rangle \tag{2.9.10}$$

在这个极限式中的收敛性称为是依分布的。特别地，用 $\exp(ixs)$ 作为 $f(x)$，我们得知特征函数列趋向于特征函数，因此 X_n 的概率密度趋向于 X 的概率密度。

2.9.5 各极限之间的关系

可以证明以下关系成立

几乎必然收敛 $\Longrightarrow$ 随机收敛。

均方收敛 $\Longrightarrow$ 随机收敛。

随机收敛 $\Longrightarrow$ 依分布收敛。

所有这些极限都有其应用。

3. 马尔可夫过程

3.1 随机过程

第1章中给出的例子都可在数学上用随机过程描述。所谓随机过程，不严格地说，是指随时间在概率意义上发展着的系统；更精确地说，则是存在着某个与时间有关的随机变量 $\boldsymbol{X}(t)$ 的系统。我们可以在时刻 $t_1、t_2、t_3、\cdots$ 测量 $\boldsymbol{X}(t)$ 的值 $x_1、x_2、x_3、\cdots$，并假定存在一组完整地描述系统的联合概率密度

$$p(x_1, t_1;\ x_2, t_2;\ x_3, t_3;\ \cdots) \tag{3.1.1}$$

借助于这些联合概率密度函数可以确定条件概率密度：

$$\begin{aligned} &p(x_1, t_1;\ x_2, t_2;\ \cdots \mid y_1, \tau_1;\ y_2, \tau_2;\ \cdots) \\ &\quad = p(x_1, t_1;\ x_2, t_2;\ \cdots;\ y_1, \tau_1;\ y_2, \tau_2;\ \cdots) \\ &\qquad / p(y_1, \tau_1;\ y_2, \tau_2;\ \cdots) \end{aligned} \tag{3.1.2}$$

这些定义的有效性与时间次序无关，虽然通常只考虑自右至左递增的时刻，即

$$t_1 \geqslant t_2 \geqslant t_3 \geqslant \cdots \geqslant \tau_1 \geqslant \tau_2 \geqslant \cdots \tag{3.1.3}$$

发展方程的概念使我们把条件概率看作是给定过去知识（在时刻 $\tau_1、\tau_2、\cdots$ 的值 $y_1、y_2、\cdots$）时对 $\boldsymbol{X}(t)$ 的未来值（在时刻 $t_1、t_2、\cdots$ 的值 $x_1、x_2、\cdots$）的预测。

一般的随机过程概念是很宽松的，为了规定随机过程我们至少需要知道所有可能的(3.1.1) 式那种联合概率。如果这些知识真正规定了过程，它被称为是可分随机过程。本书中所考虑的所有过程都假定是可分的。

最简单的一类随机过程是完全独立的过程，

$$p(x_1, t_1; x_2, t_2; x_3, t_3; \cdots) = \prod_i p(x_i, t_i) \tag{3.1.4}$$

这意味着X在时刻t的值与它在过去(或未来)的值完全独立。一个更加特殊的情况是$p(x_i, t_i)$与t_i无关，所以在所有的时刻由同样的概率法则支配着过程，这时我们就有了所谓伯努利(Bernoulli)试验，其中在相继的时刻不断重复着一个概率过程。

下一个最简单的概念是马尔可夫过程，在这种过程中仅由现在的知识决定未来。

3.2 马尔可夫过程

马尔可夫假设是用条件概率来陈述的。我们要求，如果各个时刻满足(3.1.3)的次序，则条件概率完全由最近时刻的条件所确定，即

$$\begin{aligned} &p(x_1, t_1; \ x_2, t_2; \ \cdots | y_1, \tau_1; \ y_2, \tau_2; \ \cdots) \\ &\quad = p(x_1, t_1; \ x_2, t_2; \ \cdots | y_1, \tau_1) \end{aligned} \tag{3.2.1}$$

这正是爱因斯坦和斯莫路乔夫斯基等人所作假设的更精确的陈述。从这假设本身就可知道它是非常强有力的，因为它意味着我们可以依据简单的条件概率$p(x_1, t_1 | y_1, \tau_1)$限定所有的东西。例如，按条件概率密度的定义

$$\begin{aligned} &p(x_1, t_1; \ x_2, t_2 | y_1, \tau_1) \\ &\quad = p(x_1, t_1 | x_2, t_2; \ y_1, \tau_1) p(x_2, t_2 | y_1, \tau_1) \end{aligned}$$

并应用马尔可夫假设(3.2.1)，我们求得

$$p(x_1, t_1; \ x_2, t_2 | y_1, \tau_1) = p(x_1, t_1 | x_2, t_2) p(x_2, t_2 | y_1, \tau_1). \tag{3.2.2}$$

不难看出任何联合概率可简单地表示为

$$\begin{aligned} &p(x_1, t_1; \ x_2, t_2; \ x_3, t_3; \ \cdots; \ x_n, t_n) \\ &\quad = p(x_1, t_1 | x_2, t_2) p(x_2, t_2 | x_3, t_3) p(x_3, t_3 | x_4, t_4) \\ &\qquad \cdots p(x_{n-1}, t_{n-1} | x_n, t_n) p(x_n, t_n) \end{aligned} \tag{3.2.3}$$

只要满足条件

$$t_1 \geqslant t_2 \geqslant t_3 \geqslant \cdots \geqslant t_{n-1} \geqslant t_n \tag{3.2.4}$$

3.2.1 相容性——查普曼-柯尔莫哥洛夫方程

由2.3.3节，要求在联合概率中同一类的所有互斥事件求和时消去该变量，即

$$\sum_B P(A\cap B\cap C\cdots)=P(A\cap C\cdots) \tag{3.2.5}$$

把它用于随机过程，我们可得到两个相似而易使人误解的方程：

$$\begin{aligned} p(x_1,t_1)&=\int dx_2\, p(x_1,t_1;\ x_2,t_2)\\ &=\int dx_2\, p(x_1,t_1|x_2,t_2)p(x_2,t_2) \end{aligned} \tag{3.2.6}$$

该方程是对所有随机过程都成立的恒等式，它是不同层次的方程中的第一层次的方程，第二层次的方程是

$$\begin{aligned} p(x_1,t_1|x_3,t_3)&=\int dx_2\, p(x_1,t_1;\ x_2,t_2|x_3,t_3)\\ &=\int dx_2\, p(x_1,t_1|x_2,t_2;x_3,t_3)p(x_2,t_2|x_3,t_3) \end{aligned} \tag{3.2.7}$$

这个方程也总是正确的。现在我们引入马尔可夫假设，如果$t_1\geqslant t_2\geqslant t_3$，我们可以去掉两重条件概率中对$t_3$的依赖，写成

$$p(x_1,t_1|x_3,t_3)=\int dx_2\, p(x_1,t_1|x_2,t_2)p(x_2,t_2|x_3,t_3) \tag{3.2.8}$$

这就是查普曼-柯尔莫哥洛夫方程。

式(3.2.8)和式(3.2.6)之间的实质性区别是什么？显而易见的答案是：式(3.2.6)是关于无条件概率的，而式(3.2.7)是关于条件概率的。方程(3.2.8)是把所有条件概率$p(x_i,t_i|x_j,t_j)$相互联系起来的一个相当复杂的非线性泛函方程，而式(3.2.6)只是在给定条件概率$p(x_1,t_1|x_2,t_2)$时构成了t_2的未来时刻t_1处的单一时间概率。

查普曼-柯尔莫哥洛夫方程有许多解。3.4.1节中将在某些相当轻微的条件下导出方程的微分形式，由此可以很好地理解那

些不同的解。

3.2.2 离散状态空间

在离散变量的场合，我们将使用符号 $\boldsymbol{N}=(N_1, N_2, N_3, \cdots)$，其中 N_i 是取整值的随机变量。显然，现在我们要作替换

$$\int \mathrm{d}x \longleftrightarrow \sum_n \tag{3.2.9}$$

并对这样的过程把查普曼-柯尔莫哥洛夫方程写成为

$$P(n_1, t_1 | n_3, t_3)=\sum_{n_2} P(n_1, t_1 | n_2, t_2) P(n_2, t_2 | n_3, t_3) \tag{3.2.10}$$

这是矩阵的乘法，可以是无限阶的矩阵。

3.2.3 更一般的测度

更一般的公式要假定有一个测度 $\mathrm{d}\mu(x)$ 取代 $\mathrm{d}x$，这可以作多种多样的选择。例如，如果 $\mu(x)$ 是在 x 的整值处有阶梯的阶梯函数，我们就重新得出了离散状态空间形式。绝大多数的数学研究工作试图做得尽可能地一般化，但对应用而言，这种一般性可能导致缺乏明晰性，所以只要有可能，我们宁愿用较为特殊的记号。

3.3 随机过程连续性

随机变量 $\boldsymbol{X}(t)$ 是否有连续的可能值域这个问题完全不同于 $\boldsymbol{X}(t)$ 的样本路径是否是 t 的连续函数的问题。例如，在速度为 $\boldsymbol{V}(t)$ 的分子组成的气体中，显然 $\boldsymbol{V}(t)$ 的所有可能值在原则上是可实现的，所以 $\boldsymbol{V}(t)$ 的值域是连续的。然而，常常要考虑由刚球组成的气体中的瞬时碰撞模型，在这种模型中，碰撞前的速度 v_i 将在撞击的瞬间改变为另一个值 v_f，所以 $\boldsymbol{V}(t)$ 的样本路径是不连续的。不过，可以预期在这个模型中的气体分子的位置 $\boldsymbol{X}(t)$ 是连续的。

现在出现了一个主要问题：具有连续样本路径的马尔可夫过程究竟是否实际存在？请注意马尔可夫性和连续性的组合。在古

典图景(亦即非量子力学)中,这几乎是肯定的,所有具有连续值域的变量有着连续样本路径。即使是上述理想化的刚球气体中,更实际一些,也必须允许某种势能起作用,它会使撞击过程中的转向是连续的。但是也会有这样的情况:如果我们在如此精细的时间尺度上作观测,过程或许就不是马尔可夫的了。对于预测未来,甚至是概率性的未来,整个系统的当前一段历史几乎肯定是需要的。所有从力学导出马尔可夫概率方程的企图都证实了这一点。导出的方程真正是马尔可夫的则是罕见的——确切地说是存在某个特定的记忆时间,在各个时刻之前那么长的历史是重要的(哈克(Haake)[3.1])。

这意味着作为马尔可夫过程的事物实际上并不存在,倒是可能有这样的系统,与实施观测的时间尺度相比较其记忆时间是如此之小,以至合理地认为它们可用马尔可夫过程很好地近似。但在这种情况下,样本路径是否真正连续的问题不是实质性的。近似的马尔可夫过程的样本路径确实未必是连续的。即使分子碰撞不是用精确的刚球模型,在碰撞时间内,发生了一个有限的速度变更,这将作为一个离散阶梯出现于近似马尔可夫过程。在这个时间尺度上,甚至位置也可以不连续地变化,这就给出了如同爱因斯坦所模型化的布朗运动的图景。

在化学反应中,例如说,一个单独的进行到底的反应所花的时间——与分子碰撞时间的数量级大体上相同——也提供了另一个最小时间,因为在这一段时间内存在着不能用个别分子描述的状态。因此在这里,借助于个别分子状态的真实描述须要考虑某个最小时间尺度。

然而,具有连续样本路径的马尔可夫过程在数学上确实存在,并在描述现实世界时是有用的。上面提及的气体模型给出了一个有用的例子。分子的位置用离散跳跃的不连续变更去模型化的确很可能是最好的,但与行进距离相比较,这些跳跃是无穷小的。连续曲线对样本路径则提供了一个良好的近似。另一方面,速度的可能的改变量有着与实践中可达到的典型值相同的数量级。气

体中分子的平均速度约为 1000 m/s，并在碰撞中能轻而易举地改变其符号。速度简直无法(以任何有意义的概率)达到可以把速度改变量认为是非常之小的那种值，因此用连续路径来描述速度是没有什么意义的。

3.3.1 连续马尔可夫过程的数学定义

对于一个马尔可夫过程，可以证明[3.2]，如果对任何 $\varepsilon>0$ 有

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|x-z|>\varepsilon} dx\, p(x, t+\Delta t|z, t)=0 \tag{3.3.1}$$

对 x, t 和 Δt 一致地成立，则样本路径以概率 1 是 t 的连续函数。

上述条件意味着当 Δt 趋向于零时，最终位置 x 与 z 有有限差距的概率比 Δt 更快地趋向于零[方程(3.3.1) 有时被称为林德贝格条件]。

例

i) 爱因斯坦对他的 $f(x, t)$ 求得的解(1.2.1 节) 实际上是条件概率 $p(x, t|0, 0)$。按他的方法我们会求得

$$p(x, t+\Delta t|z, t)=(4\pi D\Delta t)^{-\frac{1}{2}}\exp\left[\frac{-(x-z)^2}{4D\Delta t}\right] \tag{3.3.2}$$

易于检验在此场合(3.3.1)是满足的，所以，布朗运动在爱因斯坦的公式表示下具有连续样本路径。

ii) 柯西过程，假设

$$p(x, t+\Delta t|z, t)=\frac{\Delta t}{\pi}\Big/[(x-z)^2+\Delta t^2] \tag{3.3.3}$$

这个条件概率并不满足(3.3.1)，样本路径是不连续的。

然而，在两种情况中都如相容性所要求的有

$$\lim_{\Delta t\to 0} p(x, t+\Delta t|z, t)=\delta(x-z) \tag{3.3.4}$$

易于证明查普曼-柯尔莫哥洛夫方程在两种情况下都满足。

前述两种过程之间的差别的例示见图 3.1，其中给出了两个过程的模拟，两者之间的差别是令人感兴趣的。但要注意即使是

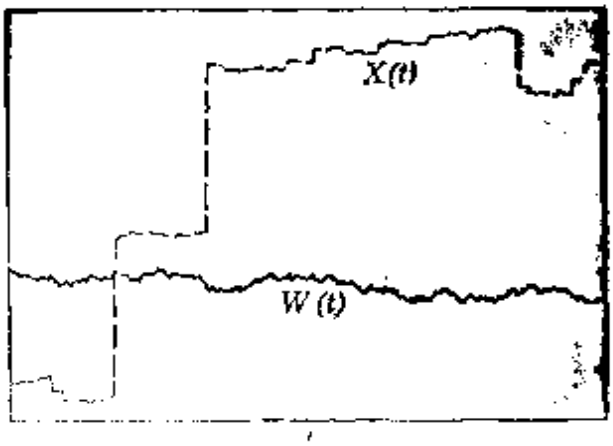


图 3.1　柯西过程 $X(t)$(虚线)和布朗运动 $W(t)$(实线)的样本路径的图示

布朗运动曲线也是极端不规则的，虽然它是连续的——事实上是处处不可微的。当然，柯西过程曲线是杂乱地不连续的。

3.4　微分查普曼-柯尔莫哥洛夫方程

在适当的假设下，查普曼-柯尔莫哥洛夫方程可简化成为微分方程。所作的假设与所考虑的过程的连续性质有密切的联系。由于连续性条件的形式 (3.3.1)，人们被引向考虑一个方法将可微性条件分为若干部分，一部分对应于代表点的连续运动，其他的对应于不连续运动。

对所有的 $\varepsilon>0$ 我们要求下列条件成立：

i) $$\lim_{\Delta t\to 0} p(\boldsymbol{x},\ t+\Delta t|\boldsymbol{z},\ t)/\Delta t = W(\boldsymbol{x}|\boldsymbol{z},\ t) \tag{3.4.1}$$

对 $|\boldsymbol{x}-\boldsymbol{z}|\geqslant\varepsilon$ 关于 $\boldsymbol{x}$, $\boldsymbol{z}$, 和 t 一致地成立；

ii) $$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|\boldsymbol{x}-\boldsymbol{z}|<\varepsilon} \mathrm{d}\boldsymbol{x}\ (x_i-z_i)p(\boldsymbol{x},\ t+\Delta t|\boldsymbol{z},\ t)$$

$$= A_i(\boldsymbol{z},\ t)+O(\varepsilon) \tag{3.4.2}$$

iii) $$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|\boldsymbol{x}-\boldsymbol{z}|<\varepsilon} \mathrm{d}\boldsymbol{x}\ (x_i-z_i)(x_j-z_j)p(\boldsymbol{x},\ t+\Delta t|\boldsymbol{z},\ t)$$

$$= B_{ij}(\boldsymbol{z},\ t)+O(\varepsilon) \tag{3.4.3}$$

后两式关于 $\boldsymbol{z}$, ε 和 t 一致地成立。

注意(3.4.2, 3)形式的高阶系数必定为零. 例如，考虑由

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|x-z|<\varepsilon} \mathrm{d}x(x_i-z_i)(x_j-z_j)(x_k-z_k)\,p(x,\ t+\Delta t|z,\ t)$$

$$=C_{ijk}(z,\ t)+O(\varepsilon) \tag{3.4.4}$$

定义的三阶量，因 C_{ijk} 关于 i、j、k 对称，考虑

$$\sum_{i,j,k}\alpha_i\alpha_j\alpha_k C_{ijk}(z,\ t)\equiv\bar{C}(\alpha,\ z,\ t) \tag{3.4.5}$$

故有

$$C_{ijk}(z,\ t)=\frac{1}{3!}\frac{\partial^3}{\partial\alpha_i\partial\alpha_j\partial\alpha_k}\bar{C}(\alpha,\ z,\ t) \tag{3.4.6}$$

因此

$$\begin{aligned}|\bar{C}(\alpha,\ z,\ t)|&\leqslant\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|x-z|<\varepsilon}|\alpha\cdot(x-z)|\,[\alpha\cdot(x-z)]^2\\&\quad\times p(x,\ t+\Delta t|z,\ t)\mathrm{d}x+O(\varepsilon)\\&\leqslant|\alpha|\,\varepsilon\lim_{\Delta t\to 0}\int[\alpha\cdot(x-z)]^2\\&\quad\times p(x,\ t+\Delta t|z,\ t)\mathrm{d}x+O(\varepsilon)\\&=\varepsilon|\alpha|\,[\sum_{i,j}\alpha_i\alpha_j B_{ij}(z,\ t)+O(\varepsilon)]+O(\varepsilon)\\&=O(\varepsilon)\end{aligned} \tag{3.4.7}$$

所以 C 为零. 相似地，我们可以证明所有相应的高阶量也是零.

按连续性条件(3.3.1)，如果 $W(x|z,\ t)$ 对所有 $x\neq z$ 为零，则过程仅有连续路径. 这样，这个函数必然以某种方式描述不连续运动，同时量 A_i 和 B_{ij} 必定与连续运动相联系.

3.4.1 微分查普曼-柯尔莫哥洛夫方程的推导

我们考虑两次连续可微的函数 $f(z)$ 的期望值的时间发展，这样有

$$\begin{aligned}&\partial_t\int\mathrm{d}x\,f(x)p(x,\ t|y,t')\\&=\lim_{\Delta t\to 0}\Big\{\int\mathrm{d}x\,f(x)\,[p(x,\ t+\Delta t|y,\ t')\\&\quad-p(x,\ t|y,\ t')]\Big\}\Big/\Delta t\end{aligned} \tag{3.4.8}$$

$$= \lim_{\Delta t \to 0} \left\{ \int dx \int dz f(x) p(x, t+\Delta t | z, t) p(z, t | y, t') - \int dz f(z) p(z, t | y, t') \right\} \Big/ \Delta t \tag{3.4.9}$$

这里我们已将查普曼-柯尔莫哥洛夫方程用于(3.4.8)的加项，产生了(3.4.9)式中相应的项。

现在我们把关于 x 的积分分成积分区域为 $|x-z| \geqslant \varepsilon$ 和 $|x-z| < \varepsilon$ 的两部分。当 $|x-z| < \varepsilon$，因 $f(z)$ 由假设是两次连续可微的，我们可写出

$$\begin{aligned} f(x) = f(z) &+ \sum_i \frac{\partial f(z)}{\partial z_i} (x_i - z_i) \\ &+ \sum_{i,j} \frac{1}{2} \frac{\partial^2 f(z)}{\partial z_i \partial z_j} (x_i - z_i)(x_j - z_j) \\ &+ |x-z|^2 R(x, z) \end{aligned} \tag{3.4.10}$$

这里我们有(仍由两次可微性)

$$\text{当 } |x-z| \to 0 \text{ 时 } |R(x, z)| \to 0 \tag{3.4.11}$$

现在代入(3.4.9)，

$$\begin{aligned} (3.4.9) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Bigg\{ & \iint_{|x-z|<\varepsilon} dx\, dz \Big[\sum_i (x_i - z_i) \frac{\partial f}{\partial z_i} \\ & + \sum_{i,j} \frac{1}{2} (x_i - z_i)(x_j - z_j) \frac{\partial^2 f}{\partial z_i \partial z_j} \Big] \\ & \times p(x, t+\Delta t | z, t) p(z, t | y, t') \\ & + \iint_{|x-z|<\varepsilon} dx dz\, |x-z|^2 R(x, z) \\ & \times p(x, t+\Delta t | z, t) p(z, t | y, t') \\ & + \iint_{|x-z|\geqslant\varepsilon} dx\, dz f(x) p(x, t+\Delta t | z, t) \\ & \times p(z, t | y, t') \\ & + \iint_{|x-z|<\varepsilon} dx dz f(z) p(x, t+\Delta t | z, t) \\ & \times p(z, t | y, t') \end{aligned}$$

$$-\iint \mathrm{d}x\mathrm{d}z f(z)p(x,\ t+\Delta t|z,\ t)p(z,\ t|y,\ t')\Big\} \tag{3.4.12}$$

[注意 $p(x,\ t+\Delta t|z,\ t)$ 是概率密度，最后一项中关于 x 的积分给出值为 1，所以这一项正是(3.4.9)中的最后一项]。

现在我们逐行来考虑上式。

第 1～3 行：由一致收敛的假定，在积分号内取极限可得[应用 3.4 节的条件(ii)和(iii)]

$$\int \mathrm{d}z\Big[\sum_i A_i(z,\ t)\frac{\partial f}{\partial z_i}+\frac{1}{2}\sum_{i,j}B_{ij}(z,\ t)\frac{\partial^2 f}{\partial z_i \partial z_j}\Big]$$
$$\times p(z,\ t|y,\ t')+O(\varepsilon) \tag{3.4.13}$$

第 4、5 行：这是一个余项，当 $\varepsilon\to 0$ 时趋于零，因为

$$\Big|\frac{1}{\Delta t}\int_{|x-z|<\varepsilon} \mathrm{d}x|x-z|^2 R(x,\ z)p(x,t+\Delta t|z,t)\Big|$$
$$\leqslant\Big[\frac{1}{\Delta t}\int_{|x-z|<\varepsilon}\mathrm{d}x|x-z|^2 p(x,\ t+\Delta t|z,\ t)\Big]$$
$$\times \underset{|x-z|<\varepsilon}{\mathrm{Max}}\ |R(x,\ z)| \tag{3.4.14}$$
$$\to\{\sum_i B_{ii}(z,\ t)+O(\varepsilon)\}\{\underset{|x-z|<\varepsilon}{\mathrm{Max}}\ |R(x,\ z)|\}$$

由(3.4.11)知，当 $\varepsilon\to 0$ 时花括弧中的因子趋向于零。

第 6～10 行：我们可以把这几项全都合在一起，得到

$$\iint_{|x-z|>\varepsilon}\mathrm{d}x\mathrm{d}z f(z)[W(z|x,\ t)p(x,\ t|y,\ t')$$
$$-W(x|z,\ t)p(z,\ t|y,\ t')] \tag{3.4.15}$$

(3.4.12)式的右端整体上是同 ε 无关的，因此，取 $\varepsilon\to 0$，我们得到

$$\partial_t\int \mathrm{d}z f(z)p(z,\ t|y,\ t')$$
$$=\int \mathrm{d}z\Big[\sum_i A_i(z,\ t)\frac{\partial f(z)}{\partial z_i}+\frac{1}{2}\sum_{i,j}B_{ij}(z,\ t)\frac{\partial^2 f(z)}{\partial z_i \partial z_j}\Big]$$
$$\times p(z,\ t|y,\ t')$$
$$+\int \mathrm{d}z f(z)\Big\{\int \mathrm{d}x[W(z|x,\ t)p(x,\ t|y,\ t')$$

$$-W(x|z,\ t)p(z,\ t|y,\ t')]\Big\} \tag{3.4.16}$$

但要注意，我们用了

$$\lim_{\varepsilon\to 0}\int_{|x-z|>\varepsilon} dxF(x,\ z)\equiv\int dxF(x,\ z) \tag{3.4.17}$$

来定义函数 $F(x,\ z)$ 的主值积分。为使(3.4.16)有其意义，这种积分应该存在。方程(3.4.1)仅对 $x\neq z$ 作出定义，因而留下了在 $x=z$ 时取值无穷大的可能性，3.3.1 节中讨论的柯西过程的情况正是这样，对此有

$$W(x|z,\ t)=1/[\pi(x-z)^2] \tag{3.4.18}$$

然而，如果 $p(x,\ t|y,\ t')$ 为连续且一次可微，则主值积分存在。本书中此后将不再明显地写出主值积分形式，因为难得有必要去考虑这种奇异场合。

最后一步是作偏积分，我们得到

$$\begin{aligned}&\int dz f(z)\partial_t p(x,\ t|y,\ t')\\&=\int dz\ f(z)\Big\{-\sum_i\frac{\partial}{\partial z_i}A_i(z,\ t)p(z,\ t|y,\ t')\\&+\sum_{i,j}\frac{1}{2}\frac{\partial^2}{\partial z_i\partial z_j}B_{ij}(z,\ t)p(z,\ t|y,\ t')\\&+\int dx[W(z|x,\ t)p(x,\ t|y,t')\\&-W(x|z,\ t)p(z,\ t|y,\ t')]\Big\}\\&+\text{曲面项}\end{aligned} \tag{3.4.19}$$

我们还未指定积分区域，假定过程被限制于具有表面 S 的区域 R，则显然有

$$p(x,\ t|z,\ t')=0\quad\text{除非 } x \text{ 和 } z \text{ 均}\in R \tag{3.4.20}$$

由定义我们又显然有

$$W(x|z,\ t)=0\quad\text{除非 } x \text{ 和 } z \text{ 均}\in R \tag{3.4.21}$$

但是，关于(3.4.2, 3)所定义的函数 $A_i(z,\ t)$ 和 $B_{ij}(z,\ t)$，上述条件可能导致它们为不连续，这是因为当 z 穿过边界时条件概率

$p(x, t+\Delta t|z, t')$ 可以合情合理地不连续变化，反映了不允许从 R 外向 R 内转移的事实。

在分部积分时，我们必须对 A_i 和 B_{ij} 都进行微分，而由上面的推理，不可能假定在区域的边界上这种微分可行，因此，我们选择 $f(z)$ 仅在一个完全包含在 R 中的区域 R' 上可任意取值，此外都为零。因此我们可推断，对所有 R 的内部的点 z 有

$$\begin{aligned}\partial_t p(z, t|y, t') = &-\sum_i \frac{\partial}{\partial z_i}[A_i(z, t)p(z, t|y, t')] \\ &+\sum_{i,j}\frac{1}{2}\frac{\partial^2}{\partial z_i \partial z_j}[B_{ij}(z, t)p(z, t|y, t')] \\ &+\int dx[W(z\ x, t)p(x, t|y, t') \\ &-W(x|z, t)p(z, t|y, t')].\end{aligned} \tag{3.4.22}$$

曲面项不存在，因为它们必为零。

这个方程在文献中似乎没有统一的名称。因为它纯粹是查普曼-柯尔莫哥洛夫方程的微分形式，我建议称它为微分查普曼-柯尔莫哥洛夫方程。

3.4.2 微分查普曼-柯尔莫哥洛夫方程的地位

前面的推导尚未令我们澄清这样的问题：微分查普曼-柯尔莫哥洛夫方程的解到什么样的限度内也是查普曼-柯尔莫哥洛夫方程自己的解，或者到什么样的限度内解的确存在。然而，确凿无疑的是，一组服从查普曼-柯尔莫哥洛夫方程的条件概率一定能产生一个马尔可夫过程，由这些条件概率生成的联合概率满足所有的概率公理。

可以证明 [3.3]，在某些条件下，如果我们指定 $A(x, t)$，$B(x, t)$（它必须是半正定）和 $W(x|y, t)$（它必须是非负的），则微分查普曼-柯尔莫哥洛夫方程存在一个非负解，而且此解也满足查普曼-柯尔莫哥洛夫方程。应满足的是初始条件

$$p(\boldsymbol{x}, t|\boldsymbol{y}, t) = \delta(\boldsymbol{y}-\boldsymbol{x})$$

它由自条件概率密度的定义，以及任何适当的边界条件。对最一般形式的方程这些条件是很难指定的，但在第 5 章中就福克-普朗克方程的场合(3.5.2 节)给出了条件。

3.5 条件和结论的解释

现在可以看到 3.4 节的各个条件(i)、(ii)、(iii)产生方程的不同组成部分，它们的解释是相当直截了当的。我们可以识别出有三种过程发生，分别称为跳跃、漂移和扩散。

3.5.1 跳过程：主方程

考虑

$$\boldsymbol{A}_i(\boldsymbol{x}, t) = B_{ij}(\boldsymbol{x}, t) = 0 \tag{3.5.1}$$

的情况，我们有主方程

$$\partial_t p(\boldsymbol{x}, t|\boldsymbol{y}, t') = \int d\boldsymbol{x}\,[W(\boldsymbol{x}|\boldsymbol{x}, t)p(\boldsymbol{x}, t|\boldsymbol{y}, t') - W(\boldsymbol{x}|\boldsymbol{x}, t)p(\boldsymbol{x}, t|\boldsymbol{y}, t')] \tag{3.5.2}$$

下面来近似求解它至 Δt 的一阶项，注意到

$$p(\boldsymbol{x}, t|\boldsymbol{y}, t) = \delta(\boldsymbol{y}-\boldsymbol{x}) \tag{3.5.3}$$

因此

$$p(\boldsymbol{x}, t+\Delta t|\boldsymbol{y}, t) = \delta(\boldsymbol{y}-\boldsymbol{x})\left[1 - \int d\boldsymbol{x} W(\boldsymbol{x}|\boldsymbol{y}, t)\Delta t\right] + W(\boldsymbol{x}|\boldsymbol{y}, t)\Delta t \tag{3.5.4}$$

可见对任何 Δt，粒子停留在原来位置 $\boldsymbol{y}$ 的概率是一个有限值，由(3.5.4)式中 $\delta(\boldsymbol{y}-\boldsymbol{x})$ 的系数给出。而那些未停留在 $\boldsymbol{y}$ 的粒子的分布由适当地归一化以后的 $W(\boldsymbol{x}|\boldsymbol{y}, t)$ 所给出。这样，典型路径 $\boldsymbol{X}(t)$ 将由直线段 $\boldsymbol{X}(t)$ = 常数给出，其中散置着不连续跳跃，其分布由 $W(\boldsymbol{x}|\boldsymbol{y}, t)$ 给出。出于这个理由，该过程被称为跳过程，其路径在离散点上不连续。

在状态空间仅由整数构成的场合，主方程取

$$\partial_t P(n, t|n', t') = \sum_m [W(n|m, t)P(m, t|n', t') - W(m|n, t)P(n, t|n', t')] \quad (3.5.5)$$

的形式，毫无疑问只可能出现跳跃，因为状态变量 $N(t)$ 只能取离散值。然而，极为重要的是应该明白，既使变量 $X(t)$ 可取值于连续值域，纯粹的跳过程依然可能出现。

3.5.2 扩散过程——福克-普朗克方程

如果我们假定量 $W(z|x, t)$ 为零，微分查普曼-柯尔莫哥洛夫方程简化成为福克-普朗克方程：

$$\frac{\partial p(z, t|y, t')}{\partial t} = -\sum_i \frac{\partial}{\partial z_i}[A_i(z, t)p(z, t|y, t')] + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial z_i \partial z_j}[B_{ij}(z, t)p(z, t|y, t')] \quad (3.5.6)$$

相应的过程在数学上称为扩散过程。向量 $A(z, t)$ 称为漂移向量，矩阵 $\underset{\sim}{B}(z, t)$ 称为扩散矩阵。作为定义(3.4.3)的结果，扩散矩阵是半正定对称的。不难从 $W(x|z, t)$ 的定义(3.4.1)看到，如果 $W(x|z, t)$ 为零，则样本路径连续性的要求(3.3.1)是满足的。因此，福克-普朗克方程描述了一个具有连续样本路径的过程 $X(t)$。事实上，我们可以启发式地给出过程的多得多的明确描述。给定

$$p(z, t|y, t) = \delta(z-y) \quad (3.5.7)$$

我们来考虑 $p(z, t+\Delta t|y, t)$ 的计算。对于微小的 Δt，福克-普朗克方程的解仍具有完整的陡峭尖峰形状，因此，与 p 的导数相比较，$A_i(z, t)$ 和 $B_{ij}(z, t)$ 的导数可忽略不计。这样一来，我们可近似地把问题简化成求解

$$\frac{\partial p(z, t|y, t')}{\partial t} = -\sum_i A_i(y, t)\frac{\partial p(z, t|y, t')}{\partial z} + \sum_{i,j}\frac{1}{2}B_{ij}(y, t)\frac{\partial^2 p(z, t|y, t')}{\partial z_i \partial z_j} \quad (3.5.8)$$

其中对于微小 $t-t'$ 值我们也已忽略了 A_i 和 B_{ij} 的时间依赖性。

现在可以求出方程(3.5.8)的满足初始条件(3.5.7)的解，我们得到

$$p(\boldsymbol{z}, t+\Delta t|\boldsymbol{y}, t) = (2\pi)^{-N/2}\{\det[\underset{\sim}{\boldsymbol{B}}(\boldsymbol{y}, t)]\}^{-1/2}[\Delta t]^{-1/2} \times \exp\left\{-\frac{1}{2}\right. \left.\times \frac{[\boldsymbol{z}-\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{y}, t)\Delta t]^T[\underset{\sim}{\boldsymbol{B}}(\boldsymbol{y}, t)]^{-1}[\boldsymbol{z}-\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{y}, t)\Delta t]}{\Delta t}\right\} \tag{3.5.9}$$

即是一个具有方差矩阵 $\underset{\sim}{\boldsymbol{B}}(\boldsymbol{y}, t)$ 和均值 $\boldsymbol{y}+\boldsymbol{A}(\boldsymbol{y}, t)\Delta t$ 的高斯分布。因此我们得到了这样的系统运动图景：它以速度 $\boldsymbol{A}(\boldsymbol{y}, t)$ 作整体漂移，并叠加上具有协方差矩阵 $\underset{\sim}{\boldsymbol{B}}(\boldsymbol{y}, t)$ 的高斯涨落，即可写成

$$\boldsymbol{y}(t+\Delta t) = \boldsymbol{y}(t) + \boldsymbol{A}(\boldsymbol{y}(t), t)\Delta t + \boldsymbol{\eta}(t)\Delta t^{1/2} \tag{3.5.10}$$

这里

$$\langle \boldsymbol{\eta}(t)\rangle = 0 \tag{3.5.11}$$

$$\langle \boldsymbol{\eta}(t)\boldsymbol{\eta}(t)^T\rangle = \underset{\sim}{\boldsymbol{B}}(\boldsymbol{y}, t) \tag{3.5.12}$$

不难看出这个描述给出了：

i) 恒为连续的样本路径，因为当 $\Delta t \to 0$ 时显然有 $\boldsymbol{y}(t+\Delta t) \to \boldsymbol{y}(t)$；

ii) 处处不可微的样本路径，因为 $\Delta t^{1/2}$ 出现于(3.5.10) 中。

稍后在第 4 章中将会看到，启发式的描述(3.5.10)可以大大精确化并导致随机微分方程的概念。

3.5.3 确定性过程——刘维方程

在微分查普曼-柯尔莫哥洛夫方程 (3.4.22) 中可能只有首项不为零，这导致刘维方程的特殊情况：

$$\frac{\partial p(\boldsymbol{z}, t|\boldsymbol{y}, t')}{\partial t} = -\sum_i \frac{\partial}{\partial z_i}[A_i(\boldsymbol{z}, t)p(\boldsymbol{z}, t|\boldsymbol{y}, t')] \tag{3.5.13}$$

这方程出现在经典力学中，它刻划了一个完全确定性的运动，就是

说，如果 $x(y, t)$ 是常微分方程

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = A[x(t), t] \tag{3.5.14}$$

的满足

$$x(y, t') = y \tag{3.5.15}$$

的解，则(3.5.13)的满足初始条件

$$p(z, t'|y, t') = \delta(z-y) \tag{3.5.16}$$

的解为

$$p(z, t|y, t') = \delta[z-x(y, t)] \tag{3.5.17}$$

证明这一断言的最好方法是直接代入。由

$$-\sum_i \frac{\partial}{\partial z_i}\{A_i(z, t)\delta[z-x(y, t)]\} \tag{3.5.18}$$

$$= -\sum_i \frac{\partial}{\partial z_i}\{A_i[x(y, t), t]\delta[z-x(y, t)]\} \tag{3.5.19}$$

$$= -\sum_i \left\{A_i[x(y, t), t]\ \frac{\partial}{\partial z_i}\delta[z-x(y, t)]\right\} \tag{3.5.20}$$

以及

$$\frac{\partial}{\partial t}\delta[z-x(y, t)]$$

$$= -\sum_i \frac{\partial}{\partial z_i}\delta[z-x(y, t)]\ \frac{\mathrm{d}x_i(y, t)}{\mathrm{d}t} \tag{3.5.21}$$

并利用(3.5.14)，可见(3.5.20, 21)是相等的。这样，如果粒子在时刻 t' 位于完全确定的初始位置 y，它将保持在求解常微分方程(3.5.14)所得的轨道上。

因此，由(3.5.14)形式的一阶微分方程所定义的确定性运动是马尔可夫过程的一个初等形式。当然，解(3.5.17)仅是(3.5.9)所近似表示的那种过程的特殊情况，其中高斯部分为零。

3.5.4 一般过程

一般地，$A(z, t)$、$B(z, t)$ 和 $W(x|z, t)$ 中的量未必为零，这种情况下所得过程的样本路径如图 3.2 所示，即是一个分段连续

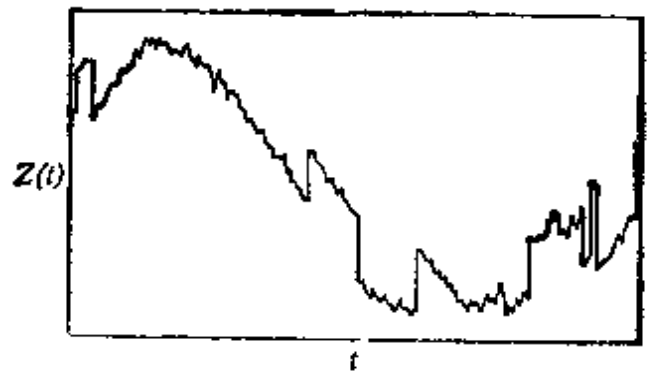


图 3.2　一般的马尔可夫过程样本路径的图示，其中存在漂移、扩散和跳跃

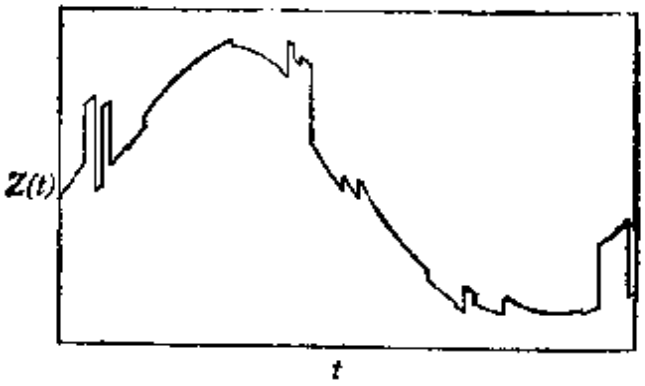


图 3.3　仅有漂移和跳跃的马尔可夫过程的样本路径

路径，由附有非零漂移的扩散过程所对应的曲线段叠加上涨落部分所构成。

$A(x, t)$ 非零，但 $\underset{\sim}{B}(x, t)$ 为零也是可能的，这时样本路径由叠加上间断的光滑曲线段[(3.5.14)的解]组成，如图 3.3 所示。这很像人们在稀薄气体中会预期的图景，其中粒子在反复碰撞之间自由运动，这种碰撞引起动量的瞬时变化，而不是位置的变化。

3.6　初始时刻的时间发展方程——后向方程

比 3.4 节简单得多，我们可以推导一些方程，它们给出关于 $p(x, t|y, t')$ 的初始变量 y，t' 的时间发展。

考虑

$$\lim_{\Delta t'\to 0}\frac{1}{\Delta t'}[p(x,\ t|y,\ t'+\Delta t')-p(x,\ t|y,\ t) \tag{3.6.1}$$

$$=\lim_{\Delta t'\to 0}\frac{1}{\Delta t'}\int dz\, p(z,t'+\Delta t'|y,\ t')[p(x,\ t|y,\ t'+\Delta t')$$

$$-p(x,\ t|z,\ t'+\Delta t')] \tag{3.6.2}$$

其中第二项用到了查普曼-柯尔莫哥洛夫方程，并注意第一项给出了$1\times p(x,\ t|y,\ t'+\Delta t')$。

现在必须假设所有有关的导数都存在，且$p(x,\ t|y,\ t')$对某个区域$t-t'>\delta>0$关于x、t、t'为连续有界，那末我们可以继续写出

$$=\lim_{\Delta t'\to 0}\frac{1}{\Delta t'}\int dz\, p(z,\ t'+\Delta t'|y,\ t')$$

$$\times[p(x,\ t|y,\ t')-p(x,\ t|z,\ t')] \tag{3.6.3}$$

应用类似于3.4.1节中的技巧，最终导得

$$\frac{\partial p(x,\ t|y,\ t')}{\partial t'}=-\sum_i A_i(y,\ t')\frac{\partial p(x,\ t|y,\ t')}{\partial y_i}$$

$$-\frac{1}{2}\sum_{ij}B_{ij}(y,\ t')\frac{\partial^2 p(x,\ t|y,\ t')}{\partial y_i\partial y_j}$$

$$+\int dz W(z|y,\ t')[p(x,\ t|y,\ t')$$

$$-p(x,\ t|z,\ t')] \tag{3.6.4}$$

这个方程称为后向微分查普曼-柯尔莫哥洛夫方程。在数学意义上，它比对应的前向方程(3.4.22)有着更好处理的定义。前向和后向两个方程的合适的初始条件都是对所有的t有

$$p(x,\ t|y,\ t)=\delta(x-y) \tag{3.6.5}$$

这个初始条件描述了一个显然的事实：如果在时刻t粒子位于y，则在同样时刻在x处找到它的概率密度是$\delta(x-y)$。

前向和后向方程是相互等价的。因为前向方程的满足初始条件(3.6.5)[或(3 .5.3)] 与适当的边界条件的解，如3.4.2节所述，产生查普曼-柯尔莫哥洛夫方程的解。但我们刚刚证明了这些恰好产生出后向方程。(对福克-普朗克方程在适当的边界条件之间的关系在5.2.1，4节中加以处理)。基本差别在于哪一组变

量保持固定．在前向方程的场合，我们使(y, t')固定不变，并存在$t \geqslant t'$的解，所以(3.6.5)是前向方程的初始条件．对于后向方程，存在$t' \leqslant t$的解．由于后向方程表示关于t'的发展，在这种情况下(3.6.5)式最好称为终端条件．

因为它们是等价的，所以前向和后向方程都是有用的．前向方程更直接地给出作为观测时间t的函数的可测量的量值，在应用中往往用得更普遍．后向方程则在首次通过时间或出口问题的研究中用得最多，在这类研究中我们寻求一个粒子在给定时间内离开某个区域的概率．

3.7 平稳齐次马尔可夫过程

在1.4.3节中我们遇到过平稳过程的概念，它表示一个已处于稳定状态的系统的随机运动，其随机性质与进行测量的时间是相互独立的．平稳性可在多种不同的程度上下定义，但我们将把术语"平稳过程"专用于一个严格的定义，即是，如果随机过程$\boldsymbol{X}(t)$与过程$\boldsymbol{X}(t+\varepsilon)$对任何$\varepsilon$有相同的统计特性，则$\boldsymbol{X}(t)$是平稳的．这个定义等价于下面的说法：所有的联合概率密度满足时间平移不变性，就是

$$p(x_1, t_1; x_2, t_2; x_3, t_3; \cdots; x_n, t_n) = p(x_1, t_1+\varepsilon; x_2, t_2+\varepsilon; x_3, t_3+\varepsilon; \cdots; x_n, t_n+\varepsilon) \tag{3.7.1}$$

因此这些概率仅是时间差$t_i - t_j$的函数．特别，单个时间的概率与时间独立，可以简单地写成

$$p_s(x) \tag{3.7.2}$$

而双时间的联合概率可写成

$$p_s(x, t_1 - t_2; x_2, 0) \tag{3.7.3}$$

最后，条件概率也可写成

$$p_s(x_1, t_1 - t_2 | x_2, 0) \tag{3.7.4}$$

对于马尔可夫过程，因为所有的联合概率可以写成双时间条件概率与单时间概率的乘积，所以平稳性的一个充分必要条件是能够

将单和双时间概率写成(3.7.2, 3)所给出的形式。

3.7.1 遍历性质

如果我们有一个平稳过程，则能合乎情理地预期，平均值的测量可以取变量 x 按时间顺序的值来构成，并对它们的各种函数取平均。这是一个关于大数法则(如 2.5.2 节所述)的行之有效的信念：它可用于随机过程中顺次测量所定义的变量。

定义变量 $\overline{X}(T)$ 为

$$\overline{X}(T)=\frac{1}{2T}\int_{-T}^{T}\mathrm{d}t\,x(t) \tag{3.7.5}$$

这里 $x(t)$ 是一个平稳过程，并考虑 $T\to\infty$ 时的极限。它表示了在全部时间内作平均测量均值的一个可能的模型。显然

$$\langle\overline{X}(T)\rangle=\langle x\rangle_{\mathrm{S}} \tag{3.7.6}$$

现在计算 $\overline{X}(T)$ 的方差，这应是

$$\langle\overline{X}(T)^2\rangle=\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T}\mathrm{d}t_1\mathrm{d}t_2\langle x(t_1)x(t_2)\rangle \tag{3.7.7}$$

又若过程是平稳的，

$$\langle x(t_1)x(t_2)\rangle=R(t_1-t_2)+\langle x\rangle^2 \tag{3.7.8}$$

这里 R 是双时间相关函数，因此

$$\langle\overline{X}(T)^2\rangle-\langle x\rangle^2=\frac{1}{4T^2}\int_{-2T}^{2T}\mathrm{d}\tau R(\tau)(2T-|\tau|) \tag{3.7.8}^{(*)}$$

其中最后一个因子来自变量代换

$$\tau=t_1-t_2$$

$$t=t_1 \tag{3.7.9}$$

并关于 t 取积分。

现在等式左端已是 $\overline{X}(T)$ 的方差，可以证明，在某些条件下，当 $T\to\infty$ 时它趋向于零。最直截了当的要求是

$$\lim_{T\to\infty}\frac{1}{T}\int_{-2T}^{2T}\mathrm{d}\tau\left(1-\frac{|\tau|}{2T}\right)R(\tau)=0 \tag{3.7.10}$$

不过它有点含糊不清。但是

(*) 算式编号重复，原文如此。——译注

$$\int_0^\infty \mathrm{d}\tau\,|R(\tau)|<\infty \tag{3.7.11}$$

显然是使极限为零的一个充分条件，这就要求当 $|t_1-t_2|\to\infty$ 时相关函数 $\langle x(t_1), x(t_2)\rangle$ 足够迅速地趋向于零。在令人感兴趣的实例中 $R(\tau)$ 常常有渐近状态

$$R(\tau)\sim \mathrm{Re}\{A\ \exp(-\tau/\tau_c)\} \tag{3.7.12}$$

其中 τ_c 是一个(可能为复的)参数，称为相关时间。显然判别准则(3.7.11)是满足的，此时 $\overline{X}(T)$ 的方差趋向于零，应用(3.7.6)和(2.9.4)我们可以写出

$$\underset{T\to\infty}{\mathrm{ms}-\lim}\ \overline{X}(T)=\langle x\rangle. \tag{3.7.13}$$

这意味着求平均的步骤(3.7.5)确是合理的。不难把这个结论推广到离散场合，对离散时间 $t_n=t_0+n\Delta t$ 处的测量值的无限集合取平均。

另外的遍历假设是易於阐述的，最值得关心的两个量是自相关函数和分布函数。在1.4.2节中早已指出，测算相关函数的最自然的方法是先确定

$$G(\tau,\ T)=\frac{1}{T}\int_0^T \mathrm{d}t\ x(t)x(t+\tau) \tag{3.7.14}$$

然后可颇为容易地用同样的推理方法证明

$$\underset{T\to\infty}{\mathrm{ms}-\lim}\ G(\tau,\ T)=\langle x(t)x(t+\tau)\rangle. \tag{3.7.15}$$

但要满足下述条件，即是说，定义 $\rho(\tau,\ \lambda)$ 为

$$\begin{aligned}&\langle x(t+\lambda+\tau)x(t+\lambda)x(t+\tau)x(t)\rangle.\\ &\quad=\rho(\tau,\ \lambda)+\langle x(t+\tau)x(\tau)\rangle^2.\end{aligned} \tag{3.7.16}$$

则要求

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-2T}^{2T}\left(1-\frac{|\lambda|}{2T}\right)\rho(\tau,\ \lambda)\mathrm{d}\lambda=0 \tag{3.7.17}$$

我们可以看清，这意味着对充分大的 λ，四时间平均(3.7.16)分解因式成为双时间平均的乘积，其"误差项" $\rho(\tau,\ \lambda)$ 当 $\lambda\to\infty$ 时足够迅速地趋向于零。如(3.7.12)所给出的那种指数渐近行为就是充分条件，而且是常见的。

相仿地，对于1.4节那样由傅里叶变换

$$S(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-i\omega\tau}G(\tau)d\tau \tag{3.7.18}$$

所给出的谱，也可由

$$S(\omega)=\lim_{T\to\infty}\frac{1}{2\pi T}\left|\int_0^T dt\ e^{-i\omega t}x(t)\right|^2 \tag{3.7.19}$$

这样的步骤求得。

最后，测算分布函数的实用方法是考虑一个区间 (x_1, x_2)，并反复测量 $x(t)$ 以决定它是否落在该范围内，这给出 $\int_{x_1}^{x_2}dx\ p_s(x)$ 的一个测量值。实质上，我们是在测量由

$$\begin{aligned}\chi(x)&=1 \quad x_1<x<x_2\\&=0 \quad \text{其他}\end{aligned} \tag{3.7.20}$$

所定义的函数 $\chi(x)$ 的时间平均值，改动一些〈x〉的遍历性的证明方法可得知，在满足

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-2T}^{T}d\tau\left(1-\frac{|\tau|}{2T}\right)\int_{x_1}^{x_2}dx'p_s(x')\times\left\{\int_{x_1}^{x_2}dx[p(x,\ \tau|x',\ 0)-p_s(x)]\right\}=0 \tag{3.7.21}$$

时分布是遍历的。这里最显而易见的充分条件是

$$\lim_{\tau\to\infty}p(x,\ \tau|x',\ 0)=p_s(x) \tag{3.7.22}$$

而且这个极限过程足够迅速地达到。实际上指数地趋向于极限是常见的，如同均值的情形一样，是足够地快的。

事实上这个条件对马尔可夫过程的均值和自相关函数的遍历性是充分的，因为所有的均值可以用条件概率表达，极限(3.7.22)能足够迅速地达到的事实易于看出对(3.7.17)和(3.7.10)是充分的。如果这个相当强的条件满足，我们就称马尔可夫过程为"遍历的"。

3.7.2 齐次过程

假若一个平稳马尔可夫过程满足条件 (3.7.22)，那末我们有

一个方法从平稳马尔可夫过程去构成一个非平稳过程，当时间趋向于无穷大时其极限是一个平稳过程。我们只要对

$$t,\ t' > t_0 \tag{3.7.23}$$

用

$$p(x,\ t) = p_s(x,\ t\,|\,x_0,\ t_0) \tag{3.7.24}$$

和

$$p(x,\ t\,|\,x',\ t') = p_s(x,\ t\,|\,x',\ t') \tag{3.7.25}$$

以及从它们用马尔可夫过程的常规方法所得到的所有其他的联合概率来定义这个非平稳过程。显然，如果(3.7.22)满足，我们发现当 $t\to\infty$ 或 $t_0\to-\infty$ 时

$$p(x,\ t)\to p_s(x)$$

所有其他的概率也成为平稳的，因为条件概率是平稳的。这种过程称为齐次过程。

物理解释是相当明白的。我们有一个随机系统，在时刻 t_0 它的变量 x 因某种外界作用而固定取值 x_0，然后随时间的推移经演变而返回成为一个平稳系统，这正是现实世界中许多平稳系统所形成的方式。

从微分查普曼-柯尔莫哥洛夫方程的观点来看，我们发现平稳分布函数 $p_s(x)$ 是平稳微分查普曼-柯尔莫哥洛夫方程的解。该方程具有

$$\begin{aligned} 0 = &-\sum_i \frac{\partial}{\partial z_i}[A_i(z)p(z,t\,|\,y,t')] \\ &+\frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial z_i\partial z_j}[B_{ij}(z)p(z,\ t\,|\,y,\ t')] \\ &+\int dx[W(z\,|\,x)p(x,\ t\,|\,y,\ t') \\ &-W(x\,|\,z)p(z,\ t\,|\,y,\ t')] \end{aligned} \tag{3.7.26}$$

的形式。这里我们已用到了过程为齐次的事实，它指明了(3.4.1～3)中定义的 A，B 和 W 是与 t 无关的。这是齐次过程的又一定义。

3.7.3 趋向于平稳过程

相反的问题也存在。假设 $\boldsymbol{A}$、$\boldsymbol{B}$和W 与时间无关，又有 $p_s(\boldsymbol{z})$ 满足(3.7.26)，那末在什么条件下微分查普曼-柯尔莫哥洛夫方程的解确实趋向于平稳解 $p_s(\boldsymbol{z})$？

对此问题看来并无完整的解答，但是我们可给出一个相当好的描述如下。定义微分查普曼-柯尔莫哥洛夫方程的任何两个解 p_1 和 p_2 的李雅普诺夫泛函

$$K=\int \mathrm{d}\boldsymbol{x}\, p_1(\boldsymbol{x},t)\log[p_1(\boldsymbol{x},t)/p_2(\boldsymbol{x},t)] \tag{3.7.27}$$

并暂时假定 p_1 和 p_2 都不取零值，这时可以证明K总是为正而 $\mathrm{d}K/\mathrm{d}t$ 总是为负。

首先，注意到 $p_2(\boldsymbol{x},t)$ 和 $p_1(\boldsymbol{x},t)$都已归一化，可以写成

$$K[p_1,p_2,t]=\int \mathrm{d}\boldsymbol{x}\, p_1(\boldsymbol{x},t)\{\log[p_1(\boldsymbol{x},t)/p_2(\boldsymbol{x},t)] + p_2(\boldsymbol{x},t)/p_1(\boldsymbol{x},t)-1\} \tag{3.7.28}$$

利用对于所有 $z>0$ 成立的不等式

$$-\log z+z-1\geqslant 0 \tag{3.7.29}$$

即证得 $K\geqslant 0$。

现在来证明 $\mathrm{d}K/\mathrm{d}t\leqslant 0$。我们有(已采用简化记号)

$$\frac{\mathrm{d}K}{\mathrm{d}t}=\int \mathrm{d}\boldsymbol{x}\left\{\frac{\partial p_1}{\partial t}[\log p_1+1-\log p_2]-\frac{\partial p_2}{\partial t}[p_1/p_2]\right\} \tag{3.7.30}$$

再一一计算微分查普曼-柯尔莫哥洛夫方程中漂移项、扩散项和跳跃项给予 $\mathrm{d}K/\mathrm{d}t$ 的贡献：

$$\left(\frac{\mathrm{d}K}{\mathrm{d}t}\right)_{\text{漂}}=\sum_i\int \mathrm{d}\boldsymbol{x}\left\{-[\log(p_1/p_2)+1]\frac{\partial}{\partial x_i}(A_i p_1)+(p_1/p_2)\frac{\partial}{\partial x_i}(A_i p_2)\right\} \tag{3.7.31}$$

经整理可给出

$$\left(\frac{\mathrm{d}K}{\mathrm{d}t}\right)_{\text{漂}}=\sum_i\int \mathrm{d}\boldsymbol{x}\,\frac{\partial}{\partial x_i}[-A_i p_i\log(p_1/p_2)] \tag{3.7.32}$$

相仿地，我们可以算出

$$\left(\frac{dK}{dt}\right)_{扩} = \frac{1}{2}\sum_{i,j}\int dx \left\{[\log(p_1/p_2)+1]\frac{\partial^2}{\partial x_i \partial x_j}(B_{ij}p_1) - (p_1/p_2)\frac{\partial^2}{\partial x_i \partial x_j}(B_{ij}p_2)\right\} \quad (3.7.33)^{(*)}$$

经整理后可写成

$$\left(\frac{dK}{dt}\right)_{扩} = -\frac{1}{2}\sum_{i,j}\int dx\, p_1 B_{ij}\left\{\frac{\partial}{\partial x_i}[\log(p_1/p_2)]\right\} \times\left\{\frac{\partial}{\partial x_j}[\log(p_1/p_2)]\right\} + \frac{1}{2}\sum_{i,j}\int dx \frac{\partial^2}{\partial x_i \partial x_j}[p_1 B_{ij}\log(p_1/p_2)] \quad (3.7.34)$$

最后，同样可以算出

$$\begin{aligned}\left(\frac{dK}{dt}\right)_{跳} = \int dx\, dx' \{&[W(x|x')p_1(x', t) \\ &- W(x'|x)p_1(x, t)] \\ &\times\{\log[p_1(x, t)/p_2(x, t)]+1\} \\ &-[W(x|x')p_2(x', t) \\ &- W(x'|x)p_2(x, t)]p_1(x, t)/p_2(x, t)\}\end{aligned} \quad (3.7.35)$$

经整理得

$$\left(\frac{dK}{dt}\right)_{跳} = \int dx dx'\, W(x|x')\{p_2(x', t) \times[\phi'\log[\phi/\phi'] - \phi + \phi']\} \quad (3.7.36)$$

这里

$$\phi = p_1(x, t)/p_2(x, t) \quad (3.7.37)$$

而 ϕ' 同样地由 x' 来定义。

现在我们考虑最简单的情况。假设存在一个除无穷远外处处

(*) 原文中此式等号右边的系数 $\frac{1}{2}$ 前有一负号，应去掉，——译注

非零的平稳解 $p_s(x)$，在无穷远处它自身及其一阶导数均成为零，那末我们选定 $p_2(x, t) = p_s(x)$。(3.7.32)与(3.7.34) 的第二项对 dK/dt 的贡献部分可以进行积分以给出曲面项，这些项在无穷远处为零，所以我们得到

$$\left(\frac{dK}{dt}\right)_{跳} = 0 \tag{3.7.38a}$$

$$\left(\frac{dK}{dt}\right)_{扩} \leqslant 0 \tag{3.7.38b}$$

$$\left(\frac{dK}{dt}\right)_{跳} \leqslant 0 \tag{3.7.38c}$$

上面最后一个不等式可在(3.7.29)中置 $z = \phi\phi'$ 而得到。

现在必须考虑(3.7.38)中在什么情况下实际上成立等式。细察(3.7.36)表明，当且仅当对几乎所有的使 $W(x|x') \neq 0$ 的 x 和 x' 有 $\phi = \phi'$ 时该项才为零。因此，如果$W(x|x')$恒不为零，即是在任何一对状态间的双向转移都可能发生的话，则跳跃部分的作用为零意味着 $\phi(x)$与 x 无关，也就是

$$p_1(x, t)/p_s(x) = 常数 \tag{3.7.39}$$

这个常数必定等于 1，因为 $p_1(x, t)$ 和 $p_s(x)$ 都是归一化的。

如果 B_{ij} 是几乎处处正定的话，扩散项将严格为负，因此，我们已经证实，在相当强的条件下，就是

$$\begin{aligned} &以概率1 \quad p_s(x) \neq 0 \\ &以概率1 \quad W(x|x') \neq 0 \\ &以概率1 \quad B_{ij}(x)为正定 \end{aligned} \tag{3.7.40}$$

时，任何微分查普曼-柯尔莫哥洛夫方程的解当 $t \to \infty$ 时趋向于平稳解 $p_s(x)$。

在下面两类基本系统中，这个结论未能成立。

a) 不连通状态空间

最好是在 A_i 和 B_{ij} 为零时阐明结论，因而我们有一个纯粹跳跃系统。假设空间分成两个区域 R_1 和 R_2，从 R_1 到 R_2 的正反转移都是不可能的，故若 x 和 x' 不是都在 R_1 中或 R_2 中，则 $W(x|x') = 0$，因此，如果

$$p_1(x,\ t) = \lambda_1 p_s(x) \qquad x \in R_1$$
$$= \lambda_2 p_s(x) \qquad x \in R_2 \tag{3.7.41}$$

则 $dK/dt = 0$ 是可能的，以致存在不是唯一的平稳分布。两个区域互不连通，各有单独的随机过程发生，并各有一个唯一的平稳解。存在于 R_1 或 R_2 的相对概率不因过程而改变。

一般，如果 B_{ij} 和 A_i 在 R_1 和 R_2 之间的边界上为零，相似的结论也是成立的。

b) $p_s(x)$ 在某个确定区域上为零

如果我们有

$$p_s(x) = 0 \qquad x \in R_1$$
$$\neq 0 \qquad x \in R \tag{3.7.42}$$

并再次令 A_i 和 B_{ij} 为零。因 $p_s(x)$ 满足平稳方程(3.7.26)，可得

$$W(x|y) = 0 \qquad x \in R_1,\ y \in R_2 \tag{3.7.43}$$

换句话说，从平稳分布取正值的区域 R_2 到平稳分布为零的区域 R_1 的转移是不可能的。进一步，其逆显然也是正确的，亦即(3.7.43)蕴涵(3.7.42)。

用记号 y、y' 等表示 R_1 中的点，x、x' 表示 R_2 中的点，并用记号 $r(y)$ 表示 $p(y)$，那末我们可以将运动方程写成

$$\frac{\partial p(x)}{\partial t} = \int dx' [W(x|x')p(x') - W(x'|x)p(x)]$$
$$- \int dy' W(x|y')r(y') \tag{3.7.43}$$(*)

$$\frac{\partial r(y)}{\partial t} = \int dy' [W(y|y')r(y') - W(y'|y)r(y)]$$
$$- r(y) \int dx' W(x'|y) \tag{3.7.44}$$

为方便起见，定义

$$\mu(y) = \int dx' W(x'|y) \tag{3.7.45}$$

则(3.7.44)的解由

(*) 标号重复，原文如此。——译注

$$r(\boldsymbol{y}, t) = \exp[-\mu(\boldsymbol{y})t]q(\boldsymbol{y}, t) \tag{3.7.46}$$

所给出，其中 $q(\boldsymbol{y}, t)$ 是仅关于区域 R_1 的主方程

$$\frac{\partial q(\boldsymbol{y})}{\partial t} = \int d\boldsymbol{y}'[W(\boldsymbol{y}|\boldsymbol{y}')q(\boldsymbol{y}') - W(\boldsymbol{y}'|\boldsymbol{y})q(\boldsymbol{y})] \tag{3.7.47}$$

的解。因此产生了趋向于平衡的结果。首先，$\mu(\boldsymbol{y})>0$，所以当 $t\to\infty$ 时 $r(\boldsymbol{y}, t)\to 0$，於是(3.7.43)中的非齐次项在时间很大时趋于消失，给我们留下一个通常的主方程，把(a)的推理过程用于这个方程，不出所料，解具有(3.7.42)的形式。

3.7.4 马尔可夫过程的自相关函数

对任何马尔可夫过程，我们可以写出自相关函数的一个非常精致的公式。定义

$$\langle \boldsymbol{X}(t)|[\boldsymbol{x}_0, t_0]\rangle = \int d\boldsymbol{x}\, \boldsymbol{x}\, p(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0) \tag{3.7.48}$$

以及自相关矩阵

$$\langle \boldsymbol{X}(t)\boldsymbol{X}(t_0)^T\rangle = \int d\boldsymbol{x}\, d\boldsymbol{x}_0\, \boldsymbol{x}\, \boldsymbol{x}_0^T\, p(\boldsymbol{x}, t; \boldsymbol{x}_0, t_0) \tag{3.7.49}$$

$$= \int d\boldsymbol{x}_0 \langle \boldsymbol{X}(t)|[\boldsymbol{x}_0, t_0]\rangle \boldsymbol{x}_0^T p(\boldsymbol{x}_0, t_0) \tag{3.7.50}$$

可见(3.7.48)是在 t_0 时刻 $\boldsymbol{X}$ 具有值 $\boldsymbol{x}_0$ 的条件下确定了 $\boldsymbol{X}(t)$ 的均值，而(3.7.50)告诉我们将这个条件平均(乘以 $\boldsymbol{x}_0^T$)在时刻 t_0 取平均就得到自相关矩阵。按定义，这些结论对任何随机过程都正确。

然而，在马尔可夫过程的场合我们有着确定整个过程的唯一的条件概率，这样，对于马尔可夫过程我们可以称 $\langle \boldsymbol{X}(t)|[\boldsymbol{x}_0, t_0]\rangle$ 是一个唯一确定的量，因为在时刻 t_0 的 $\boldsymbol{x}_0$ 的知识完全确定了过程的未来。该性质的最令人注意的应用在于平稳自相关函数的计算。为说明这个唯一性是多么重要，让我们考虑具有联合概率

$$p_s(\boldsymbol{x}_1, t_1; \boldsymbol{x}_2, t_2; \cdots \boldsymbol{x}_n, t_n) \tag{3.7.51}$$

的非马尔可夫平稳过程，当然，它只依赖于时间差。现在，只选择

在时刻 $t=0$ 通过点 $x=a$ 的样本路径，以建立一个对应的非平稳过程。如此，我们定义

$$p_a(x_1,\ t_1;\ x_2,\ t_2;\ \cdots;\ x_n,\ t_n)$$
$$=p_s(x_1,\ t_1;\ x_2,\ t_2;\ \cdots;\ x_n,\ t_n|a,\ 0) \tag{3.7.52}$$

对此过程我们注意

$$\langle \boldsymbol{X}(t)\,|\,[x_0,\ t_0]\rangle_a=\int dx\, x p_s(x,\ t|x_0,\ t_0;\ a,\ 0) \tag{3.7.53}$$

它包含着对 a 的依赖，用取平均的尖括弧旁边的下标 a 来表示。如果原来的平稳过程具有适当的遍历性质，则有

$$\lim_{\tau\to\infty} p_s(x,\ t+\tau|x_0,\ t_0+\tau;\ a,\ 0)$$
$$=p_s(x,\ t-t_0|x_0,\ 0) \tag{3.7.54}$$

所以我们也有 x 的平稳条件平均

$$\langle \boldsymbol{X}(t)\,|\,[x_0,\ t_0]\rangle_s=\lim_{\tau\to\infty}\langle \boldsymbol{X}(t+\tau)\,|\,[x_0,\ t_0+\tau]\rangle_a \tag{3.7.55}$$

以及由

$$\langle \boldsymbol{X}(t)\boldsymbol{X}(t_0)^T\rangle_s=\int dx_0\, x_0^T\langle \boldsymbol{X}(t)\,|\,[x_0,\ t_0]\rangle_s p_s(x_0)\ (*) \tag{3.7.56}$$

$$=\lim_{\tau\to\infty}\langle \boldsymbol{X}(t+\tau)\boldsymbol{X}(t_0+\tau)^T\rangle_a$$
$$=\lim_{\tau\to\infty}\int dx_0\, x_0^T\langle x(t+\tau)\,|\,[x_0,\ t_0+\tau]\rangle_a$$
$$\times p_a(x_0,\ t_0+\tau) \tag{3.7.57}$$

所给出的平稳自相关矩阵。但当过程是马尔可夫过程时，这些繁复的极限过程是不必要的，因为

$$\text{马尔可夫性质}\Longrightarrow\langle \boldsymbol{X}(t)\,|\,[x_t,\ t_0]\rangle_s=\langle \boldsymbol{X}(t)\,|\,[x_0,\ t_0]\rangle_a$$
$$=\langle \boldsymbol{X}(t)\,|\,[x_0,\ t_0]\rangle \tag{3.7.58}$$

方程(3.7.50)是应用于马尔可夫过程的一个回归定理，它是对于

(*) 等式右边应为 $\int dx_0\langle \boldsymbol{X}(t)\,|\,[x_0,\ t_0]\rangle_s x_0^T p_s(x_0)$，(3.7.57) 式有同样的毛病。——译注

线性系统的更强的回归定理的基础。关于线性系统我们是指对均值有着线性运动方程

$$\mathrm{d}\langle \boldsymbol{X}(t) \mid [\boldsymbol{x}_0, t_0] \rangle / \mathrm{d}t = -A\langle \boldsymbol{X}(t) \mid [\boldsymbol{x}_0, t_0] \rangle \tag{3.7.59}$$

它在有实际兴趣的系统中是很常见的，可作为一个精确结论或者作为一个近似。(3.7.59)的初始条件显然是

$$\langle \boldsymbol{X}(t_0) \mid [\boldsymbol{x}_0, t_0] \rangle = \boldsymbol{x}_0 \tag{3.7.60}$$

因此由(3.7.50, 59)

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \boldsymbol{X}(t)\boldsymbol{X}(t_0)^T \rangle = -A\langle \boldsymbol{X}(t)\boldsymbol{X}(t_0)^T \rangle \tag{3.7.61}$$

并带有初始条件$\langle \boldsymbol{X}(t_0)\boldsymbol{X}(t_0)^T \rangle$。时间相关矩阵

$$\langle \boldsymbol{X}(t)\boldsymbol{X}(t_0)^T \rangle - \langle \boldsymbol{X}(t) \rangle \langle \boldsymbol{X}(t_0)^T \rangle = \langle \boldsymbol{X}(t), \boldsymbol{X}(t_0)^T \rangle \tag{3.7.62}$$

显然遵循同样的方程，具有时刻 t_0 处的协方差矩阵所给出的初始条件。在一个平稳系统中我们有这样的结论：如果 $G(t)$ 是平稳时间相关函数，σ 是平稳协方差矩阵，则有

$$\mathrm{d}G(t)/\mathrm{d}t = -AG(t) \tag{3.7.63}$$

和

$$G(0) = \sigma \tag{3.7.64}$$

或者

$$G(t) = \exp[-At]\sigma \tag{3.7.65}$$

这就是最简单形式的回归定理。我们再次强调它是对马尔可夫过程有效的，其中均值服从(3.7.59)那样的线性发展方程。

对于非马尔可夫过程没有什么简单的步骤，我们必须实施包含在(3.7.57)中的复杂的运算步骤。

3.8 马尔可夫过程的例子

在此我们介绍几种微分查普曼-柯尔莫哥洛夫方程的一些基本解答备考，它们在本书所有余下的章节中有着广泛的应用。

3.8.1 维纳过程

这种过程的取名是因为 N. 维纳(Wiener)对它作了详尽的研究。从本章的观点，它是 3.5.2 节中所讨论的福克-普朗克方程的解，在方程中仅有一个变量 $W(t)$，漂移项为零，且扩散系数是 1。这样，在此场合福克-普朗克方程是

$$\frac{\partial}{\partial t}\,p(w,\ t|w_0,\ t_0)=\frac{1}{2}\,\frac{\partial^2}{\partial w^2}\,p(w,\ t|w_0,\ t_0) \tag{3.8.1}$$

利用关于条件概率的初始条件

$$p(w,\ t_0|w_0,\ t_0)=\delta(w-w_0) \tag{3.8.2}$$

我们用特征函数

$$\phi(s,\ t)=\int \mathrm{d}w\ p(w,\ t|w_0,\ t_0)\exp(isw) \tag{3.8.3}$$

来求解(3.8.1)，它满足

$$\frac{\partial\phi}{\partial t}=-\frac{1}{2}\,s^2\phi \tag{3.8.4}$$

所以

$$\phi(s,\ t)=\exp\left[-\frac{1}{2}\,s^2(t-t_0)\right]\phi(s,\ t_0) \tag{3.8.5}$$

由(3.8.2)，初始条件是

$$\phi(s,\ t_0)=\exp(isw_0)$$

故有

$$\phi(s,\ t)=\exp\left[isw_0-\frac{1}{2}s^2(t-t_0)\right] \tag{3.8.6}$$

施行傅里叶逆变换，我们得到(3.8.1)的解为

$$p(w,\ t|w_0,\ t_0)$$
$$=[2\pi(t-t_0)]^{-1/2}\exp[-(w-w_0)^2/2(t-t_0)] \tag{3.8.7}$$

这表示一个高斯变量，并有

$$\langle W(t)\rangle=w_0 \tag{3.8.8}$$

$$\langle[W(t)-w_0]^2\rangle=t-t_0 \tag{3.8.9}$$

故得一个初始尖峰分布随时间而展平，如图 3.4 所示。

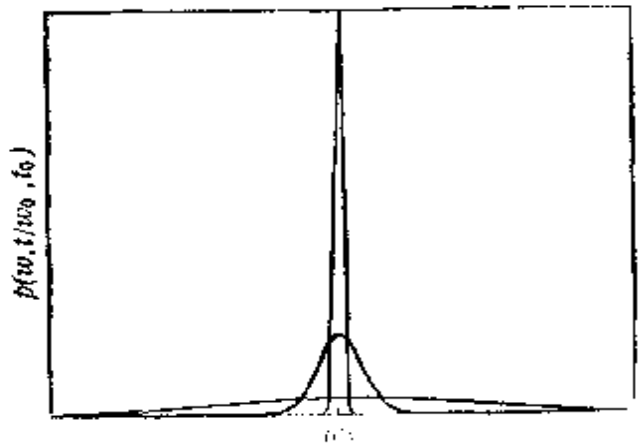


图 3.4　维纳过程：初始尖峰分布 $p(\omega, t|\omega_0, t_0)$ 随时间 $t-t_0$ 增加而展平

多元维纳过程可定义为

$$\boldsymbol{W}(t) = [W_1(t), W_2(t), \cdots, W_n(t)] \tag{3.8.10}$$

它满足多元福克-普朗克方程

$$\frac{\partial}{\partial t} p(\boldsymbol{w}, t|\boldsymbol{w}_0, t_0) = \frac{1}{2}\sum_i \frac{\partial^2}{\partial w_i^2} p(\boldsymbol{w}, t|\boldsymbol{w}_0, t_0) \tag{3.8.11}$$

其解为

$$p(\boldsymbol{w}, t|\boldsymbol{w}_0, t_0) = [2\pi(t-t_0)]^{-n/2}\exp[-(\boldsymbol{w}-\boldsymbol{w}_0)^2/2(t-t_0)] \tag{3.8.12}$$

表示一个多元高斯变量，具有

$$\langle \boldsymbol{W}(t)\rangle = \boldsymbol{w}_0 \tag{3.8.13}$$

以及

$$\langle [W_i(t) - w_{0i}][W_j(t) - w_{0j}]\rangle = (t-t_0)\delta_{ij} \tag{3.8.14}$$

一元维纳过程通常简称为布朗运动，因为维纳过程方程(3.8.1)同1.2节中指出的，由爱因斯坦证明的，布朗运动所遵循的扩散微分方程完全一样，但是用词并不一致。

关于维纳过程的若干要点是：

a) 样本路径的不规则性

虽然 $W(t)$ 的均值为零，但当 $t\to\infty$ 时均方成为无穷大，这意味着 $W(t)$ 的样本路径是非常变化多端的，事实上确是惊人地多变。图 3.5 中我们给出了几个具有相同初始点的不同的样本道

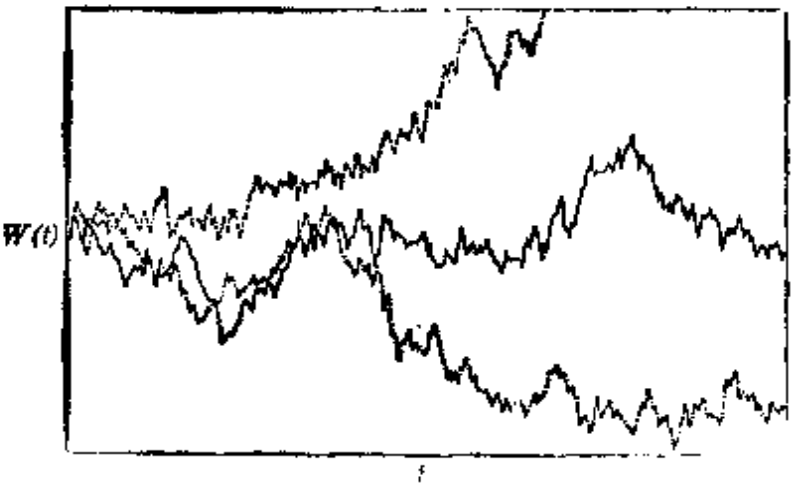


图 3.5　维纳过程的三个极其多变的模拟样本路径

路，用以说明路径的极端的不可再生性。

b) 样本路径的不可微性

维纳过程是一个扩散过程，因而 $W(t)$ 的样本路径是连续的，然而它们是不可微的。考虑

$$\text{Prob}\{|[W(t+h)-W(t)]/h|>k\} \tag{3.8.15}$$

由条件概率解，这个概率是

$$2\int_{kh}^{\infty} dw(2\pi h)^{-1/2}\exp(-w^2/2h) \tag{3.8.16}$$

它当 $h\to 0$ 时的极限为 1。这意味着不管 k 取什么值，$|[W(t+h)-W(t)]/h|$ 几乎必然要大于 k，就是说，在任何点的导数几乎必然是无穷大。这个论断与 3.5.2 节所给出的相似的直观说明和图 3.5 中逐点给出的模拟路径显然是相符的。自然，这对应着布朗粒子具有极端不规则运动的众所周知的实验事实。然而，这显然是一种理想化，因若 $W(t)$ 表示布朗粒子的位置，这就意味着其速度几乎必然是无穷大。奥恩斯坦-乌伦贝克过程是布朗运动的一个更真实的模型(3.8.4 节)。

c) 增量独立性

维纳过程是研究扩散过程的基础，借助于随机微分方程，我们可以用维纳过程将任何扩散过程表示出来。

$W(t)$ 的增量的统计独立性是特别重要的。更精确地说，

因为维纳过程是马尔可夫过程，联合概率密度可写成为

$$p(w_n, t_n; w_{n-1}, t_{n-1}; w_{n-2}, t_{n-2}; \cdots; w_0, t_0)$$
$$= \prod_{i=0}^{n-1} p(w_{i+1}, t_{i+1} | w_i, t_i)\, p(w_0, t_0) \tag{3.8.17}$$

再应用条件概率的显式(3.8.7)，可见

$$p(w_n, t_n; w_{n-1}, t_{n-1}; w_{n-2}, t_{n-2}; \cdots; w_0, t_0)$$
$$= \prod_{i=0}^{n-1} \{[2\pi(t_{i+1}-t_i)]^{-1/2}\exp[-(w_{i+1}-w_i)^2$$
$$+2(t_{i+1}-t_i)]\}\, p(w_0, t_0) \tag{3.8.18}$$

如果我们定义变量

$$\Delta W_i \equiv W(t_i) - W(t_{i-1}) \tag{3.8.19}$$

$$\Delta t_i \equiv t_i - t_{i-1} \tag{3.8.20}$$

则 ΔW_i 的联合概率密度为

$$p(\Delta w_n; \Delta w_{n-1}; \Delta w_{n-2}; \cdots; \Delta w_1; w_0)$$
$$= \prod_{i=1}^{n} \{(2\pi\Delta t_i)^{-1/2}\exp(-\Delta w_i^2/2\Delta t_i)\}\, p(w_0, t_0) \tag{3.8.21}$$

由 2.3.4 节所给出的统计独立性的定义，上式证实了变量 Δw_i 相互独立，并且同 $W(t_0)$ 独立。

在 4.2 节中将给出随机积分的定义，其中很重要的一个方面是具有独立增量 ΔW_i。

d) 自相关函数

一个令人极感兴趣的量是自相关函数，这在 1.4.2 节和 3.7.4 节中已讨论过，定义公式是

$$\langle W(t)W(s) | [w_0, t_0]\rangle$$
$$= \int dw_1 dw_2 w_1 w_2 p(w_1, t; w_2, s | w_0, t_0) \tag{3.8.22}$$

这是在初值为 $W(t_0) = w_0$ 的条件下 $W(t)$ 和 $W(s)$ 的乘积平均。我们还可看到，设 $t > s$，则有

$$\langle W(t)W(s) | [w_0, t_0]\rangle = \langle [W(t) - W(s)]W(s)\rangle$$
$$+ \langle [W(s)]^2\rangle \tag{3.8.23}$$

利用增量独立性，第一项平均为零，第二项由(3.8.9)给出，所以，

一般地有

$$\langle W(t)W(s)\,|\,[w_0,\ t_0]\rangle \cdots \min(t-t_0,\ s-t_0)+w_0 \qquad (3.8.24)$$

3.8.2 一维随机游动

这是一个很著名的、现已被视为经典的问题。一个人沿一直线以相等的概率随机地向左或向右跨步，步长为 l，所以他的位置只能取值 nl，其中 n 是整数。我们希望得知经过给定的时间后他到达离原点为 nl 距离的点的概率。

该问题可以用两种方式定义。第一种较为传统，允许走动者在时刻 $N\tau$(N 是整数）跨步，在这种时刻他必须以相等的概率向左或向右跨步。第二种允许他在单位时间内以概率 d 向左或向右跨步，这意味着他在每一点停留时间是有变动的。第二种方式由主方程来描述。

为了作出对该问题的主方程论述，我们认为单位时间内的转移概率有如下形式：

$$W(n+1\,|\,n,\ t)=W(n-1\,|\,n,\ t)=d \qquad (3.8.25)$$

否则 $W(n\,|\,m,\ t)=0$。所以按 3.5.1 节，该男子由 $n'l$ 出发而位于 nl 的主方程是

$$\partial_t P(n,\ t\,|\,n',\ t')=d[P(n+1,\ t\,|\,n',\ t')+P(n-1,\ t\,|\,n',\ t') -2P(n,\ t\,|\,n',\ t')] \qquad (3.8.26)$$

随机游动的较经典的形式并未假定那个男子按主方程来向左右跨跃，而是假定他在时刻 $N\tau$ 以相等的概率向左或向右跨跃，所以时间是离散变量。在此场合，我们可以写出

$$P(n,\ (N+1)\tau\,|\,n',\ N'\tau)=\frac{1}{2}[P(n+1,\ N\tau\,|\,n',\ N'\tau) +P(n-1,\ N\tau\,|\,n',\ N'\tau)] \qquad (3.8.27)$$

如果 τ 甚小，我们可利用在 $t=N\tau$，$t'=N'\tau$ 和 $d=\frac{1}{2}\tau^{-1}$ 时成立的

$$P(n,\ (N+1)\tau|n',\ N'\tau) \simeq P(n,\ N\tau|n',\ N'\tau) + \tau\partial_t P(n,\ t|n',\ t') \qquad (3.8.28)$$

来把(3.8.26, 27)视为彼此近似，所以在主方程模型中的单位时间转移概率对应于离散时间模型中等待时间倒数的一半。

引入特征函数

$$G(s,\ t) = \langle e^{isn}\rangle = \sum_n P(n,\ t|n',\ t')e^{isn} \qquad (3.8.29)$$

后两种系统都易于求解，这时主方程给出

$$\partial_t G(s,\ t) = d(e^{is} + e^{-is} - 2)G(s,\ t) \qquad (3.8.30)$$

而离散时间方程成为

$$G(s,\ (N+1)\tau) = \frac{1}{2}(e^{is} + e^{-is})G(s,\ N\tau) \qquad (3.8.31)$$

假设该男子在时刻 $t'=0$ 从原点 $n'=0$ 出发，两种情况下我们都有

$$G(s,\ 0) = 1 \qquad (3.8.32)$$

所以(3.8.30)的解是

$$G_1(s,\ t) = \exp[(e^{is} + e^{-is} - 2)td] \qquad (3.8.33)$$

而(3.8.31)的解是

$$G_2(s,\ N\tau) = \left[\frac{1}{2}(e^{is} + e^{-is})\right]^N \qquad (3.8.34)$$

它可写成为

$$G_2(s,\ t) = \left[1 + \frac{dt}{N}(e^{is} + e^{-is} - 2)\right]^N \qquad (3.8.35)$$

利用惯常的指数极限

$$\lim_{N\to\infty}\left(1 + \frac{a}{N}\right)^N = e^a \qquad (3.8.36)$$

在 s 充分小的时候(*)我们有

$$\lim_{N\to\infty} G_2(s,\ t) = G_1(s,\ t) \qquad (3.8.37)$$

由 2.6 节中特征函数的性质(V)，此式意味着概率分布之间的收敛性。

(*) "s 充分小"的条件是多余的。——译注

将 $G_1(s,\ t)$ 和 $G_2(s,\ N\tau)$ 按 $\exp(is)$ 的幂展开就可得到适当的概率分布，我们有

$$P_1(n,\ t|0,\ 0)=e^{-2td}I_n(4ld)\ (*) \tag{3.8.38}$$

$$P_2(n,\ N\tau|0,\ 0)=\left(\frac{1}{2}\right)^N N!\left[\left(\frac{N-n}{2}\right)!\left(\frac{N+n}{2}\right)!\right]^{-1} \tag{3.8.39}$$

离散时间分布也称为贝努里分布，它给出投掷均匀硬币 N 次得到 n 次正面的概率。

连续空间的极限也是令人感兴趣的。如果令移动距离为

$$x=nl \tag{3.8.40}$$

可得 x 的分布的特征函数是

$$\phi_1(s,\ t)=\langle e^{isx}\rangle=G_1(ls,\ t)=\exp[(e^{ils}+e^{-ils}-2)td] \tag{3.8.41}$$

无穷小步长 $l\to 0$ 时的极限是

$$\phi_1(s,\ t)\to\exp(-s^2tD) \tag{3.8.42}$$

其中

$$D=\lim_{l\to 0}(l^2d) \tag{3.8.43}$$

这是形为

$$p(x,\ t|0,\ 0)=(4\pi Dt)^{-1/2}\exp(-x^2/4Dt) \tag{3.8.44}$$

的高斯分布的特征函数(2.8.1 节)，当然也就是维纳过程 (3.8.1 节)或在 1.2 节中提及的布朗运动的分布。这样，维纳过程可看作是在无穷小步长的极限意义下连续时间随机游动的极限。

离散时间随机游动在

$$l\to 0,\ \tau\to 0\quad 以及\quad D=\lim_{l\to 0}(l^2/\tau)\ (**) \tag{3.8.45}$$

时的极限给出同样的结果。由上式显然可见 D 的表达式是单位时间内通过距离平方的平均。

(*) 等式右边应为 $e^{-2td}I_n(2td)$ 其中 I_n 为纯虚数变量场合导出的 n 阶贝塞尔函数。——译注

(**) 此式括弧中应为$(l^2/2\tau)$。——译注

我们也可以更直接地看到，将(3.8.26)的右端作为 x 的函数展开到 l 的两阶项，给出

$$\partial_t p(x, t|0, 0) = (l^2 d)\partial_x^2 p(x, t|0, 0) \tag{3.8.46}$$

这样，三个过程最终在两种程度上相互联系起来，即在所考虑的极限下，随机方程互相逼近，又在同样的极限下，这些方程的解互相逼近。这种极限正是爱因斯坦所用过的。跟 1.2 节比较一下表明，爱因斯坦用了离散时间离散空间随机游动来作出布朗运动的模型，不过他是用展开分布函数的时间发展方程这一方法导出了维纳过程模型。

本节的一些极限结论是爱因斯坦方法的较为严格的叙述。这些结论可推广到较少特殊性的情况，那是一个似乎可信的陈述：几乎任何跳过程以某种扩散过程为极限。然而，精确的极限并不总是这样简单，确有极限使某种跳过程成为确定性的并受刘维方程(3.5.3 节)所支配，而不是完整的福克-普朗克方程。这些结果将在7.2节中给出。

3.8.3 泊松过程

我们早已在 1.4.1 节中提及泊松过程，在电子到达阳极或顾客来到商店的过程中，单位时间内到达概率为 d，这种过程所遵循的主方程中应有

$$W(n+1|n, t) = d \tag{3.8.47}$$

其他情况下

$$W(n|m, t) = 0 \tag{3.8.48}$$

这个主方程成为

$$\partial_t P(n, t|n', t') = d[P(n-1, t|n', t') - P(n, t|n', t')] \tag{3.8.49}$$

跟(3.8.26)相比较，可见它还表示一个“单边”随机游动，其中走动者以单位时间内的概率 d 向右跨步。

特征函数方程是同(3.8.30)相仿的：

$$\partial_t G(s, t) = d[\exp(is) - 1]G(s, t) \tag{3.8.50}$$

它有解

$$G(s,\ t) = \exp\{td[\exp(is) - 1]\} \tag{3.8.51}$$

在时刻 $t=0$ 没有顾客(或电子)到达的初始条件下，可得出

$$P(n,\ t|0,\ 0) = \exp(-td)\,(td)^n/n! \tag{3.8.52}$$

这是以

$$\langle N(t)\rangle = td \tag{3.8.53}$$

为均值的泊松分布。与随机游动相对照，仅在

$$dl \equiv v \tag{3.8.54}$$

保持固定时存在 $l\to 0$ 的极限(*)，且极限特征函数是

$$\lim_{t\to 0}\{\exp[td(e^{ils}-1)]\} = \exp(ilvs) \tag{3.8.55}$$

其解为

$$p(x,\ t|0,\ 0) = \delta(x - vt) \tag{3.8.56}$$

在同样的取极限方式之下我们也发现会从主方程(3.8.49)得到刘维方程，它的解就是已经导出的确定性运动。

我们可以作出更精细的分析，将特征函数展开到 s 的二阶指数项，得出

$$\phi(s,\ t) = G(ls,\ t) \simeq \exp[t(ivs - s^2D/2)] \tag{3.8.57}$$

其中如前一节一样

$$D = l^2d$$

这是带有方差 Dt 和均值 vt 的高斯特征函数，所以现在有

$$p(x,\ t|0,\ 0) \simeq (2\pi Dt)^{-1/2}\exp[-(x-vt)^2/2Dt] \tag{3.8.58}$$

显然这个解也是

$$\partial_t p(x,\ t|0,\ 0) = -v\partial_x p(x,\ t|0,\ 0) + \frac{1}{2}D\partial_x^2 p(x,\ t|0,\ 0) \tag{3.8.59}$$

的解，这个方程由展开主方程(3.8.49)至 l^2 阶，即

$$P(n-1,\ t|0,\ 0) = dp(x-l,\ t|0,\ 0) \simeq dp(x,\ t|0,\ 0) - ld\partial_x p(x,\ t|0,\ 0)$$

(*) 这里有条件 $\lim_{l\to 0} dl = v$ 就可以了。——译注

$$+\frac{1}{2}\,l^2 d\partial_x p(x,\ t|0,\ 0) \qquad (3.8.60)$$

而得到。然而，这是一个逼近，或是一个展开，而不是一个极限。$l\to 0$ 时的极限给出具有纯粹确定性解 (3.8.56)的刘维方程。实际上，在适当定义 v 时的求极限过程 $l\to 0$ 对应着 $D=0$。上面提到的那种逼近是范坎彭的系统大小展开方法的特殊情况，我们将在 7.2.3 节中作出完整的论述。

3.8.4 奥恩斯坦-乌伦贝克过程

至今的例子全都没有平稳分布，即是，当 $t\to\infty$ 时在任何有限点处的分布趋向于零，由此可见，点以概率 1 移向无穷远处。

如果我们在维纳过程上加上线性漂移项，就得到

$$\partial_t p = \partial_x(kxp) + \frac{1}{2}\,D\partial_x p \qquad (3.8.61)$$

形式的福克-普朗克方程，其中 p 表示 $p(x,\ t|x_0,\ 0)$。这是奥恩斯坦-乌伦贝克过程[3.5]。特征函数

$$\phi(s) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}sx} p(x,\ t|x_0,\ 0)\mathrm{d}x \qquad (3.8.62)$$

所满足的方程为

$$\partial_t\phi + ks\,\partial_s\phi = -\frac{1}{2}\,Ds^2\phi \qquad (3.8.63)$$

可以用特征方法求解这个方程，即是，如果

$$u(s,\ t,\ \phi) = a \quad 与 \quad v(s,\ t,\ \phi) = b \qquad (3.8.64)$$

是辅助方程

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}s}{ks} = \frac{\mathrm{d}\phi}{-\frac{1}{2}\,Ds^2\phi} \qquad (3.8.65)$$

的两个积分(以 a 和 b 为任意常数)，则(3.8.63)的通解为

$$f(u,\ v) = 0$$

对于含有 $\mathrm{d}t$ 和 $\mathrm{d}s$ 的方程求积分，又对于含有 $\mathrm{d}s$ 和 $\mathrm{d}\phi$ 的方程求积分，易于求得两个特积分为

$$u(s,\ t,\ \phi) = s\exp(-kt) \qquad (3.8.66)$$

与

$$v(s,\ t,\ \phi) = \phi \exp(Ds^2/4k) \tag{3.8.67}$$

显然通解可以取 $v = g(u)$ 的形式，其中 $g(u)$ 是 u 的任意函数。这样，通解为

$$\phi(s,\ t) = \exp(-Ds^2/4k)\, g[s\exp(-kt)] \tag{3.8.68}$$

边界条件

$$p(x,\ 0|x_0,\ 0) = \delta(x - x_0) \tag{3.8.69}$$

显然要求

$$\phi(s,\ 0) = \exp(ix_0 s) \tag{3.8.70}$$

并给出

$$g(s) = \exp(Ds^2/4k + ix_0 s)$$

因此

$$\phi(s,\ t) = \exp\left[\frac{-Ds^2}{4k}(1 - e^{-2kt}) + isx_0 e^{-kt}\right] \tag{3.8.71}$$

由 2.8.1 节，它对应于一个高斯变量，具有

$$\langle X(t)\rangle = x_0 \exp(-kt) \tag{3.8.72}$$

$$\mathrm{var}\,\{X(t)\} = \frac{D}{2k}[1 - \exp(-2kt)] \tag{3.8.73}$$

显然当 $t \to \infty$ 时均值和方差分别趋向于极限 0 和 $D/2k$，这给出了一个极限平稳解。这个解也可以令 $\partial_t p = 0$ 而直接求得，于是 p 满足平稳福克-普朗克方程

$$\partial_x\left[kxp + \frac{1}{2} D\partial_x p\right] = 0 \tag{3.8.74}$$

积分一次，我们得到

$$\left[kxp + \frac{1}{2} D\partial_x p\right]_{-\infty}^{\infty} = 0 \tag{3.8.75}$$

在 $-\infty$ 处 p 连同它的导数为零的要求对于正规化是必须的。因此有

$$\frac{1}{p}\partial_x p = -\frac{2kx}{D} \tag{3.8.76}$$

故得

$$p_s(x) = (\pi D/k)^{-1/2}\exp(-kx^2/D) \tag{3.8.77}$$

这是以 0 为均值、$D/2k$ 为方差的高斯变量，正如依赖于时间的解所预期的那样。

显然，如果确实存在平稳解，它总是可对单变量系统用这样的积分方法求得。但若平稳解不存在，这个方法给出一个非正规化的解。

奥恩斯坦-乌伦贝克过程的时间相关函数，可以计算如同联系于维纳过程的那种时间相关函数，在绝大多数随机系统中这是一批可测量的数据。但是除了用定义

$$\langle X(t)X(s)\,|\,[x_0,\ t_0]\rangle$$

$$= \iint \mathrm{d}x_1\mathrm{d}x_2\, x_1x_2p(x_1,\ t;x_2,\ s\,|\,x_0,\ t_0) \tag{3.8.78}$$

以及由马尔可夫性质

$$= \iint \mathrm{d}x_1\mathrm{d}x_2\, x_1x_2p(x_1,\ t\,|\,x_2,\ s)\,p(x_2,\ s\,|\,x_0,\ t_0) \tag{3.8.79}$$

在

$$t \geqslant s \geqslant t_0 \tag{3.8.80}$$

的假设下成立之外，我们没有容易的方法去作计算。

通常，在一个确定的初始条件下的相关函数并不如平稳相关函数那样令人感兴趣，后者可让系统趋向于平稳分布而获得。这可以在遥远的过去时刻置初始条件而达到，如 3.7.2 节中指出的那样，令 $t_0 \to -\infty$，我们有

$$\lim_{t_0\to-\infty} p(x_2,\ s\,|\,x_0,\ t_0) = p_s(x_2) = (\pi D/k)^{-1/2}\exp(-kx_2^2/D) \tag{3.8.81}$$

再直接代入和积分，并注意到平稳均值是零，我们得到

$$\langle X(t)X(s)\rangle_s = \langle X(t),\ X(s)\rangle_s$$

$$= \frac{D}{2k}\exp(-k|t-s|) \tag{3.8.82}$$

这个结论展示了平稳过程的一般性质：相关函数仅依赖于时间差。还有一个一般结论是 [3.6]：本节所叙述的过程是仅有的

单变量高斯马尔可夫过程。

应用第4章中将要展开的随机微分方程的方法，本节的结论是很容易得到的。

奥恩斯坦-乌伦贝克过程是一个简单的，可作显式表示的具有平稳解的过程。在其平稳态下，常用它来作为实际噪声信号的模型，其中 $X(t)$ 和 $X(s)$ 仅在

$$|t-s|\sim 1/k\equiv\tau \tag{3.8.83}$$

时显著地相关。（更精确地说，对任何过程 $X(s)$ 而言称之为相关时间的 τ 定义为

$$\tau=\int_0^\infty \mathrm{d}t\langle X(t),\ X(0)\rangle_s/\mathrm{var}\{X\} \tag{3.8.84}$$

它与相关函数的精确形式无关。）

3.8.5 随机电报过程

我们考虑一个信号 $X(t)$，它可取 a 和 b 两个值之一，以确定的单位时间概率从一个转换成为另一个。这样就有主方程

$$\left.\begin{aligned}\partial_t P(a,\ t|x,\ t_0)&=-\lambda P(a,\ t|x,\ t_0)+\mu P(b,\ t|x,\ t_0)\\ \partial_t P(b,\ t|x,\ t_0)&=\lambda P(a,\ t|x,\ t_0)-\mu P(b,\ t|x,\ t_0)\end{aligned}\right\} \tag{3.8.85}$$

它的解易于求得，只要注意

$$P(a,\ t|x,t_0)+P(b,\ t|x,\ t_0)=1$$

又对 $\lambda P(a,\ t|x,\ t_0)-\mu P(b,\ t|x,\ t_0)$ 可以导出一个简单的方程，在初始条件

$$P(x',\ t_0|x,\ t_0)=\delta_{x,x'} \tag{3.8.86}$$

下，它的解是

$$\begin{aligned}&\lambda P(a,\ t|x,\ t_0)-\mu P(b,\ t|x,\ t_0)\\ &\quad=\exp[-(\lambda+\mu)(t-t_0)(\lambda\delta_{a,x}-\mu\delta_{b,x})\end{aligned} \tag{3.8.87}$$

所以

$$\left.\begin{aligned} P(a,\ t|x,\ t_0) &= \frac{\mu}{\lambda+\mu} + \exp[-(\lambda+\mu)(t-t_0)] \\ &\quad \times\left(\frac{\lambda}{\lambda+\mu}\delta_{a,x} - \frac{\mu}{\lambda+\mu}\delta_{b,x}\right) \\ P(b,\ t|x,\ t_0) &= \frac{\lambda}{\lambda+\mu} - \exp[-(\lambda+\mu)(t-t_0)] \\ &\quad \times\left(\frac{\lambda}{\lambda+\mu}\delta_{a,x} - \frac{\mu}{\lambda+\mu}\delta_{b,x}\right) \end{aligned}\right\} \tag{3.8.88}$$

显然在 $t_0 \to -\infty$ 时这个过程有平稳解

$$\left.\begin{aligned} P_s(a) &= \frac{\lambda}{\lambda+\mu} \\ P_s(b) &= \frac{\lambda}{\lambda+\mu} \end{aligned}\right\} \tag{3.8.89}$$

从主方程它们当然是显而易见的。

$X(t)$ 的均值和方差可直接算出：

$$\begin{aligned} \langle X(t)|[x_0,\ t_0]\rangle &= \sum x P(x,\ t|x_0,\ t_0) \\ &= \frac{a\mu+b\lambda}{\mu+\lambda} + \exp[-(\lambda+\mu)(t-t_0)] \\ &\quad \times\left(x_0 - \frac{a\mu+b\lambda}{\mu+\lambda}\right) \end{aligned} \tag{3.8.90}$$

所以

$$\langle X\rangle_s = \frac{a\mu+b\lambda}{\mu+\lambda} \tag{3.8.91}$$

可以算出方差，但它是一个非常繁复的式子。平稳方差则易于算出，应为

$$\mathrm{var}\{X\}_s = \frac{(a-b)^2\mu\lambda}{(\lambda+\mu)^2} \tag{3.8.92}$$

为了计算平稳时间相关函数，我们写出（设 $t \geqslant s$）

$$\langle X(t)X(s)\rangle_s = \sum_{xx'} xx' P(x,\ t|x's) P_s(x') \tag{3.8.93}$$

$$= \sum_{x'} x' \langle X(t)|[x',\ s]\rangle P_s(x') \tag{3.8.94}$$

现在用(3.8.90～92)可求得

$$\langle X(t)X(s)\rangle_s = \langle X\rangle_s + \exp[-(\lambda+\mu)(t-s)]$$

$$\times (\langle X^2\rangle_s - \langle X\rangle_s^2) \tag{3.8.95}$$

$$= \left(\frac{a\mu + b\lambda}{\mu + \lambda}\right)^2 + \exp[-(\lambda+\mu)(t-s)]$$

$$\times \frac{(a-b)^2\mu\lambda}{(\lambda+\mu)^2} \tag{3.8.96}$$

因此

$$\langle X(t),\ X(s)\rangle_s = \langle X(t)X(s)\rangle_s - \langle X\rangle_s^2$$

$$= \exp[-(\lambda+\mu)|t-s|] - \frac{(a-b)^2\mu\lambda}{(\lambda+\mu)^2} \tag{3.8.97}$$

注意这个时间相关函数同奥恩斯坦-乌伦贝克过程恰有相同的形式。当然高阶相关函数是不同的，但因这个简单的相关函数和双状态过程的简明性，随机电报方程也在构筑模型中得到了广泛的应用。

4. 伊藤分析和随机微分方程

4.1 动机

在1.2.2节中我们初次遇见了一种方程，它是现在称之为郎之万方程的原型，可以尝试表示为一个常微分方程，其中出现一个急速而无规则的时间涨落函数（即郎之万原先的方程中的$X(t)$项）。郎之万所给出的关于爱因斯坦的结论的推导是极为简明的，这种简明性本身就是一种充分的动力，试图将这种方程的概念置于合理而精确的基础之上。

思路简明的郎之万方程通常可以写成

$$\frac{\mathrm{d}x}{\mathrm{d}t} = a(x, t) + b(x, t)\xi(t) \tag{4.1.1}$$

的形式，其中x是所关心的变量，$a(x, t)$和$b(x, t)$是确定的已知函数，而$\xi(t)$是急速涨落的随机项。所谓"急速地变化，高度不规则函数"概念的理想化的数学形式是对$t \neq t'$，$\xi(t)$和$\xi(t')$是统计独立的。我们也要求$\langle \xi(t) \rangle = 0$，因为任何非零均值可被并入$a(x, t)$的定义中；这样一来就要求

$$\langle \xi(t)\xi(t') \rangle = \delta(t - t') \tag{4.1.2}$$

它满足在不同时刻不相关的要求，此外，它又有着$\xi(t)$的方差是无穷大的相当病态的结论。从实际的观点来看，我们知道没有什么量可以具有这样的无穷方差，但是白噪声的概念作为实际涨落信号的一个理想化确有某些意义，在1.4.2节中已指出过它同电路中的约翰逊噪声的联系。我们已经见到过两种可被认为是几乎不相关的实际形式的噪声源，即奥恩斯坦-乌伦贝克过程和随机电报信号。二者的2阶相关函数都有

$$\langle X(t),\ X(t')\rangle = \frac{\gamma}{2}\,\mathrm{e}^{-\gamma|t-t'|} \tag{4.1.3}$$

的形式，至多差一常数因子．它们之间的实质性差别在于随机电报信号的样本路径是不连续的，而奥恩斯坦-乌伦贝克过程则不然。如果(4.1.1)被视为真实的微分方程，其中$\xi(t)$不是以δ函数为相关的白噪声，确切地说是一个具有有限相关时间的噪声，则选择$\xi(t)$为连续函数看来是必要的，以使这一方程有现实意义，我们确实不希望$\mathrm{d}x/\mathrm{d}t$不连续地变化。相关函数(4.1.3)当$\gamma\to 0$时的极限显然是狄拉克δ函数，因为

$$\int_{-\infty}^{\infty}\frac{\gamma}{2}\,\mathrm{e}^{-\gamma|t-t'|}\mathrm{d}t' = 1 \tag{4.1.4}$$

又对$t\neq t'$

$$\lim_{\gamma\to\infty}\frac{\gamma}{2}\,\mathrm{e}^{-\gamma|t-t'|} = 0 \tag{4.1.5}$$

这意味着$\xi(t)$的一个可能的模型可由取$\gamma\to\infty$时奥恩斯坦-乌伦贝克过程的某种极限而得到。按3.8.4节的记号，应该对应于

$$k\to\infty \tag{4.1.6}$$

时的极限，其中$D=k^2$．这个极限绝不会存在。任何这种极限显然必须在计算可测量的量之后才可作出．这种程序是可能的，但是太繁琐以至无法作为运算工具来使用。

我们需要有别的途径。既然列出了微分方程(4.1.1)，我们必定希望它是可积分的，因此必定希望

$$u(t) = \int_0^t \mathrm{d}t'\xi(t') \tag{4.1.7}$$

存在。

假定我们现在要求有积分的通常性质：$u(t)$是t的连续函数。这推出$u(t)$是一个马尔可夫过程，因为我们可以写出

$$u(t') = \int_0^t \mathrm{d}s\xi(s) + \int_t^{t'}\mathrm{d}s\xi(s) \tag{4.1.8}$$

$$= \lim_{\varepsilon\to 0}\left[\int_0^{t-\varepsilon}\mathrm{d}s\xi(s)\right] + \int_t^{t'}\mathrm{d}s\xi(s) \tag{4.1.9}$$

对任何 $\varepsilon>0$，第一个积分中的 $\xi(s)$ 与第二个积分中的 $\xi(s)$ 相独立。因此，由连续性得知 $u(t)$ 跟 $u(t')-u(t)$ 是统计独立的，进而对所有 $t''<t$，$u(t')-u(t)$ 与 $u(t'')$ 独立。这意味着 $u(t')$ 由 $u(t)$ 值的知识所完全（概率地）确定，而与任何过去的值无关。因此 $u(t)$ 是一个马尔可夫过程。

因为 $u(t)$ 的样本函数是连续的，我们一定可以用福克-普朗克方程来描述 $u(t)$。对此过程可以用 3.5.2 节的公式来计算其漂移和扩散系数。可以写出

$$\langle u(t+\Delta t)-u_0 \mid [u_0,\ t]\rangle = \left\langle \int_t^{t+\Delta t} \xi(s)\mathrm{d}s \right\rangle = 0 \tag{4.1.10}$$

以及

$$\langle [u(t+\Delta t)-u_0]^2 \mid [u_0,\ t]\rangle = \int_t^{t+\Delta t} \mathrm{d}s \int_t^{t+\Delta t} \mathrm{d}s' \langle \xi(s)\xi(s')\rangle \tag{4.1.11}$$

$$= \int_t^{t+\Delta t} \mathrm{d}s \int_t^{t+\Delta t} \mathrm{d}s' \delta(s-s')$$

$$= \Delta t \tag{4.1.12}$$

所以漂移和扩散系数是

$$A(u_0,\ t) = \lim_{\Delta t\to 0} \frac{\langle u(t+\Delta t)-u_0 \mid [u_0,\ t]\rangle}{\Delta t} = 0 \tag{4.1.13}$$

$$B(u_0,\ t) = \lim_{\Delta t\to 0} \frac{\langle [u(t+\Delta t)-u_0]^2 \mid [u_0,\ t]\rangle}{\Delta t} = 1 \tag{4.1.14}$$

于是维纳过程的福克-普朗克方程是我们所要求的，可以写出

$$\int_0^t \xi(t')\mathrm{d}t' = u(t) = W(t) \tag{4.1.15}$$

这样，我们有了一个结论：$\xi(t)$ 的积分是 $W(t)$，但 $W(t)$ 自己却如 3.8.1 节所指出的那样是不可微的。这意味着用数学的语言，郎之万方程(4.1.1)不存在，然而相应的积分方程

$$x(t)-x(0) = \int_0^t a[x(s),\ s]\mathrm{d}s + \int_0^t b[x(s),\ s]\xi(s)\mathrm{d}s \tag{4.1.16}$$

能被无矛盾地解释。

我们来作一个代换，它直接来自将$\xi(t)$的积分解释为维纳过程$W(t)$，即

$$\mathrm{d}W(t) \equiv W(t+\mathrm{d}t) - W(t) = \xi(t)\mathrm{d}t \tag{4.1.17}$$

由此可将第二个积分写成为

$$\int_0^t b[x(s), s]\mathrm{d}W(s) \tag{4.1.18}$$

这是关于$W(t)$的样本函数的一种随机司蒂阶 (Stieltjes) 积分。这种积分可以作出定义，我们将在下一节中进行。

在作定义之前，应该指出$u(t)$为连续的要求尽管是很自然的，它还是可以放宽，产生一种方法将跳过程定义为随机微分方程。这一点早在1.4.1节中关于散粒噪声的论述中就有所暗示，然而它似乎并不太有用，本书中将不再述及，有兴趣的读者可以参考[4.1]。

最后我们应该指出，人们通常假设$\xi(t)$是高斯型的，同时满足条件(4.1.2)。前文中未有这个要求，事实上高斯性质是在假设的$u(t)$的连续性之后的性质。取什么样的假设，在严格意义上只是偏好的问题。但是$u(t)$的连续性似乎是比$\xi(t)$的高斯性质自然得多的假设，因为后者原则上涉及任意高阶的矩。

4.2 随机积分

4.2.1 随机积分的定义

假设$G(t)$是任一时间函数，$W(t)$是维纳过程，我们把随机积分

$$\int_{t_0}^{t} G(t')\mathrm{d}W(t')$$

定义为一种黎曼-司蒂阶积分。就是说，将区间$[t_0, t]$用分点

$$t_0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_{n-1} \leqslant t_n = t \tag{4.2.1}$$

分划成n个子区间(见图4.1)，并确定各中间点τ_i为

$$t_{i-1} \leqslant \tau_i \leqslant t_i \tag{4.2.2}$$

随机积分

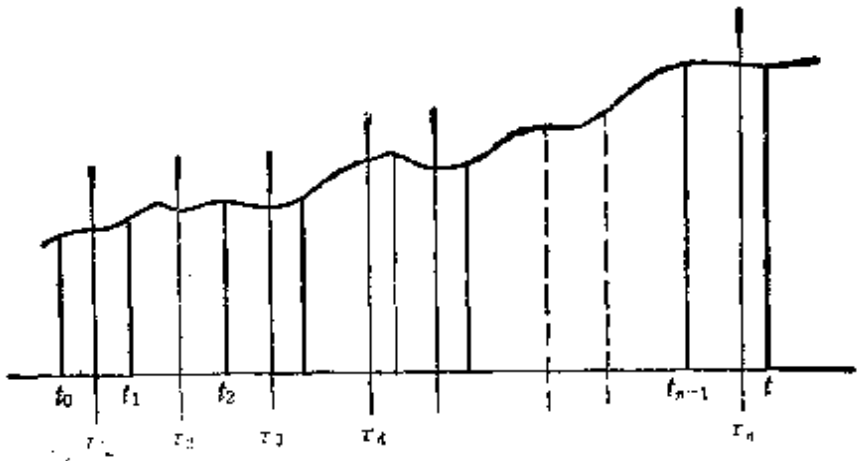


图 4.1　随机积分的定义中所用的时间区间的分割

$$\int_{t_0}^{t} G(t')\mathrm{d}W(t')$$

定义为部分和

$$S_n = \sum_{i=1}^{n} G(\tau_i)[W(t_i) - W(t_{i-1})] \tag{4.2.3}$$

的极限. 探索一下颇为容易地看到，定义为 S_n 的极限的积分一般与中间点 τ_i 的特定选择有关。如果选用 $G(\tau_i) = W(\tau_i)$，则

$$\langle S_n \rangle = \left\langle \sum_{i=1}^{n} W(\tau_i)[W(t_i) - W(t_{i-1})] \right\rangle \tag{4.2.4}$$

$$= \sum_{i=1}^{n} [\min(\tau_i, t_i) - \min(\tau_i, t_{i-1})] \tag{4.2.5}$$

$$= \sum_{i=1}^{n} (\tau_i - t_{i-1}) \tag{4.2.6}$$

如对所有的 i 选择

$$\tau_i = \alpha t_i + (1-\alpha)t_{i-1} \quad (0<\alpha<1) \tag{4.2.7}$$

则有

$$\langle S_n \rangle = \sum_{i=1}^{n} (t_i - t_{i-1})\alpha = (t - t_0)\alpha \tag{4.2.8}$$

这样一来积分的均值可以是 0 和 $(t-t_0)$ 之间的任何值，有赖于中间点的选择。

因此我们按 $\alpha = 0$ 的特性选择中间点，即是，选择

$$\tau_i = t_{i-1} \tag{4.2.9}$$

并由此定义函数 $G(t)$ 的伊藤随机积分为

$$\int_{t_0}^{t} G(t')\mathrm{d}W(t')$$

$$= \underset{n\to\infty}{\text{ms-lim}} \left\{ \sum_{i=1}^{n} G(t_{i-1})[W(t_i) - W(t_{i-1})] \right\} \tag{4.2.10}$$

按 2.9.2 节的定义，ms-lim 表示均方极限。

4.2.2 例子 $\int_{t_0}^{t} W(t')\mathrm{d}W(t')$

精确的计算是做得到的。我们可以写出[记 W_i 为 $W(t_i)$]

$$S_n = \sum_{i=1}^{n} W_{i-1}(W_i - W_{i-1}) = \sum_{i=1}^{n} W_{i-1}\Delta W_i \tag{4.2.12}$$

$$= \frac{1}{2}\sum_{i=1}^{n}[(W_{i-1} + \Delta W_i)^2 - (W_{i-1})^2 - (\Delta W_i)^2] \tag{4.2.13}$$

$$= \frac{1}{2}\,[W(t)^2 - W(t_0)^2] - \frac{1}{2}\sum_{i=1}^{n}(\Delta W_i)^2$$

可以算出最后一项的均方极限。注意到

$$\langle \Sigma(\Delta W_i)^2 \rangle = \sum_i \langle (W_i - W_{i-1})^2 \rangle$$

$$= \sum_i (t_i - t_{i-1}) = t - t_0 \tag{4.2.14}$$

由此

$$\langle [\sum_i (W_i - W_{i-1})^2 - (t - t_0)]^2 \rangle$$

$$= \langle \sum_i (W_i - W_{i-1})^4 + 2\sum_{i>j}(W_i - W_{i-1})^2(W_j - W_{j-1})^2$$

$$- 2(t - t_0)\sum_i (W_i - W_{i-1})^2 + (t - t_0)^2 \rangle \tag{4.2.15}$$

注意 $W_i - W_{i-1}$ 是高斯变量并与 $W_j - W_{j-1}$ 独立，故可分解因子，这样有

$$\langle (W_i - W_{i-1})^2 (W_j - W_{j-1})^2 \rangle = (t_i - t_{i-1})(t_j - t_{j-1}) \tag{4.2.16}$$

又有[用(2.8.6)式]

$$\langle (W_i - W_{i-1})^4 \rangle = 3\langle (W_i - W_{i-1})^2 \rangle^2 = 3(t_i - t_{i-1})^2 \tag{4.2.17}$$

与(4.2.16)合在一起给出

$$\langle[\sum_i(W_i-W_{i-1})^2-(t-t_0)]^2\rangle=2\sum_i(t_i-t_{i-1})^2$$
$$+\sum_{i,j}[(t_i-t_{i-1})-(t-t_0)][(t_j-t_{j-1})-(t-t_0)]$$
$$=2\sum_i(t_i-t_{i-1})^2 \qquad (4.2.18)$$
$$\to 0 \quad \text{当 } n\to\infty$$

这样

$$\underset{n\to\infty}{\text{ms-lim}}\sum_i(W_i-W_{i-1})^2=t-t_0 \qquad (4.2.19)$$

按均方极限的定义，所以有

$$\boxed{\int_{t_0}^{t}W(t')\mathrm{d}W(t')=\frac{1}{2}[W(t)^2-W(t_0)^2-(t-t_0)]} \qquad (4.2.20)$$

注解

i) $$\left\langle\int_{t_0}^{t}W(t')\mathrm{d}W(t')\right\rangle=\frac{1}{2}[\langle W(t)^2\rangle-\langle W(t_0)^2\rangle$$
$$-(t-t_0)]=0 \qquad (4.2.21)$$

按定义这也是显然的，由于对各别的项有$\langle W_{i-1}\Delta W_i\rangle$为零，这是因为如3.8.1节所证明的那样，$\Delta W_i$与$W_{i-1}$为统计独立。

ii) 积分结果与通常的黎曼-司蒂阶积分不再相同，那里$(t-t_0)$项不应出现。其理由是$|W(t+\Delta t)-W(t)|$几乎总是$\sqrt{t}$阶的，所以与通常的积分相比较，$\Delta W(t)$中二阶项在取极限时并不消失为零。

4.2.3 斯特拉托诺维奇积分

随机积分的另一种定义是由斯特拉托诺维奇(Stratonovich)[4.2]引入的，其中异常的项$(t-t_0)$不出现。这一定义将在4.3.6节中给出，就至今所考虑过的情况而言，它相当于作为$W(t)$的函数在计算被积函数时取值为

$$\frac{1}{2}[W(t_i)+W(t_{i-1})]。$$

直接可证

$$S\int_{t_0}^{t} W(t')\mathrm{d}W(t')$$

$$=\underset{n\to\infty}{\text{ms-lim}}\sum_{i=1}^{n}\frac{W(t_i)+W(t_{i-1})}{2}[W(t_i)-W(t_{i-1})] \tag{4.2.22}$$

$$=\frac{1}{2}[W(t)^2-W(t_0)^2] \tag{4.2.23}$$

然而，斯特拉托诺维奇所定义的积分(如在(4.2.22)式中一样，我们总是用前缀 S 来标明这种积分）和伊藤所定义的积分并没有什么一般的关系，就是说，对于任意的函数 $G(t)$，在两种积分之间没有什么联系。(但是，在可以规定 $G(t)$ 与某个随机微分方程有关的场合，能够给出一个公式将两者互相联系起来，见 4.3.6 节)。

4.2.4 非可料函数

非可料函数的概念很可能被复杂的记号搞得莫明其妙，其实是相当简单的。我们记住这一情况，其中所有函数可以表示为某个维纳过程 $W(t)$ 的函数或者泛函，它们通过

$$x(t)-x(t_0)=\int_{t_0}^{t} a[x(t'),t']\mathrm{d}t'$$

$$+\int_{t_0}^{t} b[x(t'),t']\mathrm{d}W(t') \tag{4.2.24}$$

形式的随机微分(或积分）方程而得到。如果对所有 s 和 t，$t<s$，一个函数 $G(t)$ 与 $W(s)-W(t)$ 统计独立，则称 $G(t)$ 为 t 的非可料函数。这意味着 $G(t)$ 是同维纳过程在 t 之后的状态相独立的。显然这是对一个物理学函数的相当合理的要求，该函数可以是(4.2.24)那样的方程的解，看来方程中清楚地启示了 $x(t)$ 仅涉及 $t'\leqslant t$ 时的 $W(t')$。

作为例子，特殊的 t 的非可料函数是，

i) $W(t)$

ii) $\int_{t_0}^{t}\mathrm{d}t'F[W(t')]$

iii) $\int_{t_0}^{t} dW(t')F[W(t')]$

iv) $\int_{t_0}^{t} dt'G(t')$

v) $\int_{t_0}^{t} dW(t')G(t')$

} 当 $G(t)$ 自己是一个非可料函数时

结论(iii)和(v)依赖于这样的事实：按 (4.2.10) 所定义的伊藤随机积分是仅涉及 $t'<t$ 的 $G(t')$ 和 $t'\leqslant t$ 的 $W(t')$ 的序列的极限。

特地考虑非可料函数的理由是：

i) 可以导出许多仅对这种函数才正确的结论。

ii) 它们在许多场合自然而然地出现，如涉及时间的微分方程研究，其中预料有某种意义下的因果性，即未知的将来不可能影响现在。

iii) 随机微分方程的定义需要这种函数。

4.2.5 $dW(t)^2=dt$ 和 $dW(t)^{2+N}=0$ 的证明

题头中的公式是将伊藤微积分作为平常的计算工具来应用的关键。但那样的写法不甚精确，真正的意义是对任意的非可料函数 $G(t)$ 有

$$\int_{t_0}^{t}[dW(t')]^{2+N}G(t') \equiv \underset{n\to\infty}{\text{ms-lim}}\sum_i G_{i-1}\Delta W_i^{2+N}$$

$$=\int_{t_0}^{t} dt'G(t') \quad (\text{对于 } N=0)$$

$$=0 \quad (\text{对于 } N>0) \tag{4.2.25}$$

证明是颇为直截了当的。对 $N=0$，定义

$$I=\lim_{n\to\infty}\langle[\sum_i G_{i-1}(\Delta W_i^2-\Delta t_i)]^2\rangle \tag{4.2.26}$$

$$=\lim_{n\to\infty}\Big\{\langle\sum_i \underbrace{(G_{i-1})^2}_{\text{独立}}(\Delta W_i^2-\Delta t_i)^2$$

$$+\sum_{i<j}\underbrace{2G_{i-1}G_{j-1}(\Delta W_i^2-\Delta t_i)(\Delta W_j^2-\Delta t_j)}_{\text{独立}}\Big\rangle\Big\} \tag{4.2.27}$$

相邻水平花括弧所标出的因式是相互独立的，这是因为维纳过程的性质，又因为 G_i 是非可料函数的值，它们与所有 $j<i$ 的 ΔW_i 相独立。(*)

应用这种独立性，可以将均值因式分解，并用

i) $\langle \Delta W_i \rangle = \Delta t_i$

ii) $\langle (\Delta W_i^2-\Delta t_i)^2\rangle = 2\Delta t_i^2$（由 ΔW_i 的高斯特性）

可得

$$I = 2\lim_{n\to\infty}\Big[\sum_i \Delta t_i^2\langle (G_{i-1})^2\rangle\Big] \tag{4.2.28}$$

在关于 $G(t)$ 的合理而轻微的条件下（例如有界性），此式意味着

$$\text{ms-}\lim_{n\to\infty}\Big(\sum_i G_{i-1}\Delta W_i^2-\sum_i G_{i-1}\Delta t_i\Big)=0 \tag{4.2.29}$$

并因为

$$\text{ms-}\lim_{n\to\infty}\sum_i G_{i-1}\Delta t_i=\int_{t_0}^{t} \mathrm{d}t' G(t') \tag{4.2.30}$$

我们有

$$\boxed{\int_{t_0}^{t}[\mathrm{d}W(t')]^2 G(t')=\int_{t_0}^{t}\mathrm{d}t' G(t')}$$

注解

i) $N>0$ 时 $\int_{t_0}^{t} G(t)\,[\mathrm{d}W(t)]^{2+N}=0$ 的证明是相似的，要用高斯变量的更高阶矩的明显表达式(2.8.1 节)。

ii) $\mathrm{d}W(t)$ 仅仅在积分中出现，所以当我们局限于非可料函数时，可以简单地写成

$$\mathrm{d}W(t)^2 \equiv \mathrm{d}t \tag{4.2.31}$$

$$\mathrm{d}W(t)^{2+N}\equiv 0 \quad (N>0) \tag{4.2.32}$$

iii) 这个结论仅对伊藤积分成立，因为我们用到了 ΔW_i 与 G_{i-1} 独立的事实。在斯特拉托诺维奇积分中，

(*) 原文为 "…G_i 是非可料函数的值，它们与所有 $j>i$ 的 ΔW_i 相独立"。已作更正。——译注

$$\Delta W_i = W(t_i) - W(t_{i-1}) \tag{4.2.33}$$

$$G_{i-1} = G\left[\frac{1}{2}(t_i + t_{i-1})\right] \tag{4.2.34}$$

虽然 $G(t)$ 是非可料的，但并不足以保证这样定义的 ΔW_i 和 G_{i-1} 的独立性。

iv) 用同样的方法可以证明

$$\int_{t_0}^{t} G(t')\mathrm{d}t'\mathrm{d}W(t') = \underset{n\to\infty}{\text{ms-lim}}\, \Sigma G_{i-1}\Delta W_i \Delta t_i = 0 \tag{4.2.35}$$

对更高的幂次是相似的。刻划这些结论的特性的最简单方法是断言 $\mathrm{d}W(t)$ 是 $\frac{1}{2}$ 阶的无穷小，并在微分计算中略去高于 1 阶的无穷小。

4.2.6 伊藤随机积分的性质

a) 存在性

可以证明，只要函数 $G(t')$ 在闭区间 $[t_0, t]$ 上是连续和非可料的，则伊藤随机积分存在[4.3]。

b) 多项式的积分

我们形式地应用 4.2.5 节的结果：

$$\mathrm{d}[W(t)]^n = [W(t) + \mathrm{d}W(t)]^n - W(t)^n$$

$$= \sum_{r=1}^{n} \binom{n}{r} W(t)^{n-r}\mathrm{d}W(t)^r$$

再利用对所有 $r>2$ 有 $\mathrm{d}W(t)^r \to 0$ 的事实

$$= nW(t)^{n-1}\mathrm{d}W(t) + \frac{n(n-1)}{2} W(t)^{n-2}\mathrm{d}t \tag{4.2.36}$$

所以

$$\boxed{\int_{t_0}^{t} W(t')^n \mathrm{d}W(t') = \frac{1}{n+1}[W(t)^{n+1} - W(t_0)^{n+1}] - \frac{n}{2}\int_{t_0}^{t} W(t)^{n-1}\mathrm{d}t} \tag{4.2.37}$$

c) 两种积分

注意到对每个 $G(t)$ 有着两种积分，即

$$\int_{t_0}^{t} G(t')\mathrm{d}t' \quad 和 \quad \int_{t_0}^{t} G(t')\mathrm{d}W(t')$$

上面的方程中二者都出现。一般，在两种积分之间并没有什么联系。

d) 一般的微分法则

在构成微分式时(如上述(b)中那样)，必须保留 $\mathrm{d}W(t)$ 的所有直至二阶的项。这意味着，例如有

$$\begin{aligned}\mathrm{d}\{\exp[W(t)]\} &= \exp[W(t)+\mathrm{d}W(t)]-\exp[W(t)]\\ &= \exp[W(t)]\left[\mathrm{d}W(t)+\frac{1}{2}\,\mathrm{d}W(t)^2\right]\\ &= \exp[W(t)]\left[\mathrm{d}W(t)+\frac{1}{2}\mathrm{d}t\right]\end{aligned} \tag{4.2.38}$$

或更一般地，

$$\begin{aligned}\mathrm{d}f[W(t),t] = &\frac{\partial f}{\partial t}\mathrm{d}t+\frac{1}{2}\,\frac{\partial^2 f}{\partial t^2}\,(\mathrm{d}t)^2+\frac{\partial f}{\partial W}\mathrm{d}W(t)\\ &+\frac{1}{2}\,\frac{\partial^2 f}{\partial W^2}\,[\mathrm{d}W(t)]^2\\ &+\frac{\partial^2 f}{\partial W\partial t}\,\mathrm{d}t\mathrm{d}W(t)+\cdots\end{aligned}$$

我们利用

$(\mathrm{d}t)^2\to 0$

$\mathrm{d}t\mathrm{d}W(t)\to 0$(4, 2.5 节，注解(iv))

$[\mathrm{d}W(t)]^2=\mathrm{d}t$

且所有更高阶的项消失，所以可得

$$\boxed{\mathrm{d}f[W(t),t]=\left(\frac{\partial f}{\partial t}+\frac{1}{2}\,\frac{\partial^2 f}{\partial W^2}\right)\mathrm{d}t+\frac{\partial f}{\partial W}\,\mathrm{d}W(t)} \tag{4.2.39}$$

e) 均值公式

对于非可料的 $G(t)$

$$\left\langle \int_{t_0}^{t} G(t')\mathrm{d}W(t') \right\rangle = 0 \tag{4.2.40}$$

因为 $G(t)$ 是非可料的，在随机积分的定义中有

$$\left\langle \sum_i G_{i-1}\Delta W_i \right\rangle = \sum_i \langle G_{i-1}\rangle \langle \Delta W_i \rangle = 0 \tag{4.2.41}$$

又由 2.9.5 节得知 ms-lim 与〈　〉两个运算可交换次序，因此取(4.2.41)的极限，我们就得出了结论。

这个结论对斯特拉托诺维奇积分是不正确的，因为 G_{i-1} 的值被取在区间的中央而可能同 ΔW_i 相关。

f) 相关函数

如果 $G(t)$ 和 $H(t)$ 是任意的连续非可料函数，

$$\left\langle \int_{t_0}^{t} G(t')\mathrm{d}W(t') \int_{t_0}^{t} H(t')\mathrm{d}W(t') \right\rangle = \int_{t_0}^{t} \mathrm{d}t' \langle G(t')H(t') \rangle \tag{4.2.42}$$

证明 注意

$$\left\langle \sum_i G_{i-1}\Delta W_i \sum_j H_{j-1}\Delta W_j \right\rangle = \left\langle \sum_i G_{i-1}H_{i-1}(\Delta W_i)^2 \right\rangle + \left\langle \sum_{i>j} (G_{i-1}H_{j-1} + G_{j-1}H_{i-1})\Delta W_j \Delta W_i \right\rangle \tag{4.2.43}$$

在第二项中，ΔW_i 因 $j<i$ 而与所有其他的项独立，G 和 H 又是非可料的，因此我们可以析出因式 $\langle \Delta W_i \rangle = 0$ 以致整个第二项为零。再应用

$$\langle \Delta W_i^2 \rangle = \Delta t_i \tag{4.2.44}$$

并交换求均值和求极限运算，即得结论。

形式上，这是与郎之万项 $\xi(t)$ 为 δ 相关并与 $G(t)$ 和 $H(t)$ 不相关的思想等价的。若重新写出

$$\mathrm{d}W(t) \to \xi(t)\mathrm{d}t \tag{4.2.45}$$

显然，如果 $G(t)$ 和 $H(t)$ 为非可料，则 $\xi(t)$ 与它们独立。我们有

$$\int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t} \mathrm{d}s' \langle G(t')H(s')\xi(t')\xi(s') \rangle$$

$$= \int_{t_0}^{t} \int_{t_0}^{t} \mathrm{d}t' \mathrm{d}s' \langle G(t') H(s') \rangle \langle \xi(t') \xi(s') \rangle$$

$$= \int_{t_0}^{t} \mathrm{d}t' \langle G(t') H(t') \rangle \qquad (4.2.46)$$

此式蕴涵着

$$\langle \xi(t) \xi(s) \rangle = \delta(t-s)$$

但是这里出现了定义的一个重要特点：研究随机微分方程时在涉及 δ 函数的积分中，经常出现 δ 函数的自变量等于积分上限或下限的情况，即是得到

$$I_1 = \int_{t_1}^{t_2} \mathrm{d}t f(t) \delta(t-t_1) \qquad (4.2.47)$$

或

$$I_2 = \int_{t_1}^{t_2} \mathrm{d}t f(t) \delta(t-t_2) \qquad (4.2.48)$$

那样的积分，关于它们的值可以作出各种约定。我们将证明，按本段内容的论述，必须作出

$$I_1 = f(t_1) \qquad (4.2.49)$$

$$I_2 = 0 \qquad (4.2.50)$$

的解释，相应于认定 δ 函数的权全都在一积分的下限处，毫不出现在上限处。为证实这一点，注意

$$\left\langle \int_{t_0}^{t} G(t') \mathrm{d}W(t') \left[\int_{t_0}^{t'} H(s') \mathrm{d}W(s') \right] \right\rangle = 0 \qquad (4.2.51)$$

这是因为，按 4.2.4 节注解(V) 由方括弧中的积分定义的函数是非可料的，于是用它再乘上非可料函数 $G(t')$ 所得整个被积函数自己也是非可料的。因此按 4.2.6e 的结论得知均值为零。

应用关于郎之万源 $\xi(t)$ 的公式，我们可以改写(4.2.51)为

$$\int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}s' \langle G(t') H(s') \rangle \delta(t'-s') = 0 \qquad , \qquad (4.2.52)$$

它对应于不计 δ 函数在上限处的权，因而全部权必须在下限处计算。

这个性质是(4.2.10)中的伊藤积分定义的直接推论，式中的增量“指向将来”，即是，我们把 $\mathrm{d}W(t)$ 理解为

$$dW(t) = W(t+dt) - W(t) \tag{4.2.53}$$

在斯特拉托诺维奇积分的场合，我们得出颇为不同的公式，这可全然不象伊藤积分场合那样容易证明，但是它认定了选用

$$\left.\begin{aligned} I_1 &= \frac{1}{2}\ f(t_1) \\ I_2 &= \frac{1}{2}\ f(t_2) \end{aligned}\right\}\text{(斯特拉托诺维奇)} \tag{4.2.54}$$

这意味着，出现于积分上、下限的 δ 函数都计算它的一半的权．虽然这个公式直观上比伊藤公式更令人满意，但用起来更加复杂，特别是在随机微分方程的摄动理论中，其中伊藤方法使得好多的项成为零．

4.3 随机微分方程 (SDE)

在 4.1 节中我们得出结论，郎之万方程

$$\frac{dx}{dt} = a(x, t) + b(x, t)\xi(t) \tag{4.3.1}$$

的最令人满意的解释是随机积分方程

$$x(t) - x(0) = \int_0^t dt' a[x(t'), t'] + \int_0^t dW(t') b[x(t'), t'] \tag{4.3.2}$$

不幸，4.1 节的推理并未给出所用随机积分的种类．伊藤积分在数学上和技巧上是最令人满意的，但遗憾的是它并不总是物理上自然的选择．斯特拉托诺维奇积分在一定意义上是一个自然的选择，这一解释假定 $\xi(t)$ 是具有有限相关时间的真实噪声(不是白噪声)，因而该相关时间在计算可测量的量之后允许成为无限小．此外，斯特拉托诺维奇解释使我们可以用通常的微积分，这对伊藤解释是不可能的．

从数学观点，作出明确的抉择是因为几乎不可能用斯特拉托诺维奇积分来作出证明．所以我们将定义伊藤随机微分方程，再导出它同斯特拉托诺维奇随机微分方程的等价性，并按具体情况

随便采用哪一个。白噪声随机微分方程和实际噪声系统之间的联系将在6.5节中说明。

4.3.1 伊藤随机微分方程：定义

一个随机量 $x(t)$ 遵循

$$\mathrm{d}x(t) = a[x(t),\ t]\mathrm{d}t + b[x(t),\ t]\mathrm{d}W(t) \tag{4.3.3}$$

形式的伊藤随机微分方程，只要对所有 t 和 t_0 有

$$x(t) = x(t_0) + \int_{t_0}^{t} \mathrm{d}t' a[x(t'),\ t'] + \int_{t_0}^{t} \mathrm{d}W(t') b[x(t'),\ t'] \tag{4.3.4}$$

在考虑 (4.3.4) 中的系数必须满足什么条件之前，明智的做法是先考虑这种方程的解意味着什么，以及解的唯一性在此意味着什么。为此，我们考虑随机微分方程的离散形式，取网格点 t_i（如图4.2所示）使

$$t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = t \tag{4.3.5}$$

并将方程写成为

$$x_{i+1} = x_i + a(x_i,\ t_i)\Delta t_i + b(x_i,\ t_i)\Delta W_i \tag{4.3.6}$$

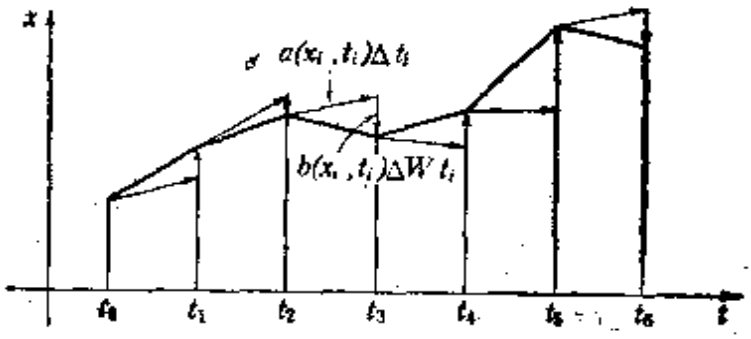


图4.2 构成随机微分方程 $\mathrm{d}x(t) = a[x(t),\ t]\mathrm{d}t + b[x(t),\ t]\mathrm{d}W(t)$ 的近似解的柯西-欧拉程序的图解

这里

$$\left.\begin{aligned} x_i &= x(t_i) \\ \Delta t_i &= t_{i+1} - t_i \\ \Delta W_i &= W(t_{i+1}) - W(t_i) \end{aligned}\right\} \tag{4.3.7}$$

从(4.3.6)可见，近似求解方程的程序计算 x_{i+1} 的方法是从已知的 x_i 加上一个确定性项

$$a(x_i, t_i)\Delta t_i \tag{4.3.8}$$

和一个随机项

$$b(x_i, t_i)\Delta W_i \tag{4.3.9}$$

该随机项包含有一个元素 ΔW_i，它是维纳过程的增量。如果(i) x_0 自身与所有 $t>t_0$ 的 $W(t)-W(t_0)$ 相互独立(这样一来，如果将初始条件认为是随机的，则必须是非可料的)，(ii) $a(x, t)$对任何固定的 x 是 t 的非可料函数，则 ΔW_i 与 x_i 为统计独立。

应用 (4.3.6) 迭代构成近似解，可见 x_i 对于 $j\geqslant i$ 总是与 ΔW_j 独立。

令网格大小趋向于零就形式地构成了解。我们说解是唯一的，意味着对随机维纳过程 $W(t)$的一个给定的样本函数，所产生的方程的特定解是唯一的。我们说解存在，意味着对维纳过程的任选样本函数以概率 1 存在一个解。

上述构成解的方法称为柯西-欧拉方法，也可用于生成模拟结果。

然而，这并不是通常用来论证唯一性和存在性的方法，虽然这样做是可能的。存在性和唯一性在此将不予证明，有兴趣的读者在[4.3]中会找到有关的证明。在时间区间 $[t_0, T]$ 中的存在性和唯一性所要求的条件是

i) 李普希兹(Lipschitz)条件：存在 k 使

$$|a(x, t)-a(y, t)|+|b(x, t)-b(y, t)|\leqslant K|x-y| \tag{4.3.10}$$

对所有 x、y 和 $[t_0, T]$ 范围中所有 t 成立。

ii) 增长条件：存在 K 使对 $[t_0, T]$ 中所有 t 成立

$$|a(x, t)|^2+|b(x, t)|^2\leqslant K^2(1+|x|^2) \tag{4.3.11}$$

在这些条件下在 $[t_0, T]$ 范围内将存在唯一的非可料解 $x(t)$。

实用中遇到的几乎所有随机微分方程都满足李普希兹条件，因为它本质上是一种平滑性条件。而增长条件则常常违背，这并

不意味着解不存在，确切地说，它意味着解可能"爆发"到无穷大，即 x 的值可能在有限时间内成为无穷大，实用中是在一个有限随机的时间内。例如，常微分方程

$$\frac{dx}{dt} = \frac{1}{2} ax^3 \tag{4.3.12}$$

中就出现这种现象，该方程在 $t=0$ 处的初始条件 $x=x_0$ 下的通解为

$$x(t) = (-at + 1/x_0^2)^{-1/2} \tag{4.3.13}$$

如果 a 是正数，当 $x_0 = (at)^{-1/2}$ 时解成为无穷大，但若 a 是负数，则绝不会爆发。不满足李普希兹条件并不确保解会爆发。需要更精密的稳定性结果的话可参阅[4.3]。

4.3.2 伊藤随机微分方程的解的马尔可夫性质

现在我们来说明，随机微分方程 (4.3.4) 的解 $x(t)$ 是一个马尔可夫过程。这个结论直观上是显然的。因为在给定初始条件 $x(t_0)$ 下，未来时间的发展是(随机意义上)唯一确定的，即是，对 $t>t_0$ 的 $x(t)$ 仅由

i) $W(t)$ 的 $t>t_0$ 时的特定样本路径；

ii) $x(t_0)$ 的值。

所确定。因为 $x(t)$ 是 t 的非可料函数，$t>t_0$ 的 $W(t)$ 与 $t<t_0$ 的 $x(t)$ 相互独立。这样，在 $x(t_0)$ 已知时 $t>t_0$ 的 $x(t)$ 的时间发展与 $t<t_0$ 的 $x(t)$ 相独立。因此 $x(t)$ 是一个马尔可夫过程。精确证明见[4.3]。

4.3.3 变量代换：伊藤公式

考虑 $x(t)$ 的一个任意函数：$f[x(t)]$，它遵循什么样的随机微分方程呢？我们用 4.2.5 节的结果来展开 $df[x(t)]$ 至 $dW(t)$ 的二阶项：

$$df[x(t)] = f[x(t) + dx(t)] - f[x(t)]$$

$$= f'[x(t)]\mathrm{d}x(t) + \frac{1}{2} f''[x(t)]\mathrm{d}x(x)^2 + \cdots$$

$$= f'[x(t)]\{a[x(t), t]\mathrm{d}t + b[x(t), t]\mathrm{d}W(t)\}$$
$$+ \frac{1}{2} f''[x(t)]b[x(t), t]^2[\mathrm{d}W(t)]^2 + \cdots$$

其中所有其他项已被排除，因为它们有更高的阶。现在我们利用

$$[\mathrm{d}W(t)]^2 = \mathrm{d}t$$

得到

$$\boxed{\begin{aligned}\mathrm{d}f[x(t)] = \{&a[x(t), t]f'[x(t)] \\ &+ \frac{1}{2} b[x(t), t]^2 f''[x(t)]\}\mathrm{d}t \\ &+ b[x(t), t]f'[x(t)]\mathrm{d}W(t)\end{aligned}} \qquad (4.3.14)$$

这个公式称为伊藤公式，它说明变量代换不是由通常的微积分所给出，除非 $f[x(t)]$ 仅是 $x(t)$ 的线性函数。

多变量 实用中伊藤公式变得非常复杂，最容易的方法就是用 $\mathrm{d}W(t)$ 是 $\frac{1}{2}$ 阶无穷小这一规则的多变量形式。用相似于4.2.5节的方法可以证明，对于 n 维维纳过程 $\boldsymbol{W}(t)$，

$$\mathrm{d}W_i(t)\mathrm{d}W_j(t) = \delta_{ij}\mathrm{d}t \qquad (4.3.15a)$$

$$[\mathrm{d}W_i(t)]^{N+2} = 0 (N>0) \qquad (4.3.15b)$$

$$\mathrm{d}W_i(t)\mathrm{d}t = 0 \qquad (4.3.15c)$$

$$\mathrm{d}t^{1+N} = 0 (N>0) \qquad (4.3.15d)$$

这意味着 $\mathrm{d}W_i(t)$ 是一个 $\frac{1}{2}$ 阶无穷小。但要注意 (4.3.15a) 是 $\mathrm{d}W_i(t)$ 和 $\mathrm{d}W_j(t)$ 的独立性的结果。为了推出满足随机微分方程

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x}, t)\mathrm{d}t + \underset{\sim}{\boldsymbol{B}}(\boldsymbol{x}, t)\mathrm{d}\boldsymbol{W}(t) \qquad (4.3.16)$$

的 n 维向量 $\boldsymbol{x}(t)$ 的函数的伊藤公式，我们简单地仿效前面的步骤。结果是

$$\boxed{\begin{aligned}\mathrm{d}f(\boldsymbol{x}) = \Big\{&\sum_i A_i(\boldsymbol{x},\ t)\partial_i f(\boldsymbol{x}) \\ &+ \frac{1}{2}\sum_{i,j}[\underset{\sim}{\boldsymbol{B}}(\boldsymbol{x},\ t)\underset{\sim}{\boldsymbol{B}}^{\mathrm{T}}(\boldsymbol{x},\ t)]_{ij}\partial_i\partial_j f(\boldsymbol{x})\Big\}\mathrm{d}t \\ &+\sum_{i,j} B_{ij}(\boldsymbol{x},\ t)\partial_i f(\boldsymbol{x})\mathrm{d}W_j(t)\end{aligned}}$$

(4.3.17)

4.3.4 福克-普朗克方程和随机微分方程之间的联系

现在我们考虑任意 $f[x(t)]$ 的时间发展。应用伊藤公式

$$\begin{aligned}\langle \mathrm{d}f[x(t)]\rangle/\mathrm{d}t = \left\langle\frac{\mathrm{d}f[x(t)]}{\mathrm{d}t}\right\rangle &= \frac{\mathrm{d}}{\mathrm{d}t}\langle f[x(t)]\rangle \\ &= \langle a[x(t),\ t]\partial_x f + \frac{1}{2}\,b[x(t),\ t]^2\partial_x f\rangle\end{aligned}$$

(4.3.18)

然而，$x(t)$ 具有条件概率密度 $p(x,\ t|x_0,\ t_0)$ 且

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle f[x(t)]\rangle &= \int \mathrm{d}x f(x)\partial_t p(x,\ t|x_0,\ t_0) \\ &= \int \mathrm{d}x[a(x,\ t)\partial_x f \\ &\quad + \frac{1}{2}\,b(x,\ t)^2\partial_x^2 f]p(x,\ t|x_0,\ t_0)\end{aligned}\qquad (4.3.19)$$

现在此式与 3.4.1 节 (3.4.16) 式相同。在跟那里同样的条件下，作部分积分并弃去曲面项得到

$$\begin{aligned}\int \mathrm{d}x f(x)\partial_t p = \int \mathrm{d}x f(x)\Big\{&-\partial_x[a(x,\ t)p] \\ &+ \frac{1}{2}\,\partial_x^2[b(x,\ t)^2 p]\Big\}\end{aligned}$$

因此，由于 $f(x)$ 是任意的，

$$\boxed{\begin{aligned}\partial_t p(x,\ t|x_0,\ t_0) = &-\partial_x[a(x,\ t)p(x,\ t|x_0,\ t_0)] \\ &+ \frac{1}{2}\,\partial_x^2[b(x,\ t)^2 p(x,\ t|x_0,\ t_0)]\end{aligned}}$$

(4.3.20)

上述结果与3.5.2节中的结果完全相似，说明了扩散过程可以用一个相似于伊藤随机微分方程的方程局部逼近。

4.3.5 多变量系统

一般，多变量随机微分方程组可以对 n 个变量由

$$\mathrm{d}\boldsymbol{x}=\boldsymbol{A}(\boldsymbol{x},\ t)\mathrm{d}t+\underset{\sim}{\boldsymbol{B}}(\boldsymbol{x},\ t)\mathrm{d}\boldsymbol{W}(t) \tag{4.3.21}$$

来定义，其中$\boldsymbol{W}(t)$是如3.8.1节所定义的 n 元维纳过程。4.3.4节中所用的推理翻版到多变量场合证实条件概率密度

$$p(\boldsymbol{x},\ t|\boldsymbol{x}_0,\ t_0)=p$$

的福克-普朗克方程是

$$\partial_t p=-\sum_i \partial_i[A_i(\boldsymbol{x},\ t)p]+\frac{1}{2}\sum_{i,j}\partial_i\partial_j\{[\underset{\sim}{\boldsymbol{B}}(\boldsymbol{x},\ t)\underset{\sim}{\boldsymbol{B}}^T(\boldsymbol{x},\ t)]_{ij}p\} \tag{4.3.22}$$

注意所有使 $\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{B}}^T$ 相同的矩阵 $\underset{\sim}{\boldsymbol{B}}$ 产生同样的福克-普朗克方程。这意味着在用 $\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{S}}$ 取代 $\underset{\sim}{\boldsymbol{B}}$，其中 $\underset{\sim}{\boldsymbol{S}}$ 为正交即 $\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{S}}^T=1$ 时，我们可以得到相同的福克-普朗克方程。还得注意 $\underset{\sim}{\boldsymbol{S}}$ 可以与 $\boldsymbol{x}(t)$ 有关。这一事实可以更直接地看清，假定 $\underset{\sim}{\boldsymbol{S}}(t)$ 是一个正交阵，具有任意的关于 t 的非可料依赖性，然后定义

$$\mathrm{d}\boldsymbol{V}(t)=\underset{\sim}{\boldsymbol{S}}(t)\mathrm{d}\boldsymbol{W}(t) \tag{4.3.23}$$

现在向量 $\mathrm{d}\boldsymbol{V}(t)$ 是高斯变量 $\mathrm{d}\boldsymbol{W}(t)$ 的线性组合，其系数 $\underset{\sim}{\boldsymbol{S}}(t)$ 是与 $\mathrm{d}\boldsymbol{W}(t)$ 独立的，这是因为 $\underset{\sim}{\boldsymbol{S}}(t)$ 是非可料的。因而对任何 $\underset{\sim}{\boldsymbol{S}}(t)$ 的固定值，$\mathrm{d}\boldsymbol{V}(t)$ 是高斯变量。由于 $\underset{\sim}{\boldsymbol{S}}(t)$ 是正交的，它们的相关矩阵是

$$\begin{aligned}\langle \mathrm{d}V_i(t)\mathrm{d}V_j(t)\rangle&=\sum_{l,m}S_{il}(t)S_{jm}(t)\langle \mathrm{d}W_l(t)\mathrm{d}W_m(t)\rangle\\&=\sum_l S_{il}(t)S_{jl}(t)\mathrm{d}t=\delta_{ij}\mathrm{d}t\end{aligned} \tag{4.3.24}$$

因此，所有的矩与 $\underset{\sim}{\boldsymbol{S}}(t)$ 独立，与 $\mathrm{d}\boldsymbol{W}(t)$ 的矩相同，所以 $\mathrm{d}\boldsymbol{V}(t)$ 本

身也是有着跟 $\mathrm{d}\boldsymbol{W}(t)$ 相同的相关矩阵的高斯变量。最后，考虑在不同时刻的因式乘积的平均值，例如，如果在

$$\sum_{i,k}\langle[\mathrm{d}W_i(t)S_{ij}(t)]^m[\mathrm{d}W_k(t')S_{kl}(t')]^n\rangle \tag{4.3.25}$$

中 $t>t'$，我们可以把 $\mathrm{d}W_i(t)$ 的不同幂次的平均值因子分解出来，因为 $\mathrm{d}W_i(t)$ 与所有其他项相独立。作出这些计算，我们将发现在对 $\mathrm{d}W_i(t)$ 取平均值后 $\underset{\sim}{S}(t)$ 的正交性给出

$$\sum_k\langle[\mathrm{d}W_j(t)]^m\rangle\langle[\mathrm{d}W_k(t')S_{kl}(t')]^n\rangle \tag{4.3.26}$$

它又类似地给出 $\langle[\mathrm{d}W_j(t)]^m[\mathrm{d}W_i(t')]^n\rangle$ (*)，因此 $\mathrm{d}\boldsymbol{V}(t)$ 也是一个维纳过程的增量。正交变换只是将过程的样本路径混合起来，并不改变其随机特性。

因此，取代(4.3.21)我们可以写出

$$\mathrm{d}\boldsymbol{x}=\boldsymbol{A}(\boldsymbol{x},\ t)\mathrm{d}t+\underset{\sim}{\boldsymbol{B}}(\boldsymbol{x},\ t)\underset{\sim}{\boldsymbol{S}}^T(t)\underset{\sim}{\boldsymbol{S}}(t)\mathrm{d}\boldsymbol{W}(t) \tag{4.3.27}$$

$$=\boldsymbol{A}(\boldsymbol{x},\ t)\mathrm{d}t+\underset{\sim}{\boldsymbol{B}}(\boldsymbol{x},\ t)\underset{\sim}{\boldsymbol{S}}^T(t)\mathrm{d}\boldsymbol{V}(t) \tag{4.3.28}$$

又因 $\boldsymbol{V}(t)$ 本身就是一个维纳过程，这一方程等价于

$$\mathrm{d}\boldsymbol{x}=\boldsymbol{A}(\boldsymbol{x},\ t)\mathrm{d}t+\underset{\sim}{\boldsymbol{B}}(\boldsymbol{x},\ t)\underset{\sim}{\boldsymbol{S}}^T(t)\mathrm{d}\boldsymbol{W}(t) \tag{4.3.29}$$

它有着完全相同的福克-普朗克方程(4.3.22)。

在4.4.6节中我们将重新提及与这一等价性有关的一些例子。

4.3.6 斯特拉托诺维奇随机微分方程

被积函数为 $x(t)$ 和 t 的函数时，斯特拉托诺维奇[4.2]曾定义随机积分为

$$\boldsymbol{S}\int_{t_0}^{t}\boldsymbol{G}[x(t'),\ t']\mathrm{d}\boldsymbol{W}(t')$$

$$=\underset{n\to\infty}{\mathrm{ms\text{-}lim}}\sum_{i=1}^{n}\cdot\boldsymbol{G}\left\{\frac{x(t_i)+x(t_{i-1})}{2},\ t_{i-1}\right\}[\boldsymbol{W}(t_i)-\boldsymbol{W}(t_{i-1})] \tag{4.3.30}$$

(*) 从(4.3.25)到此式的推导似乎不能成立。——译注

应该注意仅是对$x(t)$的相依性被取平均了。如果$G(z, t)$关于t可微，积分与范围$[t_{i-1}, t_i]$中t的特定选择值相独立。

可以用斯特拉托诺维奇积分写出随机微分方程

$$x(t) = x(t_0) + \int_{t_0}^{t} \mathrm{d}t' a[x(t'), t'] + S\int_{t_0}^{t} \mathrm{d}W(t')\beta[x(t'), t'] \tag{4.3.31}$$

我们将证实它等价于一个适当的伊藤随机微分方程。

假设$x(t)$是

$$\mathrm{d}x(t) = a[x(t), t]\mathrm{d}t + b[x(t), t]\mathrm{d}W(t) \tag{4.3.32}$$

的一个解并导出相应的α和β。在两种情况下解$x(t)$都是相同的函数。先计算在

$$S\int_{t_0}^{t} \mathrm{d}W(t')\beta[x(t'), t']$$

和

$$\int_{t_0}^{t} \mathrm{d}W(t')b[x(t'), t']$$

之间的联系。我们有

$$S\int_{t_0}^{t} \mathrm{d}W(t')\beta[x(t'), t'] \simeq \sum_i \beta\left[\frac{x(t_i)+x(t_{i-1})}{2}, t_{i-1}\right][W(t_i) - W(t_{i-1})] \tag{4.3.33}$$

在(4.3.33)中我们记

$$x(t_i) = x(t_{i-1}) + \mathrm{d}x(t_{i-1})$$

并用伊藤随机微分方程(4.3.32)写出

$$\mathrm{d}x(t_i) = a[x(t_{i-1}), t_{i-1}](t_i - t_{i-1}) + b[x(t_{i-1}), t_{i-1}][W(t_i) - W(t_{i-1})] \tag{4.3.34}$$

再用伊藤公式，我们可以写出

$$\beta\left[\frac{x(t_i)+x(t_{i-1})}{2}, t_{i-1}\right] = \beta\left[x(t_{i-1}) + \frac{1}{2}\,dx(t_{i-1}), t_{i-1}\right]$$

$$= \beta(t_{i-1}) + \left[a(t_{i-1})\partial_x\beta(t_{i-1})\right.$$

$$+\frac{1}{4}\,b^2(t_{i-1})\partial_x^2\beta(t_{i-1})\Big]$$

$$\cdot\Big[\frac{1}{2}\,(t_i-t_{i-1})\Big]$$

$$+\frac{1}{2}\,b(t_{i-1})\partial_x\beta(t_{i-1})\,[W(t_i)$$

$$-W(t_{i-1})] \qquad (4.3.35)$$

(为简明起见，只要有可能我们就用 $\beta(t_i)$ 取代 $\beta[x(t_i), t_i]$，等等)。将所有这些式子代回原来的方程 (4.3.33)并照例弃去 $\mathrm{d}t^2$，$\mathrm{d}t\mathrm{d}W$ 项，置 $\mathrm{d}W^2=\mathrm{d}t$，可得

$$S\int=\sum_i\beta(t_{i-1})\,\{W(t_i)-W(t_{i-1})\}$$

$$+\frac{1}{2}\sum b(t_{i-1})\partial_x\beta(t_{i-1})\,(t_i-t_{i-1}) \qquad (4.3.36)$$

由此我们导出

$$\boxed{S\int_{t_0}^{t}\beta[x(t'),\,t']\mathrm{d}W(t')=\int_{t_0}^{t}\beta[x(t'),\,t']\,\mathrm{d}W(t')+\frac{1}{2}\int_{t_0}^{t}b[x(t'),\,t']\partial_x\beta[x(t'),\,t']\mathrm{d}t'} \qquad (4.3.37)$$

该公式给出了函数 $\beta[x(t'),\,t']$ 的伊藤积分和斯特拉托诺维奇积分之间的联系，其中 $x(t')$是伊藤随机微分方程(4.3.31)的解。它并没有给出任意函数的伊藤积分和斯特拉托诺维奇积分之间的一般联系。

如果我们现在选择

$$a(x,\,t)=a(x,\,t)-\frac{1}{2}\,b(x,\,t)\partial_x b(x,\,t)$$

$$\beta(x,\,t)=b(x,\,t) \qquad (4.3.38)$$

我们看到

伊藤随机微分方程

$$dx = a\,dt + b\,dW(t) \tag{4.3.39a}$$

等同于斯特拉托诺维奇随机微分方程

$$dx = \left[a - \frac{1}{2} b\partial_x b\right]dt + b\,dW(t) \tag{4.3.39b}$$

或者反过来

斯特拉托诺维奇随机微分方程

$$dx = \alpha\,dt + \beta\,dW(t) \tag{4.3.40a}$$

等同于伊藤随机微分方程

$$dx = \left[\alpha + \frac{1}{2} \beta\partial_x \beta\right]dt + \beta\,dW(t) \tag{4.3.40b}$$

注解

i) 应用伊藤公式 (4.3.14) 可以证明在斯特拉托诺维奇随机微分方程中变量代换法则与通常的微积分中完全相同。从斯特拉托诺维奇随机微分方程(4.3.40a)出发，转换成伊藤随机微分方程(4.3.40b)。作新的变量代换 $y = f(x)$，其逆为 $x = g(y)$。

定义 $\bar{\alpha}(y) = \alpha[g(y)]$

$$\bar{\beta}(y) = \beta[g(y)]$$

应用伊藤公式并注意

$$df/dx = (dg/dy)^{-1}$$

可得伊藤随机微分方程

$$dy = \left[\bar{\alpha}\left(\frac{dg}{dy}\right)^{-1} + \frac{1}{2}\bar{\beta}\partial_y\bar{\beta}\left(\frac{dg}{dy}\right)^{-2} - \frac{1}{2}\frac{d^2g}{dy^2}\left(\frac{dg}{dy}\right)^{-3}\bar{\beta}^2\right]dt + \left(\frac{dg}{dy}\right)^{-1}\bar{\beta}\,dW$$

现在用(4.3.39)转换回斯特拉托诺维奇方程，我们得到

$$dy = (\bar{\alpha}\,dt + \bar{\beta}\,dW)\left(\frac{dg}{dy}\right)^{-1}$$

或

$$\mathrm{d}f[x(t)] = \{\alpha[x(t),\ t]\mathrm{d}t + \beta[x(t),\ t]\mathrm{d}W(t)\}\, f'[x(t)] \tag{4.3.41}$$

这是跟通常的微积分一样的。

ii) 多变量。如果一个多变量伊藤方程为

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x},\ t)\mathrm{d}t + \underset{\sim}{\boldsymbol{B}}(\boldsymbol{x},\ t)\mathrm{d}\boldsymbol{W}(t) \tag{4.3.42}$$

则类似地可以证明相应的斯特拉托诺维奇方程由代换

$$A_i^* = A_i - \frac{1}{2}\sum_{j,k} B_{kj}\partial_k B_{ij} \tag{4.3.43}$$

$$B_{ij}^* = B_{ij}$$

来给出。

iii) 福克-普朗克方程。应用 (4.3.43) 式和伊藤随机微分方程跟福克-普朗克方程之间的已知关系(4.3.5 节)，相应于斯特拉托诺维奇随机微分方程

$$(S)\mathrm{d}\boldsymbol{x} = \boldsymbol{A}^*(\boldsymbol{x},\ t)\mathrm{d}t + \underset{\sim}{\boldsymbol{B}^*}(\boldsymbol{x},\ t)\mathrm{d}\boldsymbol{W}(t) \tag{4.3.44}$$

的福克-普朗克方程可取

$$\boxed{\partial_t p = -\sum_i \partial_i\langle A_i^* p\rangle + \frac{1}{2}\sum_{i,j,k}\partial_i\langle B_{ik}^*\partial_j[B_{jk}^* p]\rangle} \tag{4.3.45}$$

的形式，通常称它为福克-普朗克方程的“斯特拉托诺维奇形式”。与随机微分方程的两种形式相对照，福克-普朗克方程的两种形式具有不同的外表，但(当然)是用那些常用微积分的同样法则来解释的。稍后我们将发现福克-普朗克方程的斯特拉托诺维奇形式确实在某些论述中自然而然地产生出来(6.6 节)。

iv) 伊藤积分和斯特拉托诺维奇积分的比较

在(4.3.30)中所定义的斯特拉托诺维奇积分是一个相当专门的概念，因为它仅能对二元函数 $G(z,t)$ 作出定义。按

$$G\left(x\left[\frac{1}{2}(t_i + t_{i-1})\right],\ \frac{1}{2}(t_i + t_{i-1})\right)$$

形状的更“显然”的定义在斯特拉托诺维奇的原始定义中并未采用，虽然在文献中(包括本书第一版)流传着一种见解，认为这就提

供了斯特拉托诺维奇积分的定义。但这个更显然的定义无法证明是收敛的，见 [4.6] (*)。实际上，斯特拉托诺维奇积分的精确定义从本体上来看是并不十分令人关心的，反过来按通常微积分所给出的变量代换法则的性质却是很有意义的，对此，由定义所给出的保证并不如两种积分间的联系(4.3.37, 43) 那么多。可以简单地把斯特拉托诺维奇积分定义为(4.3.37)所给出的那样，其中函数服从随机微分方程(4.3.31)，这一定义在数学上会充分令人满意，大大减少混淆。

4.3.7 对初始条件和参数的依赖性

如同确定性微分方程的场合完全一样，如果出现于随机微分方程中的函数连续地依赖于参数，则在正常情况下解也连续地依赖于参数。同样，解也连续地依赖于初始条件。对此我们来更精确地作出系统陈述。考虑一个单变量方程

$$\left.\begin{aligned} &\mathrm{d}x = a(\lambda, x, t)\mathrm{d}t + b(\lambda, x, t)\mathrm{d}W(t) \\ &\text{具有初始条件} \\ &x(t_0) = c(\lambda) \end{aligned}\right\} \qquad (4.3.46)$$

其中 λ 是一个参数。令(4.3.36)的解是 $x(\lambda, t)$。假设

i) $\operatorname{st-lim}_{\lambda\to\lambda_0} c(\lambda) = c(\lambda_0)$;

ii) $\lim_{\lambda\to\lambda_0}\{\sup_{t\in[t_0,T],|x|<N}[|a(\lambda, x, t) - a(\lambda_0, x, t)| + |b(\lambda, x, t) - b(\lambda_0, x, t)|]\} = 0$;

iii) 存在一个与 λ 无关的 K 使

$$|a(\lambda, x, t)|^2 + |b(\lambda, x, t)|^2 \leqslant K^2(1 + |x|^2)$$

则有

$$\operatorname{st-lim}_{\lambda\to\lambda_0}\{\sup_{t\in[t_0,T]}|x(\lambda, t) - x(\lambda_0, t)|\} = 0 \qquad (4.3.47)$$

其证明见[4.1]。

注解

(*) 原书末未列出该文献。——译注

i) 回忆随机收敛的定义，极限式(4.3.47)的解释为，当 $\lambda \to \lambda_0$ 时在任何有限区间 $[t_0, T]$ 内 $x(\lambda, t)$ 和 $x(\lambda_0, t)$ 间的最大偏差非零的概率趋向于零。

ii) 对初始条件的依赖性可令 a 和 b 与 λ 无关而得到。

iii) 在证明摄动展式中这个结论将是很有用的。

iv) 函数 $a(x, t)$ 和 $b(x, t)$ 本身为非随机的情况下，条件(ii)用最惯常的形式写出。常会出现 $a(x, t)$ 和 $b(x, t)$ 本身是随机(非可料)函数的情况，这时条件(ii)必须用概率性的叙述来取代。事实上，用 $\underset{\lambda \to \lambda_0}{\text{st-lim}}$ 取代 $\underset{\lambda \to \lambda_0}{\lim}$ 就够了。

4.4 一些例子和解

4.4.1 不依赖于 x 的系数

简单的方程

$$dx = a(t)dt + b(t)dW(t) \tag{4.4.1}$$

具有非随机的时间函数 $a(t)$ 和 $b(t)$，可以直接积分来求解，

$$x(t) = x_0 + \int_{t_0}^{t} a(t)dt + \int_{t_0}^{t} b(t)dW(t) \tag{4.4.2}$$

这里 x_0 可以是非随机的初始条件，或者可以是随机的，但对 $t > t_0$ 必须与 $W(t) - W(t_0)$ 独立，否则的话 $x(t)$ 不是非可料的。

按上式的构造，只要 x_0 是非随机的或它自己是高斯变量，$x(t)$ 也是高斯变量，因为

$$\int_{t_0}^{t} b(t)dW(t)$$

只是无穷小高斯变量的一个线性组合。更有甚者

$$\langle x(t) \rangle = \langle x_0 \rangle + \int_{t_0}^{t} a(t)dt$$

(因伊藤积分的均值为零)以及

$$\langle [x(t) - \langle x(t) \rangle][x(s) - \langle x(s) \rangle] \rangle \equiv \langle x(t), x(s) \rangle$$

$$= \left\langle \int_{t_0}^{t} b(t')dW(t') \int_{t_0}^{s} b(s')dW(s') \right\rangle$$

$$= \int_{t_0}^{\min(t,s)} [b(t')]^2 \mathrm{d}t' \text{(*)}$$

这里我们用到了(4.2.42)的结论，但有

$$\left.\begin{aligned} G(t') &= b(t') \\ &= 0 \end{aligned}\right\} \quad \begin{aligned} &t' < t \\ &t' \geqslant t \end{aligned}$$

$$\left.\begin{aligned} H(t') &= b(t') \\ &= 0 \end{aligned}\right\} \quad \begin{aligned} &t' < s \\ &t' \geqslant s \end{aligned}$$

这样就完全确定了该随机过程。

4.4.2 乘积线性白噪声过程

方程

$$\mathrm{d}x = cx\mathrm{d}W(t) \tag{4.4.3}$$

称为乘积线性白噪声过程。因为它关于 x 为线性，但“噪声项” $\mathrm{d}W(t)$ 乘上 x。我们可以用伊藤公式精确地求解该方程。用

$$y = \log x \tag{4.4.4}$$

定义一个新变量，所以

$$\mathrm{d}y = \frac{1}{x}\mathrm{d}x - \frac{1}{2x^2}(\mathrm{d}x)^2 = cW(t) - \frac{1}{2}c^2\mathrm{d}t \tag{4.4.5}$$

现在这个方程可直接积分，所以我们得到

$$y(t) = y(t_0) + c[W(t) - W(t_0)] - \frac{1}{2}c^2(t - t_0) \tag{4.4.6}$$

因此

$$x(t) = x(t_0)\exp\left\{c[W(t) - W(t_0)] - \frac{1}{2}c^2(t - t_0)\right\} \tag{4.4.7}$$

对任何零均值高斯变量 z 可应用有关公式算出均值

$$\langle \exp z \rangle = \exp(\langle z^2 \rangle / 2) \tag{4.4.8}$$

所以

$$\langle x(t) \rangle = \langle x(t_0) \rangle \exp\left[\frac{1}{2}c^2(t - t_0) - \frac{1}{2}c^2(t - t_0)\right] = \langle x(t_0) \rangle \tag{4.4.9}$$

(*) 此式还应加上 $\mathrm{var}\{x_0\}$ 项。——译注

根据定义这个结论也是显然的，因为

$$\langle \mathrm{d}x \rangle = \langle cx\mathrm{d}W(t) \rangle = 0$$

所以有

$$\frac{\mathrm{d}\langle x \rangle}{\mathrm{d}t} = 0$$

我们还可以算出自相关函数

$$\begin{aligned}\langle x(t)x(s) \rangle &= \langle x(t_0)^2 \rangle \left\langle \exp\left\{ c[W(t) + W(s) - 2W(t_0)] \right.\right. \\ &\quad \left.\left. - \frac{1}{2} c^2(t + s - 2t_0) \right\} \right\rangle \\ &= \langle x(t_0)^2 \rangle \exp\left\{ \frac{1}{2} c^2 [\langle [W(t) \right. \\ &\quad \left. + W(s) - 2W(t_0)]^2 \rangle - (t + s - 2t_0)] \right\} \\ &= \langle x(t_0)^2 \rangle \exp\left\{ \frac{1}{2} c^2 [t + s - 2t_0 \right. \\ &\quad \left. + 2\min(t, s) - (t + s - 2t_0)] \right\} \\ &= \langle x(t_0)^2 \rangle \exp[c^2 \min(t - t_0, s - t_0)] \end{aligned} \tag{4.4.10}$$

斯特拉托诺维奇解释 上述方程作为斯特拉托诺维奇解释下也是可以求解的，但通常的微积分应该成立。这样，取代(4.4.5)我们应有

$$\mathrm{d}y = c\mathrm{d}W(t)$$

因此

$$x(t) = x(t_0) \exp\{c[W(t) - W(t_0)]\} \tag{4.4.11}$$

在此场合

$$\langle x(t) \rangle = \langle x(t_0) \rangle \exp\left[-\frac{1}{2} c^2 (t - t_0) \right] \tag{4.4.12}$$

以及

$$\begin{aligned}\langle x(t)x(s) \rangle = \langle x(t_0)^2 \rangle \exp\left\{ \frac{1}{2} c^2 [t + s - 2t_0 \right. \\ \left. + 2\min(t - t_0, s - t_0)] \right\}\end{aligned} \tag{4.4.13}$$

可见在两种答案之间存在着明显的差别。

4.4.3 具有噪声频率的复振子

这是久保(Kubo)所给出的一个模型的简化，又是前面复变量例子的稍稍推广。我们考虑

$$\frac{\mathrm{d}z}{\mathrm{d}t} = i[\omega + \sqrt{2\gamma}\xi(t)]z \tag{4.4.14}$$

它在形式上表示一个振子的简单模型，其平均频率 ω 被噪声项 $\xi(t)$ 所扰动。

物理上，写成斯特拉托诺维奇方程是它的最好的模型

$$(S)\mathrm{d}z = i[\omega \mathrm{d}t + \sqrt{2\gamma}\mathrm{d}W(t)]z \tag{4.4.15}$$

它等价于伊藤方程(由 4.3.6 节)

$$\mathrm{d}z = [(i\omega - \gamma)\mathrm{d}t + i\sqrt{2\gamma}\mathrm{d}W(t)]z \tag{4.4.16}$$

取均值，我们马上见到有

$$\frac{\mathrm{d}\langle z\rangle}{\mathrm{d}t} = (i\omega - \gamma)\langle z\rangle \tag{4.4.17}$$

它具有阻尼振荡解

$$\langle z(t)\rangle = \exp[(i\omega - \gamma)t]\langle z(0)\rangle \tag{4.4.18}$$

在 6.6 节中我们将完满地证明为什么斯特拉托诺维奇模型是更加合适的。看清这一点的最明显的方法是注意到实际上 $\xi(t)$ 总得比白噪声平滑一些，通常的微积分应该用得上，如同在斯特拉托诺维奇解释的情况中一样。

这时求解原始斯特拉托诺维奇方程所得到的相关函数是

$$\langle z(t)z(s)\rangle = \langle z(0)^2\rangle \exp[(i\omega - \gamma)(t+s) - 2\gamma\min(t, s)] \tag{4.4.19}$$

取极限 $t, s \to \infty$，$t + \tau = s$，有

$$\lim_{t\to\infty}\langle z(t+\tau)z(t)\rangle = 0 \tag{4.4.20}$$

然而物理上感兴趣的相关函数是复相关

$$\begin{aligned}\langle z(t)z^*(s)\rangle &= \langle |z(0)|^2\rangle\langle \exp\{i\omega(t-s)\\ &\quad + i\sqrt{2\gamma}[W(t)-W(s)]\}\rangle\\ &= \langle |z(0)|^2\rangle \exp\{i\omega(t-s)\\ &\quad -\gamma[t+s-2\min(t,s)]\}\\ &= \langle |z(0)|^2\rangle \exp\{i\omega(t-s)-\gamma|t-s|\}\end{aligned} \tag{4.4.21}$$

这样，复相关函数具有一个纯粹起因于噪声的阻尼项。它可被认为是一个导致无相位差效应的噪声，由此，对大的时间差 $z(t)$ 和 $z^*(s)$ 成为相互独立。

按范坎彭的论述 [4.6]，一个实际的振子不可能由这个复振子模型来描述，然而定性的行为是很相象的，而且这个模型可被视为是具有噪声频率的振子的一个样板模型。

4.4.4 奥恩斯坦-乌伦贝克过程

利用由奥恩斯坦-乌伦贝克过程所给出的福克-普朗克方程 (3.8.4节)，应用4.3.4节的结论立即可以写出随机微分方程为

$$\mathrm{d}x = -kx\mathrm{d}t + \sqrt{D}\mathrm{d}W(t) \tag{4.4.22}$$

直接求解它，置

$$y = x\mathrm{e}^{kt} \tag{4.4.23}$$

则

$$\begin{aligned}\mathrm{d}y &= (\mathrm{d}x)\mathrm{d}(\mathrm{e}^{kt}) + (\mathrm{d}x)\mathrm{e}^{kt} + x\mathrm{d}(\mathrm{e}^{kt})\\ &= [-kx\mathrm{d}t + \sqrt{D}\mathrm{d}W(t)]k\mathrm{e}^{kt}\mathrm{d}t\\ &\quad + [-kx\mathrm{d}t + \sqrt{D}\mathrm{d}W(t)]\mathrm{e}^{kt} + kx\mathrm{e}^{kt}\mathrm{d}t\end{aligned} \tag{4.4.24}$$

注意第一个乘积为零，它仅涉及 $\mathrm{d}t^2$ 和 $\mathrm{d}W(t)\mathrm{d}t$(事实上，可以看清只要将 x 乘上确定性时间函数时总是有这种情况)。我们得到

$$\mathrm{d}y = \sqrt{D}\mathrm{e}^{kt}\mathrm{d}W(t) \tag{4.4.25}$$

积分并重新代入 y，可得

$$x(t) = x(0)\mathrm{e}^{-kt} + \sqrt{D}\int_0^t \mathrm{e}^{-k(t-t')}\mathrm{d}W(t') \tag{4.4.26}$$

如果初始条件是确定性的或是高斯分布的，则 $x(t)$ 显然是高斯变量，具有均值和方差为

$$\langle x(t)\rangle = \langle x(0)\rangle e^{-kt} \tag{4.4.27}$$

$$\mathrm{var}\{x(t)\} = \langle\{[x(0) - \langle x(0)\rangle]e^{-kt} + \sqrt{D}\int_0^t e^{-k(t-t')}dW(t')\}^2\rangle \tag{4.4.28}$$

取初始条件为非可料的，即对 $t>0$ 它与 $dW(t)$ 独立，应用4.2.6f的结论可写出

$$\begin{aligned}\mathrm{var}\{x(t)\} &= \mathrm{var}\{x(0)\}e^{-2kt} + D\int_0^t e^{-2k(t-t')}dt' \\ &= \{\mathrm{var}\{x(0)\} - D/2k\}e^{-2kt} + D/2k\end{aligned} \tag{4.4.29}$$

这些式子与3.8.4节中直接求解福克-普朗克方程所得结果是相同的，只是附加了一个非可料随机初始条件的推广。作为对解为高斯变量这一事实的补充，我们还有着正确的条件概率。

时间相关函数也能直接算出，应是

$$\begin{aligned}\langle x(t), x(s)\rangle &= \mathrm{var}\{x(0)\}e^{-k(t+s)} \\ &\quad + D\left\langle\int_0^t e^{-k(t-t')}dW(t')\int_0^s e^{-k(s-s')}dW(s')\right\rangle \\ &= \mathrm{var}\{x(0)\}e^{-k(t+s)} \\ &\quad + D\int_0^{\min(t,s)} e^{-k(t+s-2t')}dt' \\ &= \left[\mathrm{var}\{x(0)\} - \frac{D}{2k}\right]e^{-k(t+s)} \\ &\quad + \frac{D}{2k}e^{-k|t-s|}\end{aligned} \tag{4.4.30}$$

注意如果 $k>0$，当 $t, s\to\infty$ 而 $|t-s|$ 为有限，则相关函数成为平稳的，且具有3.8.4节中导出的形式。

事实上，如果我们置初始时刻于 $-\infty$ 而不是0，解(4.4.26)成为

$$x(t) = \sqrt{D}\int_{-\infty}^t e^{-k(t-t')}dW(t') \tag{4.4.31}$$

其相关函数和均值显然呈现它们的平稳值。因过程是高斯的，这使它是平稳的。

4.4.5 从直角坐标到极坐标的转换

常常用来描述光学场的模型由一对奥恩斯坦-乌伦贝克过程所给出，它们分别刻划电场的实部和虚部，即

$$\mathrm{d}E_1(t) = -\gamma E_1(t)\mathrm{d}t + \varepsilon\,\mathrm{d}W_1(t) \tag{4.4.32}$$

$$\mathrm{d}E_2(t) = -\gamma E_2(t)\mathrm{d}t + \varepsilon\,\mathrm{d}W_2(t)$$

令人感兴趣的是要转换到极坐标，我们令

$$E_1(t) = a(t)\cos\phi(t) \tag{4.4.33}$$

$$E_2(t) = a(t)\sin\phi(t)$$

为简单起见，再定义

$$\mu(t) = \log a(t) \tag{4.4.34}$$

所以

$$\mu(t) + i\phi(t) = \log[E_1(t) + iE_2(t)] \tag{4.4.35}$$

然后用伊藤微积分导出

$$\begin{aligned}\mathrm{d}(\mu + i\phi) &= \frac{\mathrm{d}(E_1 + iE_2)}{E_1 + iE_2} - \frac{1}{2}\frac{[\mathrm{d}(E_1 + iE_2)]^2}{(E_1 + iE_2)^2} \\ &= -\frac{\gamma(E_1 + iE_2)}{E_1 + iE_2}\,\mathrm{d}t + \frac{\varepsilon[\mathrm{d}W_1(t) + i\mathrm{d}W_2(t)]}{E_1 + iE_2} \\ &\quad - \frac{1}{2}\frac{\varepsilon^2[\mathrm{d}W_1(t) + i\mathrm{d}W_2(t)]^2}{(E_1 + iE_2)^2}\end{aligned} \tag{4.4.36}$$

注意到

$$\mathrm{d}W_1(t)\mathrm{d}W_2(t) = 0,$$

$$\mathrm{d}W_1(t)^2 = \mathrm{d}W_2(t)^2 = \mathrm{d}t,$$

可见最后一项为零，所以我们得到

$$\begin{aligned}\mathrm{d}[\mu(t) + i\phi(t)] = &-\gamma\mathrm{d}t + \varepsilon\exp[-\mu(t) \\ &- i\phi(t)]\{\mathrm{d}W_1(t) + i\mathrm{d}W_2(t)\}\end{aligned} \tag{4.4.37}$$

取出实部，令

$$a(t) = \exp[\mu(t)]$$

并利用伊藤微积分得到

$$\mathrm{d}a(t) = \left\{-\gamma a(t) + \frac{1}{2}\,\varepsilon^2/a(t)\right\}\mathrm{d}t$$

$$+\varepsilon\{dW_1(t)\cos\phi(t)+dW_2(t)\sin\phi(t)\} \qquad (4.4.38)$$

虚部产生

$$d\phi(t)=[\varepsilon/a(t)][-dW_1(t)\sin\phi(t)+dW_2(t)\cos\phi(t)] \qquad (4.4.39)$$

我们现在定义

$$\left.\begin{aligned} dW_a(t)&=dW_1(t)\cos\phi(t)+dW_2(t)\sin\phi(t)\\ dW_\phi(t)&=-dW_1(t)\sin\phi(t)+dW_2(t)\cos\phi(t)\end{aligned}\right\} \qquad (4.4.40)$$

注意这是4.3.5节所指出的那种正交变换，所以我们可以把$dW_a(t)$和$dW_\phi(t)$取为独立维纳过程$W_a(t)$和$W_\phi(t)$的增量。

因此，相位和振幅的随机微分方程是

$$d\phi(t)=[\varepsilon/a(t)]dW_\phi(t) \qquad (4.4.41a)$$

$$da(t)=\left[-\gamma a(t)+\frac{1}{2}\varepsilon^2/a(t)\right]dt+\varepsilon dW_a(t) \qquad (4.4.41b)$$

注解 应用4.3.6节(ii)中所给出的法则，可以将直角坐标方程(4.4.32)和极坐标方程(4.4.41)都变换成斯特拉托诺维奇形式，并发现两者都与伊藤形式完全相同。不过用通常的微积分的直接变换是不可能的，这样做的话直至(4.4.38)式前我们会得到相同的结果，但该式中

$$\left[\frac{1}{2}\varepsilon^2/a(t)\right]dt$$

项不会出现。这必须由一个额外的项来补偿，它是由斯特拉托诺维奇增量$dW_i(t)$与$\phi(t)$是相关的这个事实引起的，这样一来$dW_a(t)$和$dW_\phi(t)$不能简单地用(4.4.40)确定。我们看到了伊藤方法的好处，它保持着$dW(t)$跟在时刻t计算的变量的统计独立性。

令人遗憾的是，虽然相应的直角坐标方程是可解的，但极坐标形式的方程却是不可求解的。然而在处理激光中的极坐标方程时有所好处，其方程是相似的，但在(4.4.41b)中有正比于$a(t)^2dt$的附加项。

4.4.6 多变量奥恩斯坦-乌伦贝克过程

我们用随机微分方程

$$\mathrm{d}\boldsymbol{x}(t) = -\boldsymbol{A}\boldsymbol{x}(t)\mathrm{d}t + \boldsymbol{B}\mathrm{d}\boldsymbol{W}(t) \tag{4.4.42}$$

定义这种过程（A 和 B 是常数矩阵），方程的解是易于求得的（如 4.4.4 节那样）：

$$\boldsymbol{x}(t) = \exp(-\boldsymbol{A}t)\boldsymbol{x}(0) + \int_0^t \exp[-\boldsymbol{A}(t-t')]\boldsymbol{B}\mathrm{d}\boldsymbol{W}(t) \tag{4.4.43}$$

均值是

$$\langle \boldsymbol{x}(t)\rangle = \exp(-\boldsymbol{A}t)\langle \boldsymbol{x}(0)\rangle \tag{4.4.44}$$

相似地可得相关函数

$$\begin{aligned}\langle \boldsymbol{x}(t), \boldsymbol{x}^T(s)\rangle \equiv &\langle [\boldsymbol{x}(t) - \langle \boldsymbol{x}(t)\rangle][\boldsymbol{x}(s) - \langle \boldsymbol{x}(s)\rangle]^T\rangle \\ &= \exp(-\boldsymbol{A}t)\langle \boldsymbol{x}(0), \boldsymbol{x}^T(0)\rangle \exp(-\boldsymbol{A}s) \\ &\quad + \int_0^{\min(t,s)} \exp[-\boldsymbol{A}(t-t')]\boldsymbol{B}\boldsymbol{B}^T \\ &\quad \times \exp[-\boldsymbol{A}^T(s-t')]\mathrm{d}t'\end{aligned} \tag{4.4.45}$$

在某些特殊情况下该积分能用显式形式算出，且对特殊的低维问题，可以直接逐项地全都乘出来。本段的其余部分中我们令

$$\langle \boldsymbol{x}(0), \boldsymbol{x}^T(0)\rangle = 0,$$

相应于确定性初始条件，并计算几个特殊情况。

a) 假设 $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{A}^T\boldsymbol{A}$

由此我们可求出一个么正矩阵 S 使

$$\boldsymbol{S}\boldsymbol{S}^+ = \boldsymbol{1}$$

$$\boldsymbol{S}\boldsymbol{A}\boldsymbol{S}^+ = \boldsymbol{S}\boldsymbol{A}^T\boldsymbol{S}^+ = \mathrm{diag}(\lambda_1, \lambda_2, \cdots\lambda_n) \tag{4.4.46}$$

为简明起见设 $t \geqslant s$，则有

$$\langle \boldsymbol{x}(t), \boldsymbol{x}^T(s)\rangle = \boldsymbol{S}^+\boldsymbol{G}(t, s)\boldsymbol{S}$$

其中

$$[G(t, s)]_{ij} = \frac{(BB^T)_{ij}}{\lambda_i + \lambda_j}[\exp(-\lambda_i|t-s|)$$

$$-\exp(-\lambda_i t-\lambda_j s)]^{(*)} \tag{4.4.47}$$

b) 平稳解中的方差

如果A仅有实部为正的特征值，则存在

$$x_s(t)=\int_{-\infty}^{t}\exp[-A(t-t')]B\mathrm{d}W(t) \tag{4.4.48}$$

形式的平稳解。当然我们有

$$\left.\begin{aligned}&\langle x_s(t)\rangle=0\\ &\text{以及}\\ &\langle x_s(t),\ x_s^T(s)\rangle\\ &=\int_{-\infty}^{\min(t,s)}\exp[-A(t-t')]BB^T\exp[-A^T(s-t')]\mathrm{d}t'\end{aligned}\right\} \tag{4.4.49}$$

定义平稳协方差矩阵σ为

$$\sigma=\langle x_s(t),\ x_s^T(t)\rangle \tag{4.4.50}$$

这个量的计算可用代数方法作出：

$$\begin{aligned}&A\sigma+\sigma A^T\\ &=\int_{-\infty}^{t}A\exp[-A(t-t')]BB^T\exp[-A^T(t-t')]\mathrm{d}t'\\ &\quad+\int_{-\infty}^{t}\exp[-A(t-t')BB^T\exp[-A^T(t-t')]A^T\mathrm{d}t'\\ &=\int_{-\infty}^{t}\frac{\mathrm{d}}{\mathrm{d}t'}\{\exp[-A(t-t')]BB^T\exp[-A^T(t-t')]\}\mathrm{d}t'\end{aligned}$$

进行积分，由A的特征值实部为正的假设可知下限为零，因而仅留下上限，给出了

$$\boxed{A\sigma+\sigma A^T=BB^T} \tag{4.4.51}$$

作为平稳协方差矩阵所满足的代数方程。

c) 二维平稳方差

注意到如果A是2×2矩阵，它就满足特征方程

$$A^2-(\mathrm{Tr}\,A)A+(\mathrm{Det}\,A)I=0 \tag{4.4.52}$$

又由(4.4.49)，以及(4.4.52)蕴涵着$\exp(-At)$是关于A的一次

(*) 式中$(BB^T)_{ij}$应为$(SBB^TS^+)_{ij}$. ——译注

多项式的事实，我们一定可以写出

$$\sigma = \alpha BB^T + \beta(ABB^T + BB^TA^T) + \gamma ABB^TA^T$$

应用(4.4.52)，我们发现如果

$$\alpha + (\mathrm{Tr}\ A)\beta - (\mathrm{Det}\ A)\gamma = 0$$

$$2\beta(\mathrm{Det}\ A) + 1 = 0$$

$$\beta + (\mathrm{Tr}\ A)\gamma = 0$$

则(4.4.51)是满足的，由此得

$$\boxed{\sigma = \frac{(\mathrm{Det}\ A)BB^T + [A - (\mathrm{Tr}\ A)I]BB^T[A - (\mathrm{Tr}\ A)I]^T}{2(\mathrm{Tr}\ A)(\mathrm{Det}\ A)}} \tag{4.4.53}$$

d) 平稳态时间相关矩阵

由解答(4.4.49)，我们看到对 $t > s$ 有

$$\langle x_s(t), x_s^T(s)\rangle = \exp[-A(t-s)]\int_{-\infty}^{s} \exp[-A(s-t')]BB^T \times \exp[-A^T(s-t')]\mathrm{d}t'$$

$$= \exp[-A(t-s)]\ \sigma \quad t > s \tag{4.4.54a}$$

类似地又有

$$= \sigma \exp[-A^T(s-t)] \quad t < s \tag{4.4.54b}$$

正如平稳解的预期的形式，它们给出了对 $s-t$ 的依赖性。然后定义

$$G_s(t-s) = \langle x_s(t), x_s^T(s)\rangle \tag{4.4.55}$$

可见(记住 $\sigma = \sigma^T$)

$$G_s(t-s) = [G_s(s-t)]^T \tag{4.4.56}$$

e) 平稳态谱矩阵

谱矩阵其实是相当简单的，类似于 1.4.2 节我们定义

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-i\omega\tau}G_s(\tau)\mathrm{d}\tau \tag{4.4.57}$$

$$= \frac{1}{2\pi}\left\{\int_0 \exp[-(i\omega + A)\tau]\sigma\mathrm{d}\tau\right.$$

$$+\int_{-\infty}^{0}\sigma\exp[(-i\omega+A^T)\tau]\mathrm{d}\tau\Big\}$$

$$=\frac{1}{2\pi}[(A+i\omega)^{-1}\sigma+\sigma(A^T-i\omega)^{-1}]$$

因此，

$$(A+i\omega)S(\omega)(A^T-i\omega)=\frac{1}{2\pi}(\sigma A^T+A\sigma),$$

再利用(4.4.51)我们得到

$$\boxed{S(\omega)=\frac{1}{2\pi}(A+i\omega)^{-1}BB^T(A^T-i\omega)^{-1}} \tag{4.4.58}$$

f) 回归定理

(4.4.54a)的结论也称为回归定理，因为它断言对 $\tau>0$ 时间发展 $G(\tau)$ 受到与均值时间发展相同的规律［即(4.4.44)］所支配。这是该问题的马尔可夫线性性质的一个推论。因为

$$\frac{\mathrm{d}}{\mathrm{d}\tau}[G_s(\tau)]\mathrm{d}\tau=\frac{\mathrm{d}}{\mathrm{d}\tau}\langle\boldsymbol{x}_s(\tau),\boldsymbol{x}_s^T(0)\rangle\mathrm{d}\tau$$

$$=\langle[-A\boldsymbol{x}_s(\tau)\mathrm{d}\tau+B\mathrm{d}\boldsymbol{W}(\tau)],\boldsymbol{x}_s^T(0)\rangle$$

又因 $\tau>0$ 时 $\mathrm{d}\boldsymbol{W}(\tau)$ 与 $\boldsymbol{x}_s^T(0)$ 不相关，故有

$$\frac{\mathrm{d}}{\mathrm{d}\tau}[G_s(\tau)]=-AG_s(\tau) \tag{4.4.59}$$

这样，$G_s(\tau)$ 的计算需要 $G_s(0)$ 的知识和均值的时间发展方程。这个结果与 3.7.4 节中的结果是相仿的。

4.4.7 一般的单变量线性方程

a) 齐次场合

我们首先考虑齐次场合

$$\mathrm{d}x=[b(t)\mathrm{d}t+g(t)\mathrm{d}W(t)]x \tag{4.4.60}$$

应用常用的伊藤法则，写出

$$y=\log x \tag{4.4.61}$$

所以

$$dy = dx/x - \frac{1}{2}(dx)^2/x^2$$

$$= [b(t)dt + g(t)dW(t)] - \frac{1}{2}g(t)^2dt \qquad (4.4.62)$$

积分并作(4.4.61)的逆代换，可得

$$x(t) = x(0)\exp\left\{\int_0^t\left[b(t') - \frac{1}{2}g(t')^2\right]dt' + \int_0^t g(t')dW(t')\right\} \qquad (4.4.63)$$

$$\equiv x(0)\phi(t) \qquad (4.4.64)$$

式中已定义了$\phi(t)$。

注意[应用(4.4.8)]

$$\langle[x(t)]^n\rangle = \langle[x(0)]^n\rangle\left\langle\exp\left\{n\int_0^t\left[b(t') - \frac{1}{2}g(t')^2\right]dt' + n\int_0^t g(t')dW(t')\right\}\right\rangle$$

$$= \langle[x(0)]^n\rangle\exp\left\{n\int_0^t b(t')dt' + \frac{1}{2}n(n-1)\int_0^t g(t')^2dt'\right\} \qquad (4.4.65)$$

b) 非齐次场合

现在考虑

$$dx = [a(t) + b(t)x]dt + [f(t) + g(t)x]dW(t) \qquad (4.4.66)$$

并且令

$$z(t) = x(t)[\phi(t)]^{-1} \qquad (4.4.67)$$

其中$\phi(t)$如(4.4.64)所定义，作为齐次方程(4.4.60)的一个解。然后我们写出

$$dz = dx[\phi(t)]^{-1} + xd[\phi(t)^{-1}] + dx[d\phi(t)^{-1}]$$

注意到

$$d[\phi(t)^{-1}] = -d\phi(t)[\phi(t)]^{-2} + [d\phi(t)]^2[\phi(t)]^{-3}$$

并利用伊藤法则，可得

$$dz = \{[a(t) - f(t)g(t)]dt + f(t)dW(t)\}\phi(t)^{-1} \qquad (4.4.68)$$

它可直接求积。因此，解为

$$x(t) = \phi(t)\left\{x(0) + \int_0^t \phi(t')^{-1}\{[a(t') - f(t')g(t')]\mathrm{d}t' + f(t')\mathrm{d}W(t')\right\} \tag{4.4.69}$$

c) 矩和自相关

在导出矩的方程时，由(4.4.66)出发要比从解(4.4.69)直接计算矩和自相关更好些，因为我们有

$$\begin{aligned}\mathrm{d}[x(t)^n] &= nx(t)^{n-1}\mathrm{d}x(t) + \frac{n(n-1)}{2}x(t)^{n-2}[\mathrm{d}x(t)]^2 \\ &= nx(t)^{n-1}\mathrm{d}x(t) + \frac{n(n-1)}{2}x(t)^{n-2}[f(t) \\ &\quad + g(t)x(t)]^2\mathrm{d}t\end{aligned} \tag{4.4.70}$$

因此

$$\boxed{\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle x(t)^n\rangle &= \langle x(t)^n\rangle\left[nb(t) + \frac{n(n-1)}{2}g(t)^2\right] \\ &\quad + \langle x(t)^{n-1}\rangle[na(t) + n(n-1)f(t)g(t)] \\ &\quad + \langle x(t)^{n-2}\rangle\left[\frac{n(n-1)}{2}f(t)^2\right]\end{aligned}} \tag{4.4.71}$$

这些方程构成了一个层次体系，其中第 n 个方程涉及前两个方程的解，可以依次进行积分。

4.4.8 多变量线性方程

a) 齐次场合

方程为

$$\mathrm{d}x(t) = \left[B(t)\mathrm{d}t + \sum_i G_i(t)\mathrm{d}W_i(t)\right]x(t) \tag{4.4.72}$$

这里 $B(t)$、$G_i(t)$ 是矩阵。一般情况下这个方程没有封闭形式的解，除非所有矩阵 $B(t)$、$G_i(t)$ 之间在所有时刻都可对易，即

$$\begin{aligned}&G_i(t)G_j(t') = G_j(t')G_i(t) \\ &B(t)G_i(t') = G_i(t')B(t) \\ &B(t)B(t') = B(t')B(t)\end{aligned} \tag{4.4.73}$$

这时，解完全相仿于单变量场合的解，且有

$$x(t) = \Phi(t)x(0)$$

其中

$$\Phi(t) = \exp\left\{\int_0^t \left[B(t) - \frac{1}{2}\sum_i G_i(t)^2\right]dt + \int_0^t \sum_i G_i(t)dW_i(t)\right\} \tag{4.4.74}$$

b) 非齐次场合

采用跟一维情况完全相同的方法，我们可以把非齐次情况化成为齐次情况．这样，我们考虑

$$dx(t) = [A(t) + B(t)x]dt + \sum_i [F_i(t) + G_i(t)x]dW_i(t) \tag{4.4.75}$$

再令

$$y(t) = \psi(t)^{-1}x(t) \tag{4.4.76}$$

这里 $\psi(t)$是齐次方程 (4.4.72) 的一个矩阵解．我们先必须计算 $d[\psi^{-1}]$．因对任何矩阵有$MM^{-1}=1$，所以，展开至二阶，

$$Md[M^{-1}] + dM\ M^{-1} + dMd[M^{-1}] = 0$$

因此

$$d[M^{-1}] = -[M + dM]^{-1}dM\,M^{-1}$$

再次取到二阶

$$d[M^{-1}] = -M^{-1}dM\ M^{-1} + M^{-1}dM\ M^{-1}dM\ M^{-1} \tag{4.4.77}$$

这样一来，因为$\psi(t)$满足齐次方程

$$d[\psi(t)^{-1}] = \psi(t)^{-1}\{[-B(t) + \sum_i G_i(t)^2]dt - \sum_i G_i(t)dW_i(t)\}$$

再次取微分

$$dy(t) = \psi(t)^{-1}\{[A(t) - \sum_i G_i(t)F_i(t)]dt + \Sigma F_i(t)dW_i(t)\}$$

因此

$$x(t) = \psi(t)\{x(0) + \int_0^t \psi(t')^{-1}\{[A(t')$$

$$-\sum_i G_i(t')F_i(t')]\mathrm{d}t' + \Sigma F_i(t')\mathrm{d}W_i(t')\}\} \tag{4.4.78}$$

就实用目的而言，因为难以计算均值和相关函数，即使已知齐次解，这个解也是不太有用的。

4.4.9 与时间有关的奥恩斯坦-乌伦贝克过程

这是前述的可求解的一般线性方程的特殊情况，它将多变量奥恩斯坦-乌伦贝克过程（4.4.6节）推广为包含与时间有关的参量，即

$$\mathrm{d}\boldsymbol{x}(t) = -\boldsymbol{A}(t)\boldsymbol{x}(t)\mathrm{d}t + \boldsymbol{B}(t)\mathrm{d}\boldsymbol{W}(t) \tag{4.4.79}$$

显然在替换

$$\begin{aligned} &\boldsymbol{A}(t)\to 0\\ &\boldsymbol{B}(t)\to -\boldsymbol{A}(t)\\ &\sum_i F_i(t)\mathrm{d}W_i(t)\to \boldsymbol{B}(t)\mathrm{d}\boldsymbol{W}(t)\\ &G_i(t)\to 0 \end{aligned} \tag{4.4.80}$$

之下(4.4.75)与这方程相同。相应的齐次方程就是确定性方程

$$\mathrm{d}\boldsymbol{x}(t) = -\boldsymbol{A}(t)\boldsymbol{x}(t)\mathrm{d}t \tag{4.4.81}$$

在 $\boldsymbol{A}(t)\boldsymbol{A}(t') = \boldsymbol{A}(t')\boldsymbol{A}(t)$ 的前提下它是可解的，具有解

$$\boldsymbol{x}(t) = \boldsymbol{\psi}(t)\boldsymbol{x}(0)$$

其中

$$\boldsymbol{\psi}(t) = \exp\left[-\int_0^t \boldsymbol{A}(t')\mathrm{d}t'\right] \tag{4.4.82}$$

这样，应用(4.4.78)，

$$\begin{aligned} \boldsymbol{x}(t) = &\exp\left[-\int_0^t \boldsymbol{A}(t')\mathrm{d}t'\right]\boldsymbol{x}(0)\\ &+\int_0^t \left\{\exp\left[-\int_{t'}^t \boldsymbol{A}(s)\mathrm{d}s\right]\right\}\boldsymbol{B}(t')\mathrm{d}\boldsymbol{W}(t') \end{aligned} \tag{4.4.83}$$

这同4.4.6节中导出的不依赖于时间的奥恩斯坦-乌伦贝克过程的解(4.4.43)非常相象。

由此我们有

$$\langle x(t)\rangle = \exp\left[-\int_0^t A(t')\mathrm{d}t'\right]\langle x(0)\rangle \tag{4.4.84}$$

$$\begin{aligned}&\langle x(t), x^T(t)\rangle\\ &= \exp\left[-\int_0^t A(t')\mathrm{d}t'\right]\langle x(0), x(0)^T\rangle \exp\left[-\int_0^t A^T(t')\mathrm{d}t'\right]\\ &\quad + \int_0^t \mathrm{d}t' \exp\left[-\int_{t'}^t A(s)\mathrm{d}s\right] B(t')B^T(t')\exp\left[-\int_{t'}^t A^T(s)\mathrm{d}s\right]\end{aligned} \tag{4.4.85}$$

联系于低噪声系统中渐近方法的开发，与时间有关的奥恩斯坦-乌伦贝克过程将会自然而然地产生。

5. 福克-普朗克方程

在篇幅较长的本章中，我们要从相应的福克-普朗克方程的角度来展开连续马尔科夫过程的理论，该方程对系统给出概率密度函数的时间演化。本章分为两个主要主题——单变量过程与多变量过程。单变量系统有许多精确的结果，因此把它分离出来单独处理是恰当的。因而在5.2节及其各小节全面地考察单变量系统，在5.3节则考察多变量系统。不过有一点，即构造适当边界条件的问题对两种情况都具有基本的重要性，它放在5.2.1节中总起来进行考察。在5.2.4节则对后向福克-普朗克方程中的边界条件给出了相应的处理。在5.2节的其他小节中所集中考察的是有关平稳分布函数、本征函数的性质以及出口问题的一些精确结果，它们绝大部分在单变量情况下是能够明确而清晰地得以解决的。

5.3节及其各小节要说明多变量系统的一些精确结果。这些结果并不如单变量情况时那样清晰。这里还包含了在多变量系统中所具有的另一个特征，这就是细致平衡的概念，它在单变量情况几乎是显然的，但在多变量系统里却导致了一些十分有趣的结论。

作为本章的结束，还对多变量福克-普朗克方程阐述了出口问题中的一些精确结果。

5.1 背景

我们已经在好几处遇到过福克-普朗克方程。开头在爱因斯坦对扩散方程的原始推导和应用那里(1.2节)，再者(3.5.2节)是

作为查普曼-柯尔莫哥洛夫微分方程的一个特例，最后(4.3.4节)是随机微分方程，都与福克-普朗克方程相关联。有许多与应用福克-普朗克方程相联系的技巧可以比利用相应的随机微分方程更直接地获得一些结果；反之亦然。所以为了获得扩散过程特性的完整而生动的描述，我们必须同时研究这两种观点。

"福克-普朗克方程"这名称起源于福克(Fokker)(1914)[5.1]及普朗克(Planck)(1917)[5.2]的工作，前者研究了辐射场中的布朗运动，后者则企图基于它来建立一个关于涨落的完整理论。由于柯尔莫哥洛夫的工作为它建立了严格的数学基础[5.3]，许多专业数学方面的著作倾向于使用"柯尔莫哥洛夫方程"这个术语。又由于斯莫路乔夫斯基首先用了这个方程，有些人也用"斯莫路乔夫斯基方程"这个术语。我不对术语的是非曲直作任何评价，而只想把"福克-普朗克方程"作为本书所面向的读者所通常使用的术语。

5.2 一维福克-普朗克方程

在一维情况下，福克-普朗克方程具有简单的形式

$$\frac{\partial f(x,t)}{\partial t}=-\frac{\partial}{\partial x}[A(x,t)f(x,t)]+\frac{1}{2}\frac{\partial^2}{\partial x^2}[B(x,t)f(x,t)] \tag{5.2.1}$$

在第3.4.5节已经说明对于条件概率福克-普朗克方程也是有效的，这就是说，对于任何初值x_0, t_0，可取

$$f(x,t)=p(x,t|x_0,t_0) \tag{5.2.2}$$

它的初始条件为

$$p(x,t_0|x_0,t_0)=\delta(x-x_0) \tag{5.2.3}$$

但是，利用一下概率的定义，就有

$$p(x,t)=\int \mathrm{d}x_0 p(x,t;x_0,t_0)$$

$$=\int \mathrm{d}x_0 p(x,t|x_0,t_0)p(x_0,t_0) \tag{5.2.4}$$

可看出对于具有初始条件

$$p(x,\ t)|_{t=t_0} = p(x,\ t_0) \tag{5.2.5}$$

的 $p(x,\ t)$，福克-普朗克方程也成立，而这个初始条件的奇异性一般说比(5.2.3)要少些。

由第 4.3.4 节我们知道满足福克-普朗克方程的条件概率所刻划的随机过程等价于伊藤随机微分方程

$$\mathrm{d}x(t) = A[x(t),\ t] + \sqrt{B[x(t),\ t]}\mathrm{d}W(t)$$

而且这两种刻划被认为是彼此相互补充的。下面我们将会看到基于福克-普朗克方程的摄动理论十分不同于基于随机微分方程的摄动理论，两者都各有它们的用处。

5.2.1 边界条件

福克-普朗克方程是一个二阶抛物型偏微分方程，为了求解，我们要有诸如 (5.2.5) 的初始条件及在约束 x 的区间的端点上的边界条件，这些条件具有多种的形式。

一般地来导出边界条件比限于考察单变量情况要简单些。现考察前向方程

$$\begin{aligned}\partial_t p(z,\ t) = &-\sum_i \frac{\partial}{\partial z_i} A_i(z,\ t)p(z,\ t) \\ &+ \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial z_i \partial z_j} B_{ij}(z,\ t)p(z,\ t)\end{aligned} \tag{5.2.6}$$

我们注意，它也可以写成

$$\frac{\partial p(z,\ t)}{\partial t} + \sum_i \frac{\partial}{\partial z_i} J_i(z,\ t) = 0 \tag{5.2.7}$$

其中我们定义了概率流

$$J_i(z,\ t) = A_i(z,\ t)p(z,\ t) - \frac{1}{2}\sum_j \frac{\partial}{\partial z_j} B_{ij}(z,\ t)p(z,\ t) \tag{5.2.8}$$

方程(5.2.7)具有局部守恒方程的形式，能够按下述方法写成一个积分的形式。考虑某个以 S 为边界的区域 R，定义

$$P(R,\ t)=\int_R \mathrm{d}x p(z,\ t)$$

则(5.2.7)等价于

$$\frac{\partial P(R,\ t)}{\partial t}=-\int_S \mathrm{d}S\boldsymbol{n}\cdot\boldsymbol{J}(z,\ t) \tag{5.2.9}$$

其中 $\boldsymbol{n}$ 是 S 的向外法线。这样，(5.2.9)就表明了概率的总丧失是由 $\boldsymbol{J}$ 在 R 的边界上的曲面积分给定的。我们还可以指出流 $\boldsymbol{J}$ 事实上具有某种更强的性质，即在任何一个曲面 S 上的积分给出了通过这个曲面的概率的净流。

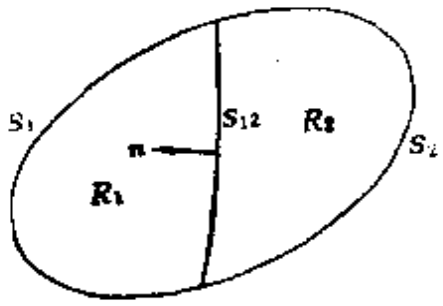


图 5.1 用以图示概率流是概率的净流的区域

这只要考虑两个邻接的由曲面 S_{12} 隔开着的区域 R_1 及 R_2。设 S_1 及 S_2 是与 S_{12} 一起分别包围着 R_1 及 R_2 的曲面(见图 5.1)。

只要注意到这里要考察的是具有连续样本路径的过程，因而在一个充分小的时间间隔 Δt 中由 R_2 穿过 S_{12} 而到 R_1 中的概率量就是 t 时刻于 R_2 中而 $t+\Delta t$ 时刻于 R_1 中的联合概率，它等于

$$\int_{R_1}\mathrm{d}x\int_{R_2}\mathrm{d}y p(x,\ t+\Delta t;\ y,\ t)$$

我们就可以来计算概率的净流量了。从 R_2 到 R_1 的概率净流量是由这个概率减去相反方向穿过的概率，然后被 Δt 去除得到的，即

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{R_1}\mathrm{d}x\int_{R_2}\mathrm{d}y[p(x,\ t+\Delta t;\ y,\ t)$$
$$-p(y,\ t+\Delta t;\ x,\ t)] \tag{5.2.10}$$

注意有等式

$$\int_{R_1}\mathrm{d}x\int_{R_2}\mathrm{d}y p(x,\ t;\ y,\ t)=0$$

这是因为等式左边为同时在 R_1 和 R_2 中的概率。于是可以改写

$$(5.2.10)=\int_{R_1}\mathrm{d}x\int_{R_2}\mathrm{d}y[\partial_{t'}p(x,\ t';\ y,\ t)$$
$$-\partial_{t'}p(y,\ t';\ x,\ t)]_{t'=t}$$

再利用(5.2.7)形式下的福克-普朗克方程，上式

$$= -\int_{R_2} \mathrm{d}x \sum_i \frac{\partial}{\partial x_i} J_i(x,\ t;\ R_2,\ t)$$

$$+\int_{R_2} \mathrm{d}y \sum_i \frac{\partial}{\partial y_i} J_i(y,\ t;\ R_1,\ t) \qquad (5.2.11)$$

其中 $J_i(x,\ t;\ R_2,\ t)$是由

$$p(x,\ t,\ R_2,\ t) = \int_{R_2} \mathrm{d}y p(x,\ t,\ y,\ t)$$

按照 (5.2.8) 中由 $p(z,\ t)$ 给出 $J(z,\ t)$ 的相同方式所给定，$J_i(y,\ t;\ R_1,\ t)$也类似地定义。现在我们可以把这些积分转换为曲面上的积分。在 S_2 上的积分中包含了因子 $p(x,\ t,\ R_2,\ t)$，其中 x(除了一个零测集外)不在 R_2 中或者在其边界上，所以这个积分为零，同样在 S_1 上的积分也为零，但是在 S_{12} 上的那些积分却不是零，因为此时积分是在 R_1 与 R_2 的共同边界上求积的。

这样一来，由 R_2 到 R_1 的净流量为

$$\int_{S_{12}} \mathrm{d}S\, n \cdot \{J(x,\ t;\ R_1,\ t) + J(x,\ t;\ R_2,\ t)\}$$

并且由于 x 属于 R_1 与 R_2 的交集 (*)，我们最后得知单位时间内从 R_2 到 R_1 的概率净流量

$$\boxed{\begin{aligned} &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{R_1} \mathrm{d}x \int_{R_2} \mathrm{d}y \left[p(x,\ t+\Delta t;\ y,\ t) \right. \\ &\left. \quad - p(y,\ t+\Delta t;\ x,\ t) \right] = \int_{S_{12}} \mathrm{d}S n \cdot J(x,\ t) \\ &\text{其中 } n \text{ 指向为从 } R_2 \text{ 到 } R_1 \end{aligned}} \qquad (5.2.12)$$

我们现在可以来一一考察几种边界条件如下：

a) 反射壁

可以认为这是粒子不能离开区域 R 时的情况，因此穿过 R 的边界 S 的概率流量是零。于是我们要求

$$n \cdot J(z,\ t) = 0,\ z \in S,\ n = S \text{ 的法线} \qquad (5.2.13)$$

(*) 这里原文有误，原文是"R_1 与 R_2 的并集"，—— 译注

其中 $J(x, t)$ 由(5.2.8)所给定。

因为粒子不能穿过 S，它必然在那里被反射，因而对这种条件赋于名称反射壁。

b) 吸收壁

此时我们假设一旦粒子到达了 S，从那时起它就从系统中消除了，因而其壁是吸收壁，因此在边界上的概率是零，也就是说

$$p(x, t) = 0, \quad x \in S \tag{5.2.14}$$

c) 在间断点的边界条件

系数 A_i 以及 B_{ij} 两者都可能在一曲面 S 上不连续，但是由于在穿过 S 时是自由运动，因而在穿过 S 时，流的法线方向分量及概率量都应该是连续的，即

$$n \cdot J(x)\,|_{S_+} = n \cdot J(x)\,|_{S_-} \tag{5.2.15}$$

$$p(x)\,|_{S_+} = p(x)\,|_{S_-} \tag{5.2.16}$$

其中作为下标 S_+ 及 S_- 的含义是分别表示从 S 的左边和右边趋近 S 时有关量的极限值。

d) 周期性边界条件

我们假设过程在某一区间[a, b]上发生，而此区间的两端点是彼此重合的（例如当扩散过程是在一圆周上发生时就出现了这种情况）。那末我们可以加上由不连续点处的边界条件所导出的边界条件，即

$$\text{I}: \lim_{x \to b-} p(x, t) = \lim_{x \to a+} p(x, t) \tag{5.2.17}$$

$$\text{II}: \lim_{x \to b-} J(x, t) = \lim_{x \to a+} J(x, t) \tag{5.2.18}$$

要加上周期性边界条件的情况最常见的是函数 $A(x, t)$ 与 $B(x, t)$ 都是这同一区间上的周期函数，因而

$$A(b, t) = A(a, t) \tag{5.2.19}$$

$$B(b, t) = B(a, t)$$

这意味着 I 与 II 简化为 $p(x, t)$ 及其导数在点 a 及 b 处的一个等式。

e) 规定好了的边界

如果扩散系数在一边界上为零，则我们有了一个边界类别可以被自动规定好的情况。假设运动仅发生在 $x>a$ 的区域中，如果 $A(x, t)$ 及 $\sqrt{B(x, t)}$ 在 $x=a$ 时满足李普希兹条件（见第 4.3.1 节），而且 $B(x, t)$ 在 $x=a$ 处是可微的，则

$$\partial_x B(a, t)=0 \tag{5.2.20}$$

而随机微分方程有解，并且可以记为

$$\mathrm{d}x(t)=A(x, t)\mathrm{d}t+\sqrt{B(x, t)}\mathrm{d}W(t) \tag{5.2.21}$$

在这种较特殊的情况下，$A(x, t)$ 的符号决定着各种情况，有如下三种情况出现。

i) 出口边界。在这种情况下我们假设

$$A(a, t)<0 \tag{5.2.22}$$

使得只要粒子到达了点 a，则它就必然要走出区域而到 $x<a$ 中去，因此边界叫做“出口边界”。

ii) 入口边界。我们假设

$$A(a, t)>0 \tag{5.2.23}$$

在这种情况下，若粒子到达点 a，则 $A(a, t)$ 的符号就表明了它要再回到 $x>a$ 的状态，故一个位于 a 右方的粒子决不可能离开这个区域，不过一个在 $x=a$ 处引进的粒子则必定会进入这个区域中。所以这时叫做“入口边界”。

iii) 自然边界。最后来考察

$$A(a, t)=0 \tag{5.2.24}$$

的情况，此时粒子一旦到达 $x=a$ 则它将停留在那里。但是可以证明粒子根本到不了这一点。这是一种既不能从它那里吸收到任何粒子又不能从它那里引进任何粒子的边界。

费勒[5.4]指出了边界一般可以分成四种类型，这就是正则边界、入口边界、出口边界及自然边界。他的边界分类一般判别法如下。定义

$$f(x)=\exp\left[-2\int_{x_0}^{x} \mathrm{d}s A(s)/B(s)\right] \tag{5.2.25}$$

$$g(x)=2/[B(x)f(x)] \tag{5.2.26}$$

$$h_1(x) = f(x)\int_{x_0}^{x} g(s)\mathrm{d}s \tag{5.2.27}$$

$$h_2(x) = g(x)\int_{x}^{x} f(s)\mathrm{d}s \tag{5.2.28}$$

这里 $x_0 \in (a, b)$ 是已给定的. 若用

$$\mathscr{L}(x_1, x_2) \tag{5.2.29}$$

表示在区间 (x_1, x_2) 上的可积函数全体所组成的空间，则边界点 a 可分类如下：

I. 正则的：	若 $f(x) \in \mathscr{L}(a, x_0)$ 且 $g(x) \in \mathscr{L}(a, x_0)$
II. 出口：	若 $g(x) \notin \mathscr{L}(a, x_0)$ 且 $h_1(x) \in \mathscr{L}(a, x_0)$
III. 入口：	若 $g(x) \in \mathscr{L}(a, x_0)$ 且 $h_2(x) \in \mathscr{L}(a, x_0)$
IV. 自然的：	所有其他情况

从下面第5.2.2节的结果可以看出，对于一个出口边界来说不存在福克-普朗克方程的可归一化平稳解，并且由(5.2.161)所表示的到达此边界的平均时间是有限值. 类似地，如果一个边界是出口，可以存在一个平稳解，不过到达此边界的平均时间是无穷大. 当情况是一个正则边界时，到达边界的平均时间是有限值，但是也就存在一个具有反射边界 a 的平稳解. 自然边界的分析是较费力的. 读者可以在文献[5.5]看到一个较完整的讨论.

f) 在无穷远处的边界

所有上面提到过的边界类型都可能出现在无穷远处，这只要我们可以同时确保概率的归一化. 若 $p(x)$ 的变化是适当程度光滑的，这就要求

$$\lim_{x\to\infty} p(x, t) = 0 \tag{5.2.30}$$

如果 $\partial_x p(x)$ 的变化是适当程度光滑的（也就是说当 $x\to\infty$ 时不是快速地作无穷次振动），则

$$\lim_{x\to\infty} \partial_x p(x, t) = 0 \tag{5.2.31}$$

所以一个在无穷远处非零的流通常要求 $A(x, t)$ 或者 $B(x, t)$ 在无穷远处要趋向无穷，对于这种情况最好的处理办法是把它改变

为在 $x=\infty$ 处取有限值的另一个变量来进行。

当允许在 $x=\pm\infty$ 处存在边界且在无穷远处有非零流时，有两种可能性使得我们不允许失去概率：

i) $J(\pm\infty, t)=0$ (5.2.32)

ii) $J(+\infty, t)=J(-\infty, t)$ (5.2.33)

它们分别是反射边界条件和周期边界条件的极限情况。

5.2.2 齐次福克-普朗克方程的平稳解

先回顾一下。在3.7.2节中曾说过，在一个齐次过程中漂移系数及扩散系数是与时间独立的，在此情况下平稳分布所满足的方程是

$$\frac{\mathrm{d}}{\mathrm{d}x}[A(x)p_s(x)]-\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}[B(x)p_s(x)]=0 \qquad (5.2.34)$$

它也可以简单地用(如5.2.1节中定义的)流改写为

$$\frac{\mathrm{d}J(x)}{\mathrm{d}x}=0 \qquad (5.2.35)$$

显然它有解

$$J(x)=\text{常数} \qquad (5.2.36)$$

假设过程发生于区间(a, b)中，则我们就必然有

$$J(a)=J(x)=J(b)\equiv J \qquad (5.2.37)$$

并且只要在边界条件中有一个是反射边界条件，则它就意味着两个边界都是反射的，而且 $J=0$。

如果这些边界不是反射的，则(5.2.37)就要求它们是周期的，此时我们就要用(5.2.17, 18)所给出的边界条件。

a) 零流——位势解

置 $J=0$，就可把(5.2.37)改写为

$$A(x)p_s(x)-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x}[B(x)p_s(x)]=0 \qquad (5.2.38)$$

它的解是

$$p_s(x)=\frac{\mathcal{N}}{B(x)}\exp\left[2\int_a^x \mathrm{d}x'\frac{A(x')}{B(x')}\right] \qquad (5.2.39)$$

其中N是使得

$$\int_a^b \mathrm{d}x p_s(x) = 1 \tag{5.2.40}$$

的归一化常数，这种解作为位势解是为人们所熟悉的，这有各种历史原因，但主要是由于这个平稳解是由一个单一积分得到的（在5.3.3节将会考察这个术语的完整含意）。

b) 周期性边界条件

此时是非零流，而(5.2.36)则改写为

$$A(x)p_s(x) - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x}[B(x)p_s(x)] = J \tag{5.2.41}$$

但是J不是任意给定的，而是要由归一化及周期的边界条件

$$p_s(a) = p_s(b) \tag{5.2.42}$$

$$J(a) = J(b) \tag{5.2.43}$$

所确定。为了方便起见定义

$$\psi(x) = \exp\left[2\int_a^x \mathrm{d}x' \frac{A(x')}{B(x')}\right] \tag{5.2.44}$$

于是容易经积分(5.2.41)来得到

$$p_s(x)B(x)/\psi(x) = p_s(a)B(a)/\psi(a) + J\int_a^x \frac{\mathrm{d}x'}{\psi(x')} \tag{5.2.45}$$

我们再加上边界条件(5.2.42)之后，就看到有

$$J = [B(b)/\psi(b) - B(a)/\psi(a)]p_s(a)\Big/\left[\int_a^b \frac{\mathrm{d}x'}{\psi(x')}\right] \tag{5.2.46}$$

从而

$$p_s(x) = p_s(a)\left[\frac{\int_a^x \frac{\mathrm{d}x'}{\psi(x')} \frac{B(b)}{\psi(b)} + \int_x^b \frac{\mathrm{d}x'}{\psi(x')} \frac{B(a)}{\psi(a)}}{\frac{B(x)}{\psi(x)}\int_a^b \frac{\mathrm{d}x'}{\psi(x')}}\right] \tag{5.2.47}$$

c) 无穷值域及奇异边界

在这两种情况的任一种下，由于发散等等原因，前面的各种可

能性中，总有这种或那种被证明是行不通了。不过要把这些可能性仔细列举出来通常是十分复杂的，我们将利用下一节中给出的一些例子来证实它。

5.2.3 一些平稳解的例子

a) 重力场中的扩散

不变重力场中的强阻力布朗粒子常常由随机微分方程

$$\mathrm{d}x = -g\mathrm{d}t + \sqrt{D}\,\mathrm{d}W(t) \tag{5.2.48}$$

来描述(见 6.4 节)，其福克-普朗克方程是

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}(gp) + \frac{1}{2}D\frac{\partial^2 p}{\partial x^2} \tag{5.2.49}$$

在具有反射边界条件的区间(a, b)上，平稳解由(5.2.39)来给出，即

$$p_s(x) = \mathcal{N}\exp[-2gx/D] \tag{5.2.50}$$

这里我们已经在$\mathcal{N}$的定义中吸收了常数因子。

显然只要a是有限值，不管B有可能是无穷值，此解就是归一的。这个结果并不比下列说明更深奥，即在一烧杯液体中作扩散的一些粒子将会落下去，并且假如烧杯是无限深的，它们就不会停止这种下落！反抗重力的向上扩散是可能发生的，扩散任何距离都可能，只是以指数分布律那样小的概率发生。

现在假设(a, b)具有周期性边界条件，把这结果代入(5.2.47)，我们就得到

$$p_s(x) = p_s(a) \tag{5.2.51}$$

这是个均匀分布。

b) 奥恩斯坦-乌伦贝克过程

我们用 3.8.4 的记号，此时福克-普朗克方程是

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}(kxp) + \frac{1}{2}D\frac{\partial^2 p}{\partial x^2} \tag{5.2.52}$$

它在具有反射壁的区间(a, b)上的平稳解是

$$p_s(x) = \mathcal{N}\exp[-kx^2/D] \tag{5.2.53}$$

只要 $k>0$，它就在$(-\infty, \infty)$上可归一化。若 $k<0$，则它只在一有限区间上有意义。

现在假设

$$a=-b<0 \tag{5.2.54}$$

此时

$$\psi(x)=\exp\left[-\frac{k}{D}(x^2-a^2)\right] \tag{5.2.55}$$

若在此区间上考虑其具有周期性边界条件，只要注意到

$$\psi(a)=\psi(-a) \tag{5.2.56}$$

我们就会发现有

$$p_s(x)=p_s(a)\psi(x)/\psi(a)$$

$$=p_s(a)\exp\left[-\frac{k}{D}(x^2-a^2)\right]$$

因此对称性会产生一个与反射壁情况下相同的解。

令 $a\to\infty$ 时，我们会看到我们仍有相同的解，只要 $k>0$，则当 $a\to\infty$ 或 $b\to\infty$，结果也这样。

c) 一个化学反应模型

尽管通常化学反应最好是由生灭主方程来刻划（如在第7章中所介绍的那样），其近似处理则常常是利用福克-普朗克方程来给出。反应

$$X+A\rightleftarrows 2X \tag{5.2.57}$$

是令人感兴趣的，因为它具有一个在 $x=0$ 处的出口边界(其中 x 是X的分子数)。显然，如果X不存在，则就不可能出现X与A之间的碰撞，因而也就不会再产生出X。

在7.6.1节中导出了它的福克-普朗克方程，它是

$$\partial_t p(x,t)=-\partial_x[(ax-x^2)p(x,t)]$$

$$+\frac{1}{2}\partial_x^2[(ax+x^2)p(x,t)] \tag{5.2.58}$$

在 $x=\alpha$ 及 $x=\beta$ 我们加上反射边界。此时，平稳解就是

$$p_s(x)=e^{-2x}(a+x)^{4a-1}x^{-1} \tag{5.2.59}$$

如果 $a=0$ 则它不是可归一的。在 $x=0$ 处的极点是那里的吸收所

造成的。事实上只要把此解与(5.2.28)相比较，就会看到

$$B(0,\ t)\equiv(ax+x^2)_{x=0}=0$$
$$A(0,\ t)\equiv(ax-x^2)_{x=0}=0$$
$$\partial_x B(0,\ t)\equiv(a+2x)_{x=0}>0 \qquad (5.2.60)$$

所以我们确有一个出口边界。只有当 $a>0$ 时平稳解才是适切的，因为除此外就是不可归一的。在这里反射壁的物理意义十分简单：不论 X 的哪一个分子消失了，我们立即就补上另一个分子进去。在图 5.2 中给出了 $p_s(x)$ 的曲线图示。事实上使所有的 x 都消失所需的时间是非常长的，实际上除了靠近 $x=0$ 的情况外平稳解 (5.2.59) 是分布的很好代表。

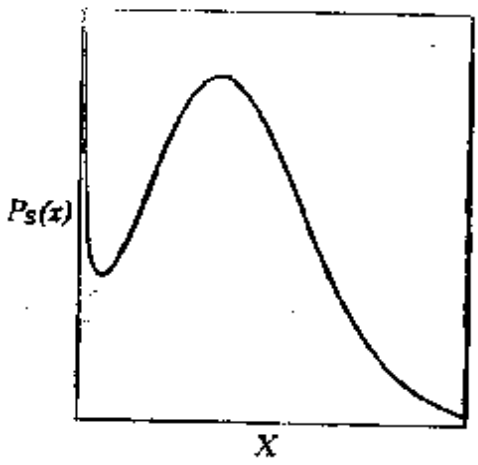


图 5.2 反应 $X+A\rightleftharpoons 2X$ 的不可归一"平稳解" $p_s(x)$

5.2.4 后向福克-普朗克方程的边界条件

我们假定对于 x, t 及 x', t' 的某个集合来说 $p(x,\ t|x',\ t')$ 服从前向福克-普朗克方程，并且过程局限于具有边界 S 的一个区域 R 中。那末只要 s 是 t 与 t' 之间的时刻，就有

$$0=\frac{\partial}{\partial s}p(x,\ t|x',\ t')$$
$$=\frac{\partial}{\partial s}\int \mathrm{d}y\,p(x,\ t|y,\ s)p(y,\ s|x',\ t') \qquad (5.2.61)$$

其中我们已经利用了查普曼-柯尔莫哥洛夫方程。我们要在积分号后取导数 $\frac{\partial}{\partial s}$，并对第二个因子用前向福克-普朗克方程而对第一个因子用后向福克-普朗克方程。为简单起见记

$$p(y,\ s)=p(y,\ s|x',\ t') \qquad (5.2.62)$$
$$\bar{p}(y,\ s)=p(x,\ t|y,\ s)$$

那末

$$0=\int_R \mathrm{d}\boldsymbol{y}\Big[-\sum_i \frac{\partial}{\partial y_i}(A_i p)+\sum_{i,j}\frac{\partial^2}{\partial y_i \partial y_j}(B_{ij}p)\Big]\bar{p}$$

$$+\int_R \mathrm{d}\boldsymbol{y}\Big[-\sum_i A_i\frac{\partial \bar{p}}{\partial y_i}-\sum_{i,j}B_{ij}\frac{\partial^2 \bar{p}}{\partial y_i \partial y_j}\Big]p \qquad (5.2.63)$$

并且作某些计算后又

$$=\int_R \mathrm{d}\boldsymbol{y}\sum_i \frac{\partial}{\partial y_i}\Big\{-A_i p\bar{p}$$

$$+\frac{1}{2}\sum_j\Big[\bar{p}\frac{\partial}{\partial y_j}(B_{ij}p)-pB_{ij}\frac{\partial \bar{p}}{\partial y_j}\Big]\Big\} \qquad (5.2.64)$$

$$=\int_S\sum_i \mathrm{d}S_i\Big\{\bar{p}\Big[-A_i p+\frac{1}{2}\sum_j\frac{\partial}{\partial y_j}(B_{ij}p)\Big]\Big\}$$

$$-\frac{1}{2}\int_S\sum_i \mathrm{d}S_i p\Big(\sum_j B_{ij}\frac{\partial \bar{p}}{\partial y_j}\Big) \qquad (5.2.65)$$

现在我们分别考察各种情况。

a) 吸收边界

这要求在边界上 $p=0$。用 $p=0$ 代入 (5.2.65)，我们就得到

$$0=\int_S \bar{p}\sum_{i,j}\mathrm{d}S_i B_{ij}\frac{\partial p}{\partial y_j} \qquad (5.2.66)$$

所以不难看出还要求在边界上 $\bar{p}(\boldsymbol{y},\ t)=0$ 是与 (5.2.65) 相一致的。不过若边界是吸收的，则显然

$$p(\boldsymbol{x},\ t|\boldsymbol{y},\ s)=0 \quad \text{对所有 } \boldsymbol{y}\in \text{边界} \qquad (5.2.67)$$

这是因为此式只不过是说对 X 而言从边界上再进入 R 的概率是零。

b) 反射边界

对前向方程加上这条件就使得 (5.2.65) 中的第一个积分为零，最后面的因子则只有当

$$\sum_{i,j}n_i B_{ij}(\boldsymbol{y})\frac{\partial}{\partial y_j}[p(\boldsymbol{x},\ t|\boldsymbol{y},\ s)]=0 \qquad (5.2.68)$$

时才会对任何 p 为零。在一维的情况下，除了 B 为零的情况外，这简化为

$$\frac{\partial}{\partial y}p(\boldsymbol{x},\ t|\boldsymbol{y},\ s)=0 \qquad (5.2.69)$$

c) 其他边界

在这一节我们不考察这些情况，其详细讨论请参看文献[5.4]。

5.2.5 本征函数方法(齐次过程)

现在我们要说明在齐次过程情况下一个解如何才能最自然地用本征函数来表示。我们考虑反射与吸收边界的情况。

a) 反射边界情况下的本征函数

对一个在区间(a, b)上进行而具有两反射边界的过程来考察福克-普朗克方程。我们假定此方程具有一个平稳解$p_s(x)$且方程形式为

$$\partial_t p(x, t) = -\partial_x[A(x)p(x, t)] + \frac{1}{2}\partial_x^2[B(x)p(x, t)] \tag{5.2.70}$$

用

$$p(x, t) = p_s(x)q(x, t) \tag{5.2.71}$$

来定义函数$q(x, t)$，通过直接代入方程的办法，我们发现$q(x, t)$满足后向福克-普朗克方程

$$\partial_t q(x, t) = A(x)\partial_x q(x, t) + \frac{1}{2}B(x)\partial_x^2 q(x, t) \tag{5.2.72}$$

现在我们希望来考虑具有形式

$$p(x, t) = P_\lambda(x)e^{-\lambda t} \tag{5.2.73}$$

$$q(x, t) = Q_\lambda(x)e^{-\lambda t} \tag{5.2.74}$$

的解，这种类型的解满足本征函数方程

$$-\partial_x[A(x)P_\lambda(x)] + \frac{1}{2}\partial_x^2[B(x)P_\lambda(x)] = -\lambda P_\lambda(x) \tag{5.2.75}$$

$$A(x)\partial_x Q_{\lambda'}(x) + \frac{1}{2}B(x)\partial_x^2 Q_\lambda(x) = -\lambda' Q_{\lambda'}(x) \tag{5.2.76}$$

这样一来，我们就可以用关于x的偏积分而一下子得到

$$(\lambda - \lambda')\int_a^b dx P_\lambda(x)Q_{\lambda'}(x)$$

$$= \left[Q_{\lambda'}(x) \left\{ -A(x)P_{\lambda}(x) + \frac{1}{2}\partial_x [B(x)P_{\lambda}(x)] \right\} \right.$$

$$\left. - \frac{1}{2} B(x)P_{\lambda}(x)\partial_x Q_{\lambda'}(x) \right]_a^b \tag{5.2.77}$$

再把反射边界条件用于 $Q_{\lambda'}(x)$的系数上，就看出此系数为零，进一步，利用借平稳解(5.2.71)表示的 $q(x, t)$的定义，我们易证

$$\frac{1}{2} B(x)\partial_x Q_{\lambda'}(x) = -A(x)P_{\lambda'}(x) + \frac{1}{2}\partial_x [B(x)P_{\lambda'}(x)] \tag{5.2.78}$$

因此这一项也是零，所以 $Q_{\lambda}(x)$与 $P_{\lambda}(x)$组成了一个正交系

$$\int_a^b \mathrm{d}x P_{\lambda}(x) Q_{\lambda'}(x) = \delta_{\lambda\lambda'} \tag{5.2.79}$$

或者存在另外两个正交系

$$\int_a^b \mathrm{d}x p_s(x) Q_{\lambda}(x) Q_{\lambda'}(x) = \delta_{\lambda\lambda'} \tag{5.2.80}$$

$$\int_a^b \mathrm{d}x [p_s(x)]^{-1} P_{\lambda}(x) P_{\lambda'}(x) = \delta_{\lambda\lambda'} \tag{5.2.81}$$

值得注意的是由于

$$P_0(x) = p_s(x) \tag{5.2.82}$$

$$Q_0(x) = 1 \tag{5.2.83}$$

当令 $\lambda = \lambda' = 0$ 时就给出了平稳解 $p_s(x)$的归一化，利用这种正交性我们就能够把任何解都用本征函数表示，因为如果

$$p(x, t) = \sum_{\lambda} A_{\lambda} P_{\lambda}(x) \mathrm{e}^{-\lambda t} \tag{5.2.84}$$

则

$$\int_a^b \mathrm{d}x Q_{\lambda}(x) p(x, 0) = A_{\lambda} \tag{5.2.85}$$

举例来说，条件概率 $p(x, t|x_0, 0)$是由初始条件

$$p(x, 0|x_0, 0) = \delta(x - x_0) \tag{5.2.86}$$

所给定的，所以

$$A_{\lambda} = \int_a^b \mathrm{d}x Q_{\lambda}(x)\delta(x - x_0) = Q_{\lambda}(x_0) \tag{5.2.87}$$

从而

$$p(x,\ t|x_0,\ 0)=\sum_{\lambda}P_\lambda(x)Q_\lambda(x_0)e^{-\lambda t} \tag{5.2.88}$$

我们还可以把自相关函数十分明确地写出来，它是

$$\langle x(t),\ x(0)\rangle=\int dx\int dx_0 x x_0 p(x,\ t|x_0,\ 0)\,p_s(x) \tag{5.2.89}$$

$$=\sum_{\lambda}\left[\int dx x P_\lambda(x)\right]^2 e^{-\lambda t} \tag{5.2.90}$$

这里我们已经利用了 $Q_\lambda(x)$ 的定义(5.2.74)。

b) 吸收边界情况下的本征函数

这是一种十分相似的情况，定义 P_λ 及 Q_λ 同上，只不过 $p_s(x)$ 不再定义为具有反射边界条件的福克-普朗克方程的平稳解，而是具有吸收边界条件的了。按此定义我们发现必然有

$$P_\lambda(a)=Q_\lambda(a)=P_\lambda(b)=Q_\lambda(b)=0 \tag{5.2.91}$$

并且有关正交性的证明仍然通得过，因而利用上述条件及本征函数方程(5.2.75, 76) 就可算出本征函数，并且所有其他结果也大体相同。不过 λ 的值域不包含 $\lambda=0$ 在内，因而当 $t\to\infty$ 时

$$p(x,\ t|x_0,\ 0)\to 0。$$

5.2.6 一些例子

a) 具有吸收边界的维纳过程

局限于区间(0, 1)上的福克-普朗克方程是

$$\partial_t p=\frac{1}{2}\partial_x^2 p \tag{5.2.92}$$

吸收边界条件要求

$$p(0,\ t)=p(1,\ t)=0 \tag{5.2.93}$$

并且合适的本征函数是 $\sin(n\pi x)$，所以可以把 p 展开为一个傅里叶正弦级数

$$p(x,\ t)=\sum_{n=1}^{\infty}b_n(t)\sin(n\pi x) \tag{5.2.94}$$

它必然满足(5.2.93)。初始条件要选择得使

$$p(x,\ 0)=\delta(x-x_0) \tag{5.2.95}$$

它的傅里叶系数是

$$b_n(0) = 2\int_0^1 dx\delta(x-x_0)\sin(n\pi x) = 2\sin(n\pi x_0) \qquad (5.2.96)$$

把傅里叶展开式(5.2.94)代入(5.2.92)，就有

$$\dot{b}_n(t) = -\lambda_n b_n(t) \qquad (5.2.97)$$

其中

$$\lambda_n = n^2\pi^2/2 \qquad (5.2.98)$$

并且解就是

$$b_n(t) = b_n(0)\exp(-\lambda_n t) \qquad (5.2.99)$$

所以我们有解（由初始条件 (5.2.95) 我们知道它就是条件概率 $p(x, t|x_0, 0)$）

$$p(x, t|x_0, 0) = 2\sum_{n=1}^{\infty} \exp(-\lambda_n t)\sin(n\pi x_0)\sin(n\pi x) \qquad (5.2.100)$$

b) 具有反射边界的维纳过程

在此边界条件简化为（在区间(0, 1)上）

$$\partial_x p(0, t) = \partial_x p(1, t) = 0 \qquad (5.2.101)$$

并且本征函数是 $\cos(n\pi s)$，所以我们可作一个傅里叶余弦展开

$$p(x, t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n(t)\cos(n\pi x) \qquad (5.2.102)$$

并具有同样的初始条件

$$p(x, 0) = \delta(x-x_0) \qquad (5.2.103)$$

所以

$$a_n(0) = 2\int_0^1 dx\cos(n\pi x)\delta(x-x_0) = 2\cos(n\pi x_0) \qquad (5.2.104)$$

用前面一样的方法，我们得到

$$a_n(t) = a_n(0)\exp(-\lambda_n t) \qquad (5.2.105)$$

其中

$$\lambda_n = n^2\pi^2/2 \qquad (5.2.106)$$

这就使得

$$p(x, t|x_0, 0) = 1 + 2\sum_{n=1}^{\infty} \cos(n\pi x_0)\cos(n\pi x)\exp(-\lambda_n t) \qquad (5.2.107)$$

若 $t\to\infty$，过程就成了平稳的，其平稳分布为

$$p_s(x)=\lim_{t\to\infty}p(x,\ t|x_0,\ 0)=1 \tag{5.2.108}$$

我们可以用

$$\langle x(t),\ x(0)\rangle_s=\int_0^1\int_0^1 dxdx_0xx_0p(x,\ t|x_0,\ 0)p_s(x) \tag{5.2.109}$$

来计算其平稳自相关函数，并且可用显式积出来，即

$$\langle x(t),\ x(0)\rangle_s=\frac{1}{4}+\frac{8}{\pi^4}\sum_{n=0}^{\infty}\exp(-\lambda_{2n+1}t)(2n+1)^{-4} \tag{5.2.110}$$

我们看到，在 $t\to\infty$ 时所有指数因子变为零并且

$$\langle x(t),\ x(0)\rangle_s\longrightarrow\frac{1}{4}=[\langle x\rangle_s]^2 \tag{5.2.111}$$

当 $t\to 0$ 时，只要考虑到(由黎曼-ζ 函数理论而得知的)恒等式

$$\sum_{n=0}^{\infty}(2n+1)^{-4}=\frac{\pi^4}{96} \tag{5.2.113}$$ (*)

就有

$$\langle x(t),\ x(0)\rangle_s\longrightarrow\frac{1}{4}+\frac{8}{\pi^4}\sum_{n=0}^{\infty}(2n+1)^{-4}=\frac{1}{3}=\langle x_s^2\rangle \tag{5.2.114}$$

c) 奥恩斯坦-乌伦贝克过程

如同在第 3.8.4 节中那样，福克-普朗克方程是

$$\partial_t p(x,\ t)=\partial_x[kxp(x,\ t)]+\frac{1}{2}D\partial_x^2p(x,\ t) \tag{5.2.115}$$

本征函数 Q_λ 所满足的本征函数方程是

$$d_x^2Q_\lambda-\frac{2kx}{D}d_xQ_\lambda+\frac{2\lambda}{D}Q_\lambda=0 \tag{5.2.116}$$

基于代换

$$y=x\sqrt{k/D}$$

它就变成了厄密多项式所满足的微分方程([5.6])

$$d_y^2Q_\lambda-2yd_yQ_\lambda+(2\lambda/k)Q_\lambda=0 \tag{5.2.117}$$

(*) 原文中没有(5.2.112)式。——译注

我们可以把它表示为

$$Q_\lambda = (2^n n!)^{-\frac{1}{2}} H_n\left(x\sqrt{\frac{k}{D}}\right) \tag{5.2.118}$$

其中

$$\lambda = nk \tag{5.2.119}$$

这些解如同(5.2.79～81)那样能够归一化。

如前所知平稳解是

$$p_s(x) = \left(\frac{k}{\pi D}\right)^{\frac{1}{2}} \exp(-kx^2/D) \tag{5.2.120}$$

而一个通解可以表示为

$$p(x, t) = \sum_n \sqrt{[k/(2^n n!\pi D)]} \cdot \exp(-kx^2/D) H_n(x\sqrt{k/D}) e^{-nkt} A_n \tag{5.2.121}$$

其中

$$A_n = \int_{-\infty}^{\infty} dx p(x, 0) H_n(x\sqrt{k/D})(2^n n!)^{-\frac{1}{2}} \tag{5.2.122}$$

这个结果也可以利用厄密多项式的母函数来求在第3.8.4节中所给出的条件概率的显式解而直接获得。我们看到，弛豫到平稳态的时间尺度是由本征值

$$\lambda_n = nk \tag{5.2.123}$$

所给定的。这里的 k 是确定性弛豫的比例常数，因而它也确定着弛豫过程中的最慢时间，我们也可以用(5.2.90)直接计算出自相关函数，利用[5.6]中的结果

$$H_1(y) = 2y \tag{5.2.124}$$

我们就会知道正交性意味着只有相应于 $n=1$ 的本征函数具有非零系数，这就要计算

$$\int x P_{\lambda 1}(x) dx = \int_{-\infty}^{\infty} \sqrt{k/(2\pi D)} \exp(-kx^2/D)(2x\sqrt{k/D})x = \sqrt{D/2k} \tag{5.2.125}$$

从而如同在节3.8.4及节4.4.4中已经得到的，

$$\langle x(t),\ x(0)\rangle_s = \frac{D}{2k}\,\mathrm{e}^{-kt} \tag{5.2.126}$$

d) 瑞利过程

我们采用在4.4.5节已经展开的振幅涨落方法来考察，这里的福克-普朗克方程是

$$\partial_t p(x,\ t) = \partial_x[(\gamma x - \mu/x)p(x,\ t)] + \mu\partial_x p(x,\ t) \tag{5.2.127}$$

其中

$$\mu = \varepsilon^2/2 \tag{5.2.128}$$

过程的值域是$(0,\ \infty)$，方程的(归一化)平稳解是

$$p_s(x) = (\gamma x/\mu)\exp(-\gamma x^2/2\mu) \tag{5.2.129}$$

$Q_\lambda(x)$的本征函数方程是

$$\mathrm{d}_x^2 Q_\lambda + \left(\frac{1}{x} - \gamma x/\mu\right)\mathrm{d}_x Q_\lambda + (\lambda/\mu)\,Q_\lambda = 0 \tag{5.2.130}$$

令

$$z = x^2\gamma/2\mu \tag{5.2.131}$$

就得到

$$z\mathrm{d}_z^2 Q_\lambda + (1-z)\mathrm{d}_z Q_\lambda + (\lambda/2\gamma)\,Q_\lambda = 0 \tag{5.2.132}$$

只要

$$\lambda = 2n\gamma \tag{5.2.133}$$

这个方程就成了拉盖尔多项式(Laguerre polynomials)所满足的微分方程([5.6])。我们可以把Q_λ表示为

$$Q_\lambda(x) = L_n(\gamma x^2/2\mu) \tag{5.2.134}$$

它是归一化了的。因而这个过程的条件概率是

$$p(x,\ t|x_0,\ 0) = \sum_{n=0}^{\infty} \frac{\gamma x}{\mu}\exp\left(-\frac{\gamma x^2}{2\mu}\right) \cdot L_n\left(\frac{\gamma x_0^2}{2\mu}\right)L_n\left(\frac{\gamma x^2}{2\mu}\right)\mathrm{e}^{-2n\gamma t} \tag{5.2.135}$$

我们可以用(5.2.90)的方法来计算其自相关函数：

$$\langle x(t),\ x(0)\rangle = \sum_{n=0}^{\infty}\left[\int_0^\infty x\mathrm{d}x\ \frac{\gamma x}{\mu}\exp\left(\frac{-\gamma x^2}{2\mu}\right)\right.$$

$$\cdot L_n\left(\frac{\gamma x^2}{2\mu}\right)\Big]^2 \exp(-2n\gamma t) \tag{5.2.136}$$

利用

$$\int_0^\infty \mathrm{d}z z^\alpha \mathrm{e}^{-z} L_n(z) = (-1)^n \Gamma(\alpha+1)\binom{\alpha}{n} \tag{5.2.137}$$

即可看出自相关系数是

$$\langle x(t), x(0)\rangle = \frac{2\mu}{\gamma}\sum_{n=0}^{\infty}\frac{\pi}{4}\binom{1}{2}^2 \exp(-2n\gamma t) \tag{5.2.138}$$

5.2.7 齐次过程的首次通过时间

一个其位置由福克-普朗克方程所刻划的粒子，能在 x 的某个区域中停留多久呢？这是人们经常感兴趣的问题。如同 3.6 节中已叙述过的那样，这个问题的解答可以利用后向福克-普朗克方程来获得。

a) 两端是吸收壁的情况

假设粒子在时刻 $t=0$ 的初始位置是 x，我们要问它在区间 (a, b) 中停留多久，其中 (a, b) 假设为包含着 x：

$$a \leqslant x \leqslant b \tag{5.2.139}$$

我们在 a 及 b 两点设定两个吸收壁，粒子一旦到达 a 或 b 就会从系统中消失。因此只要粒子仍然在区间 (a, b) 中，它就不曾离开过这个区间。

在这些条件下粒子于时刻 t 仍然在 (a, b) 中的概率是

$$\int_a^b \mathrm{d}x' p(x', t|x, 0) \equiv G(x, t) \tag{5.2.140}$$

假设粒子离开 (a, b) 的时刻是 T，则 (5.2.140) 可以改写成

$$\mathrm{Prob}(T \geqslant t) = \int_a^b \mathrm{d}x' p(x', t|x, 0) \tag{5.2.141}$$

它意味着 $G(x, t)$ 与 $\mathrm{Prob}(T \geqslant t)$ 是相同的。由于系统是齐次的，我们可以写成

$$p(x', t|x, 0) = p(x', 0|x, -t) \tag{5.2.142}$$

而且后向福克-普朗克方程可以写为

$$\partial_t p(x', t|x, 0) = A(x)\partial_x p(x', t|x, 0) + \frac{1}{2} B(x)\partial_x^2 p(x', t|x, 0) \tag{5.2.143}$$

于是 $G(x, t)$ 服从方程

$$\partial_t G(x, t) = A(x)\partial_x G(x, t) + \frac{1}{2} B(x)\partial_x^2 G(x, t) \tag{5.2.144}$$

边界条件显然是

$$p(x', 0|x, 0) = \delta(x - x')$$

因而

$$G(x, 0) = 1 \quad a \leqslant x \leqslant b$$
$$= 0 \quad \text{其他} \tag{5.2.145}$$

并且若 $x = a$ 或 b，则粒子立刻被吸收，所以

$$\text{Prob}(T \geqslant t) = 0$$

当 $x = a$ 或 b，也就是说

$$G(a, t) = G(b, t) = 0 \tag{5.2.146}$$

由于 $G(x, t)$ 是 $T \geqslant t$ 的概率，T 的任何函数的均值就是

$$\langle f(T) \rangle = -\int_0^\infty f(t) \mathrm{d}G(x, t) \tag{5.2.147}$$

这样一来，平均首次通过时间

$$T(x) = \langle T \rangle \tag{5.2.148}$$

就经分部积分后由

$$T(x) = -\int_0^\infty t\partial_t G(x, t)\mathrm{d}t \tag{5.2.149}$$

$$= \int_0^\infty G(x, t)\mathrm{d}t \tag{5.2.150}$$

所给定。

类似地，若定义

$$T_n(x) = \langle T^n \rangle \tag{5.2.151}$$

则有

$$T_n(x)=\int_0^\infty t^{n-1}G(x,\ t)\mathrm{d}t \tag{5.2.152}$$

我们可以利用(5.2.150)并在$(0,\ \infty)$上积分(5.2.144)来导出$T(x)$所满足的一个简单常微分方程。注意到

$$\int_0^\infty \partial_t G(x,\ t)\mathrm{d}t=G(x,\ \infty)-G(x,\ 0)=-1 \tag{5.2.153}$$

就可得到

$$A(x)\partial_x T(x)+\frac{1}{2}B(x)\partial_x^2 T(x)=-1 \tag{5.2.154}$$

其边界条件为

$$T(a)=T(b)=0 \tag{5.2.155}$$

类似地我们有

$$-nT_{n-1}(x)=A(x)\partial_x T_n(x)+\frac{1}{2}B(x)\partial_x^2 T_n(x) \tag{5.2.156}$$

这意味着首次通过时间的所有各阶矩都可以用多次积分来得到。

方程的解。方程(5.2.154)可以用积分直接求解，作了某些处理后此解可以用

$$\boxed{\psi(x)=\exp\left\{\int_a^x \mathrm{d}x'\left[2A(x')/B(x')\right]\right\}} \tag{5.2.157}$$

来表示出来。我们得到

$$T(x)=\frac{\left[\begin{array}{l}2\left[\left(\int_a^x\frac{\mathrm{d}y}{\psi(y)}\right)\int_x^b\frac{\mathrm{d}y'}{\psi(y')}\int_a^{y'}\frac{\mathrm{d}z\psi(z)}{B(z)}\right.\\ \left.-\left(\int_x^b\frac{\mathrm{d}y}{\psi(y)}\right)\int_a^x\frac{\mathrm{d}y'}{\psi(y')}\int_a^{y'}\frac{\mathrm{d}z\psi(z)}{B(z)}\right]\end{array}\right]}{\int_a^b\frac{\mathrm{d}y}{\psi(y)}} \tag{5.2.158}$$

b) 一端是吸收壁的情况

我们仍考虑在区间(a,b)中的运动但假定在a处是反射壁，此时边界条件变成了

$$\partial_x G(a,\ t)=0 \tag{5.2.159a}$$

$$G(b,\ t)=0 \tag{5.2.159b}$$

这是从5.2.4节中导出的后向福克-普朗克方程的条件得到的，解出具有相应的边界条件的(5.2.154)，就可得到

$$\boxed{T(x)=2\int_x^b \frac{\mathrm{d}y}{\psi(y)}\int_a^y \frac{\psi(z)}{B(z)}\,\mathrm{d}z} \tag{5.2.160}$$

式中 a 为反射壁、b 为吸收壁，$a<b$。

类似地我们还会发现有

$$\boxed{T(x)=2\int_a^x \frac{\mathrm{d}y}{\psi(y)}\int_y^b \frac{\psi(z)}{B(z)}\,\mathrm{d}z} \tag{5.2.161}$$

式中 a 为吸收壁，b 为反射壁，$a<b$。

c) 一个应用——从势垒上逸出

我们假定一个点按照福克-普朗克方程

$$\partial_t p(x,\ t)=\partial_x[U'(x)p(x,\ t)]+D\partial_x^2 p(x,\ t) \tag{5.2.162}$$

的规律运动着。如图5.3中所示，其位势有极大值与极小值。我们假设运动是在一个无穷区域上进行的，这意味着其平稳解是

$$p_s(x)=\mathcal{N}\exp[-U(x)/D] \tag{5.2.163}$$

如图5.3中所示，它是双峰的，所以它具有相对是较大的概率处于 b 的左边或 b 的右边，而不是靠近 b。那应从左边势井中逸出的平均逃逸时间是多少呢？这就是说若 x 是在 b 邻近的一点，则从 a 到达 x 的平均首次通过时间是多少？在(5.2.160)中利用代换

$$\begin{aligned} &b\to x_0\\ &a\to-\infty\\ &x\to a \end{aligned} \tag{5.2.164}$$

就可得到

$$T(a\to x_0)=\frac{1}{D}\int_a^{x_0}\mathrm{d}y\exp[U(y)/D]\int_{-\infty}^{y}\exp[-U(z)/D]\mathrm{d}z \tag{5.2.165}$$

假如 $U(x)$ 的中央极大值大而 D 值小，则 $\exp[U(x)/D]$ 在 $x=b$ 处就陡峭地耸起，而同时在 $z=b$ 的邻近处 $\exp[-U(z)/D]$ 就很小。因此

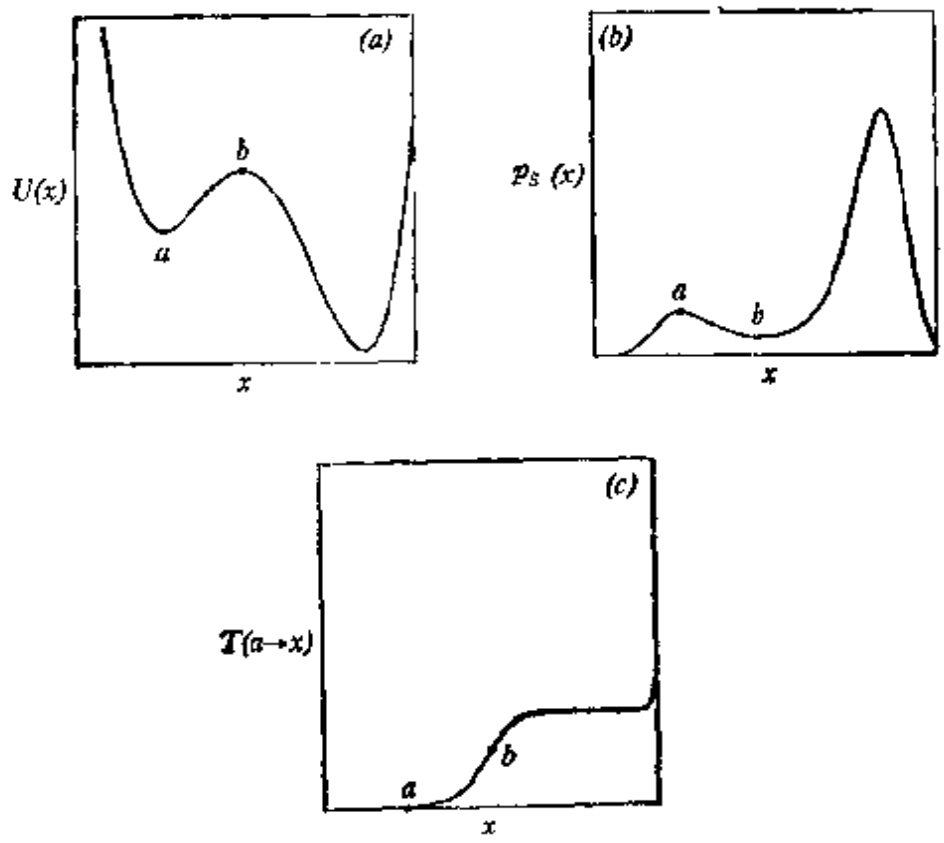


图 5.3

(a) 双井势 $U(x)$; (b) 平稳分布 $P_s(x)$;
(c) 从 a 到 x 的平均首次通过时间 $T(a-x_0)$

$$\int_{-\infty}^{y} \exp[-U(z)/D]\mathrm{d}z$$

在 $y=b$ 的附近就是 y 的一个慢变函数。这意味着积分

$$\int_{-\infty}^{y} \exp[-U(z)/D]\mathrm{d}z$$

对于那些使得 $\exp[U(y)/D]$ 之值明显地不同于零的 y 值而言近似为常值。因而我们可在被积函数中的积分里令 $y=b$ 并把这样所得到的常数因子移到关于 y 的积分号外。所以我们可以用

$$T(a\to x_0)\simeq\left\{\frac{1}{D}\int_{-\infty}^{b}\mathrm{d}z\exp[-U(z)/D]\right\}\cdot\int_{a}^{x_0}\mathrm{d}y\exp[U(y)/D]\ (*) \qquad (5.2.166)$$

(*) 原文将 $\mathrm{d}z$ 误写为 $\mathrm{d}y$. ——译注

来作为(5.2.165)的近似。

请注意，由于 $p_s(x)$ 的定义(5.2.163)，我们可以得知

$$\int_{-\infty}^{b} dy \exp[-U(z)/D] = n_a/N \tag{5.2.167}$$

这意味着 n_a 是当系统为平稳时粒子在 b 左侧的概率。

在图 5.3 中指出了相应于不同的 x_0 值的 $T(a\to x_0)$ 的图形，它说明对于在左势井中的 x_0 来说到达 x_0 的平均首次通过时间是十分小的，而对于右势井中的 x_0 则十分大。这意味着进行超越势垒而到达右势井过程的粒子实际上要把绝大部分时间花在战胜势垒的阻碍上。对于一个初始位置在 a 点而要到达 c 点附近的粒子来说，它所需的时间关于初始和最后的精确位置是十分不敏感的，所以把这个时间称为逃逸时间而加以讨论是十分有意义的。为计算其大小我们可以进一步假设在 b 附近有

$$U(x)\simeq U(b)-\frac{1}{2}\left(\frac{x-b}{\delta}\right)^2 \tag{5.2.168}$$

并且在 a 附近有

$$U(x)\simeq U(a)+\frac{1}{2}\left(\frac{x-a}{\alpha}\right)^2 \tag{5.2.169}$$

这样一来，(5.2.166)中的常数因子就可以赋值为

$$\int_{-\infty}^{b} dz \exp[-U(z)/D]\sim\int_{-\infty}^{\infty} dz \exp\left[-\frac{U(a)}{D}-\frac{(z-a)^2}{2D\alpha^2}\right] \tag{5.2.170}$$

$$\sim\alpha\sqrt{2\pi D}\exp[-U(\alpha)/D] \tag{5.2.171}$$

并且在假设了 x_0 是在中央点 b 的充分右方之后其另一个因子就可以赋值为

$$\int_{a}^{x_0} dy \exp[U(y)/D]\sim\int_{-\infty}^{\infty} dy\left[\frac{U(b)}{D}-\frac{(y-b)^2}{2D\delta^2}\right] \tag{5.2.172}$$

$$=\delta\sqrt{2\pi D}\exp[U(b)/D] \tag{5.2.173}$$

把这两个结果代入(5.2.166)我们就得到

$$T(a\to x_0)\simeq 2\alpha\delta\pi\exp\{[U(b)-U(a)]/D\} \tag{5.2.174}$$

这就是化学反应理论中的阿雷尼厄斯公式 (Arrhenius formula) 我们可以用引进一个坐标使得 $x=a$ 是物质 A 而 $x=c$ 是物质 c 的办法来建立反应过程的模型。反应就是由上述扩散过程为模型的，并且这两种不同的化学物质由在 b 处的势垒分离开。在化学反应中统计力学已给定了 D 之值为

$$D=kT \tag{5.2.175}$$

其中 k 是玻耳兹曼常数而 T 是绝对温度。我们发现反应对温度的最重要的依赖来自于通常表示为

$$\exp(\Delta E/kT) \tag{5.2.176}$$

的指数因子，它预示着对温度的十分独特的依赖关系。直观上这种结果是显然的，因为这个指数因子所表示的是系统处于热平衡时分子的能量将超过势垒能量的概率。到达了这种能量的分子将以某个有限概率进行反应。

在第 9 章我们将回到类似的问题上来作更详细的讨论。

5.2.8 经由区间的一特定端点出口的概率

一个初始位置在 (a, b) 中的 x 处的粒子经由 a 出口的概率是多大?其平均出口时间又是多少?

一个粒子在经过时间 t 后经由 a 出口的全概率是由在 a 点的概率流关于时间的积分来给定的，所以我们是由

$$g_a(x, t)=-\int_t^\infty \mathrm{d}t' J(a, t'|x, 0) \tag{5.2.177}$$

$$=\int_t^\infty \mathrm{d}t' \left\{-A(a)p(a, t'|x, 0) + \frac{1}{2}\partial_a[B(a)p(a, t'|x, 0)]\right\} \tag{5.2.178}$$

(这里之所以要取负号是因为我们要求流的指向朝左)及

$$g_b(x, t)=\int_t^\infty \mathrm{d}t'\left\{A(b)p(b, t'|x, 0) - \frac{1}{2}\partial_b[B(b)p(b, t'|x, 0)]\right\} \tag{5.2.179}$$

来定义这个概率。这些数量分别给出了粒子经由 a 或 b 出口的概率。(在已知粒子是经由 a 出口时)粒子在时刻 t 之后出口的概率则是

$$\mathrm{Prob}(T_a>t)=g_a(x,\ t)/g_a(x,\ \infty) \tag{5.2.180}$$

现在来求 $g_a(x,\ t)$ 满足的方程。利用 $p(a,\ t|x,\ 0)$ 满足后向福克-普朗克方程的事实，我们就有

$$\begin{aligned} A(x)\partial_x g_a(x,\ t)+\frac{1}{2}B(x)\partial_x g_a(x,\ t) &= -\int_t^\infty \mathrm{d}t'\partial_{t'}J(a,\ t'|x,\ 0) \\ &= J(a,\ t|x,\ 0) \\ &= \partial_t g_a(x,\ t) \end{aligned} \tag{5.2.181}$$

若已知出口是经由 a 的，则平均出口时间是

$$\begin{aligned} T(a,\ x) &= -\int_0^\infty t\partial_t \mathrm{Prob}(T_a>t)\mathrm{d}t \\ &= \int_0^\infty g_a(x,\ t)\mathrm{d}t/g_a(x,\ \infty) \end{aligned} \tag{5.2.182}$$

对(5.2.181)关于 t 作积分，就可得到

$$\boxed{\begin{aligned} A(x)\partial_x[\pi_a(x)T(a,\ x)]+\frac{1}{2}B(x)\partial_x^2[\pi_a(x)T(a,\ x)] \\ =-\pi_a(x) \end{aligned}} \tag{5.2.183}$$

其中我们定义了

$$\pi_a(x)=(\text{经由 } a \text{ 出口的概率})=g_a(x,\ \infty) \tag{5.2.184}$$

(5.2.183)的边界条件是十分明显的，因为它们可以由后向福克-普朗克方程的边界条件来得到，即

$$\pi_a(a)T(a,\ a)=\pi_a(b)T(a,\ b)=0 \tag{5.2.185}$$

第一个等式中 $T(a,\ a)$ 显然是零（因为从 a 到达 a 的所需时间是0），而在第二个等式中 $\pi_a(b)$ 是零(因为从 b 出发而经由 a 出口的概率为零)。

只要令(5.2.181)中的 $t\to\infty$，我们就看到粒子终究不会留在

$a, b)$中，因此该式的右边趋于零，从而得知

$$\boxed{A(x)\partial_x\pi_a(x)+\frac{1}{2}B(x)\partial_x\pi_a(x)=0} \tag{5.2.186}$$

此时的边界条件则成了

$$\pi_a(a)=1$$
$$\pi_a(b)=0 \tag{5.2.187}$$

(5.2.186)在赋予上述边界条件及条件

$$\pi_a(x)+\pi_b(x)=1 \tag{5.2.188}$$

时的解是

$$\pi_a(x)=\left[\int_x^b \mathrm{d}y\psi(y)\right]\Big/\int_a^b \mathrm{d}y\psi(y) \tag{5.2.189}$$

$$\pi_b(x)=\left[\int_a^x \mathrm{d}y\psi(y)\right]\Big/\int_a^b \mathrm{d}y\psi(y) \tag{5.2.190}$$

其中$\psi(x)$的定义如(5.2.157)。

这些公式可以在(第9.1.4节中的)初始时集中在一个不稳定平稳点上的分布弛豫问题中找到应用。

5.3 多维福克-普朗克方程

在多变量情况下，福克-普朗克方程所表现的行为范围比单变量的情况本质上更为复杂。其边界也不只是线段的两端点而是曲线或曲面。而且边界上的性质也会因地而异。平稳解即便是具有反射边界的，也会对应于非零的概率流，而且本征函数不再是那样简单了。

尽管这样，一维与多维情况之间的类似之处仍然是有用的。本节正是要沿着与一维情况相同的线索来加以考察。

5.3.1 变量的代换

设有以x_i为变元的福克-普朗克方程

$$\partial_t p(x,\ t)=-\sum_i \partial_i[A_i(x)p(x,\ t)]$$

$$+\frac{1}{2}\sum_{i,j}\partial_i\partial_j[B_{ij}(\boldsymbol{x})p(\boldsymbol{x},\ t)] \tag{5.3.1}$$

我们要想知道新的变元

$$y_i=f_i(\boldsymbol{x}) \tag{5.3.2}$$

所满足的方程，其中 f_i 是某些独立的可微函数，我们用 $\tilde{p}(\boldsymbol{y},\ t)$ 来表示新变量的概率密度，它由

$$\tilde{p}(\boldsymbol{y},\ t)=p(\boldsymbol{x},\ t)\left|\frac{\partial(x_1,\ x_2,\ \cdots)}{\partial(y_1,\ y_2,\ \cdots)}\right| \tag{5.3.3}$$

所给定，实现这变元变换的最简单方法是对相应的随机微分方程

$$\mathrm{d}\boldsymbol{x}(t)=\boldsymbol{A}(\boldsymbol{x})\mathrm{d}t+\sqrt{\widetilde{\boldsymbol{B}(\boldsymbol{x})}}\mathrm{d}\boldsymbol{W}(t) \tag{5.3.4}$$

使用伊藤公式，然后再如第 4.3.4 中所推导的那样从所得到的新随机微分方程计算 $\tilde{p}(\boldsymbol{y},\ t)$所满足的福克-普朗克方程。

其结果比较复杂。在某些特殊情况下直接按 (5.3.3) 做可能是可取的。除非充分地利用了对称性和作了简化，否则较为繁杂的计算是无法避免的。

例子：从直角坐标系到极坐标系的变换。作为一个例子，我们考察先前在4.4.5 节中已用随机微分方程的方法作了研究的瑞利过程是如何变换到极坐标表示的。此时福克-普朗克方程是

$$\partial_t p(\boldsymbol{E}_1,\ \boldsymbol{E}_2,\ t)=\gamma\frac{\partial}{\partial\boldsymbol{E}_1}\boldsymbol{E}_1 p+\gamma\frac{\partial}{\partial\boldsymbol{E}_2}\boldsymbol{E}_2 p$$
$$+\frac{1}{2}\varepsilon^2\left(\frac{\partial^2 p}{\partial\boldsymbol{E}_1}+\frac{\partial^2 p}{\partial\boldsymbol{E}_2^2}\right) \tag{5.3.5}$$

在定义

$$\boldsymbol{E}_1=a\cos\phi$$
$$\boldsymbol{E}_2=a\sin\phi \tag{5.3.6}$$

后我们要求出 a 及 ϕ 所满足的福克-普朗克方程。

雅可比行列式是

$$|J|=\frac{\partial(\boldsymbol{E}_1,\ \boldsymbol{E}_2)}{\partial(a,\ \phi)}=\begin{vmatrix}\cos\phi & -a\sin\phi\\ \sin\phi & a\cos\phi\end{vmatrix}=a \tag{5.3.7}$$

利用拉普拉斯算子的极坐标形式

$$\frac{\partial^2}{\partial E_1^2}+\frac{\partial^2}{\partial E_2^2}=\frac{1}{a^2}\frac{\partial^2}{\partial\phi^2}+\frac{1}{a}\frac{\partial}{\partial a}\left(a\frac{\partial}{\partial a}\right) \tag{5.3.8}$$

以及(5.3.6)的逆表示

$$\left.\begin{aligned} a&=\sqrt{E_1^2+E_2^2}\\ \phi&=\arctan(E_2/E_1)\end{aligned}\right\} \tag{5.3.9}$$

即可看到

$$\left.\begin{aligned} &\frac{\partial a}{\partial E_1}=\frac{E_1}{\sqrt{E_1^2+E_2^2}}=\cos\phi\\ &\text{类似地}\\ &\frac{\partial a}{\partial E_2}=\sin\phi\end{aligned}\right\} \tag{5.3.10}$$

以及

$$\left.\begin{aligned} &\frac{\partial\phi}{\partial E_2}=\frac{E_1}{E_1^2+E_2^2}=\cos\phi/a\\ &\text{类似地}\\ &\frac{\partial\phi}{\partial E_1}=-\sin\phi/a\end{aligned}\right\} \tag{5.3.11}$$

因而

$$\begin{aligned}&\frac{\partial}{\partial E_1}E_1p+\frac{\partial}{\partial E_2}E_2p\\ &\quad=2p+E_1\left(\frac{\partial p}{\partial a}\frac{\partial a}{\partial E_1}+\frac{\partial p}{\partial\phi}\frac{\partial\phi}{\partial E_1}\right)\\ &\qquad+E_2\left(\frac{\partial p}{\partial a}\frac{\partial a}{\partial E_2}+\frac{\partial p}{\partial\phi}\frac{\partial\phi}{\partial E_2}\right)\\ &\quad=2p+a\frac{\partial p}{\partial a}=\frac{1}{a}\frac{\partial}{\partial a}(a^2p)\end{aligned} \tag{5.3.12}$$

设用 $\tilde{p}(a,\phi)$表示以 α、ϕ 表示的概率密度函数，则雅可比行列式的公式(5.3.3)告诉了我们必有

$$\tilde{p}(a,\phi)=\left|\frac{\partial(E_1,E_2)}{\partial(a,\phi)}\right|p(E_1,E_2)=ap(E_1,E_2) \tag{5.3.13}$$

组合(5.3.5, 8, 12, 13)，就得到了

$$\frac{\partial\tilde{p}}{\partial t}=-\frac{\partial}{\partial a}\left[\left(-\gamma a+\frac{\sigma^2}{2a}\right)\tilde{p}\right]$$

$$+\frac{\varepsilon^2}{2}\left(\frac{1}{a^2}\frac{\partial^2\bar{p}}{\partial\phi^2}+\frac{\partial^2\bar{p}}{\partial a^2}\right) \tag{5.3.14}$$

这(当然)就是4.4.5节中按伊藤公式作变元变换所得到的两个随机微分方程所对应的福克-普朗克方程。

5.3.2 边界条件

在5.2.1节中我们已经一般地接触了一下边界条件问题，在那里是按照概率流的情况来考察的。至今我们对于一个任意的多维福克-普朗克方程的各种边界条件的完整情况似乎不能作出详细的说明。因此在本书中我们将主要考虑在一个曲面 S 上具有反射壁的边界条件，即

在 S 上，$n\cdot J=0$ (5.3.15)

其中 n 是曲面的法线

$$J_i(x,\ t)=A_i(x,\ t)p(x,\ t)-\frac{1}{2}\sum_j\frac{\partial}{\partial x_j}B_{ij}(x,\ t)p(x,\ t) \tag{5.3.16}$$

的情况，以及吸收壁边界条件

对 S 上的 x，$p(x,\ t)=0$ (5.3.17)

的情况。实际上有可能曲面的某一部分是反射边界而另一部分是吸收边界。在一曲面 S 上若 A_i 或 B_{ij} 是不连续的，我们要加上

在 S 上 $\quad n\cdot J_1=n\cdot J_2$ (5.3.18)

$p_1(x)=p_2(x)\quad x$ 在 S 上

的要求，切向流分量则允许是不连续的。

关于后向方程的边界条件在第5.2.4节中已导出，为完整起见这里也写出来，它们是

对吸收边界

$$p(x,\ t|y,\ t')=0\quad y\in S \tag{5.3.19}$$

对反射边界

$$\sum_{i,j}n_iB_{ij}(y)\frac{\partial}{\partial y_j}p(x,\ t|y,\ t')=0\quad y\in S \tag{5.3.20}$$

5.3.3 平稳解：位势条件

一大类由福克-普朗克方程描述的令人感兴趣的系统允许这样一种平稳解存在，即对于 R 中的所有 x 来说其概率流都是零。如果假定是这种情况的话，只要把(5.3.16)中的 J 之定义改个形式，就可以获得一个完全等价的方程

$$\frac{1}{2}\sum_j B_{ij}(x)\frac{\partial p_s(x)}{\partial x_j}=p_s(x)\Big[A_i(x)-\frac{1}{2}\sum_j\frac{\partial}{\partial x_j}B_{ij}(x)\Big] \tag{5.3.21}$$

若矩阵 $B_{ij}(x)$ 对所有 x 都存在逆阵，(5.3.21)就可以改写成

$$\frac{\partial}{\partial x_i}\log[p_s(x)]=\sum_k B_{ik}^{-1}(x)\Big[2A_k(x)-\sum_j\frac{\partial}{\partial x_j}B_{kj}(x)\Big] \tag{5.3.22}$$

$$\equiv Z_i[A,\ B,\ x] \tag{5.3.23}$$

不是对任何 $B_{ij}(x)$ 及 $A_i(x)$ 这个方程都可满足，因为其左边显然是一梯度，因而 Z_i 也必是一梯度，而且充要条件是旋度为0，也就是说

$$\frac{\partial Z_i}{\partial x_j}=\frac{\partial Z_j}{\partial x_i} \tag{5.3.24}$$

若这个条件得以满足，则平稳解可以由(5.3.22)的简单积分而得到：

$$p_s(x)=\exp\Big\{\int^x \mathrm{d}x'\cdot Z[A,B,x']\Big\} \tag{5.3.25}$$

由于这些 Z_i 是由 $\log[p_s(x)]$ 的导数来推出的，所以后者常被看成是一个位置 $\phi(x)$，这样一来我们就更确切地可以写为

$$p_s(x)=\exp[-\phi(x)] \tag{5.3.26}$$

及

$$\phi(x)=-\int^x \mathrm{d}x'\cdot Z[A,\ B,\ x'] \tag{5.3.27}$$

由此条件(5.3.24)被人们称为位势条件。

例子：极坐标下的瑞利过程。由(5.3.14)我们有

$$\boldsymbol{A}=\begin{bmatrix} -\gamma\alpha+\varepsilon^2/2\alpha \\ 0 \end{bmatrix} \tag{5.3.28}$$

$$\underset{\sim}{B}=\begin{bmatrix} \varepsilon^2 & 0 \\ 0 & \varepsilon^2/\alpha^2 \end{bmatrix} \tag{5.3.29}$$

由它们

$$\sum_j \frac{\partial}{\partial x_j} B_{\alpha,j}=\sum_j \frac{\partial}{\partial x_j} B_{\phi,j}=0 \tag{5.3.30}$$

所以

$$\boldsymbol{Z}=2\underset{\sim}{B}^{-1}\boldsymbol{A}=\begin{bmatrix} -2\gamma\alpha/\varepsilon^2+1/\alpha \\ 0 \end{bmatrix} \tag{5.3.31}$$

并且显然地

$$\frac{\partial Z_\alpha}{\partial\phi}=\frac{\partial Z_\phi}{\partial\alpha}=0 \tag{5.3.32}$$

平稳解则是

$$p_s(\alpha,\ \phi)=\exp\left[\int^{(\alpha,\phi)}(\mathrm{d}\phi Z_\phi+\mathrm{d}\alpha Z_\alpha)\right] \tag{5.3.33}$$

$$=\mathcal{N}\exp\left(-\frac{\gamma\alpha^2}{\varepsilon^2}+\log\alpha\right) \tag{5.3.34}$$

$$=\mathcal{N}\alpha\exp\left(-\frac{\gamma\alpha^2}{\varepsilon^2}\right) \tag{5.3.35}$$

5.3.4 细致平衡

a) 细致平衡的定义

某种福克-普朗克方程的平稳解是对应着一个零概率流这一事实是细致平衡的这一物理现象的一种特殊说法。粗略地说，如果一个马尔可夫过程在平稳情况下其每种可能的转移都跟与之反方向的转移相平衡，则它就是满足细致平衡要求。由于这个概念来自物理学，让我们用一个物理学上的例子更确切地来阐明它。考察具有位置 $\boldsymbol{r}$ 及速度 $\boldsymbol{v}$ 的粒子组成的气体。一个转移对应着在某个时刻 t 具有位置与速度 $(\boldsymbol{r},\ \boldsymbol{v})$ 的一个粒子在其后的某个时刻 $t+\tau$ 具有位置与速度 $(\boldsymbol{r}',\ \boldsymbol{v}')$。这种概率密度就是联合概率密度

$$p(\boldsymbol{r}',\ \boldsymbol{v}',\ t+\tau;\boldsymbol{r},\ \boldsymbol{v},\ t)$$

我们可以把这种转移符号化地表示为

$$(\boldsymbol{r}, \boldsymbol{v}, t) \longrightarrow (\boldsymbol{r}', \boldsymbol{v}', t+\tau) \tag{5.3.36}$$

反方向的转移并不只是简单地把加了撇与不加撇的量对换就会给出来的，它是变换

$$(\boldsymbol{r}', -\boldsymbol{v}', t) \longrightarrow (\boldsymbol{r}, -\boldsymbol{v}, t+\tau) \tag{5.3.37}$$

它对应着时间反演变换并且要求速度分量随之反向，只有这样从 $\boldsymbol{r}'$ 到 $\boldsymbol{r}$ 的运动才跟从 $\boldsymbol{r}$ 到 $\boldsymbol{r}'$ 的运动方向相反。

这样一来，反演变换的概率密度就是联合概率密度

$$p(\boldsymbol{r}, -\boldsymbol{v}, t+\tau; \boldsymbol{r}', -\boldsymbol{v}', t) \tag{5.3.38}$$

细致平衡原理要求系统处于稳态时这两个联合概率是相等的。所以我们可以写出

$$p(\boldsymbol{r}', \boldsymbol{v}', \tau; \boldsymbol{r}, \boldsymbol{v}, 0) = p(\boldsymbol{r}, -\boldsymbol{v}, \tau; \boldsymbol{r}', -\boldsymbol{v}', 0) \tag{5.3.39}$$

(这个原理可以从物理学上的一些定律来推导出来，请参看文献[5.7]及5.3.66节)。

对于一个马尔可夫过程来说，我们可以把(5.3.39)更为明确地写成

$$p(\boldsymbol{r}', \boldsymbol{v}', \tau | \boldsymbol{r}, \boldsymbol{v}, 0) p_s(\boldsymbol{r}, \boldsymbol{v}) = p(\boldsymbol{r}, -\boldsymbol{v}, \tau | \boldsymbol{r}', -\boldsymbol{v}', 0) p_s(\boldsymbol{r}', -\boldsymbol{v}') \tag{5.3.40}$$

这里我们已经对相应的齐次马尔可夫过程应用了条件概率的特点(如果过程不是马尔可夫型的，则这些条件概率就只能对平稳系统适用了)。

细致平衡的一般形式是借助于一些任意变元 α_i 来表达的，这些变元在时间反演变换之下按照规则

$$\alpha_i \longrightarrow \varepsilon_i \alpha_i \tag{5.3.41}$$

$$\varepsilon_i = \pm 1 \tag{5.3.42}$$

变换到反向变元，其中 ε_i 取 -1 还是 $+1$ 则依赖于各个变元在时间反演之下是奇的还是偶的。在上述例子中 $\boldsymbol{r}$ 是偶的而 $\boldsymbol{v}$ 则是奇的。

这样，我们由细致平衡就要求有

$$p_s(x, t+\tau; x', t) = p_s(\varepsilon x', t+\tau; \varepsilon x, t) \tag{5.3.43}$$

其中我们用了 εx 来表示 $(\varepsilon_1 x_1, \varepsilon_2 x_2, \cdots)$。我们看到在 (5.3.43) 中令 $\tau=0$ 就得到

$$\delta(x-x')p_s(x')=\delta(\varepsilon x-\varepsilon x')p_s(\varepsilon x) \tag{5.3.44}$$

因为在这两个 δ 函数中仅包含着符号的变化，故它们是相等的，因此

$$p_s(x)=p_s(\varepsilon x) \tag{5.3.45}$$

是细致平衡的表示式(5.3.43)的推论。若再来用条件概率改写一下它，就有

$$\boxed{p(x,\ \tau|x',\ 0)p_s(x')=p(\varepsilon x',\ \tau|\varepsilon x,\ 0)p_s(x)} \tag{5.3.46}$$

b) 细致平衡的一般推论

(5.3.45)的一个重要推论是

$$\langle x\rangle_s=\varepsilon\langle x\rangle_s \tag{5.3.47}$$

(因而所有奇变元有着零值平稳均值)，并且对于自相关函数

$$G(\tau)\equiv\langle x(\tau)x^T(0)\rangle_s$$

来说，

$$G(\tau)=\varepsilon\langle x(0)x^T(\tau)\rangle_s\varepsilon$$

因此我们有

$$\boxed{G(\tau)=\varepsilon G^T(\tau)\varepsilon} \tag{5.3.48}$$

并且只要令 $\tau=0$ 并考虑到协方差矩阵 σ 满足 $\sigma=\sigma^T$，还有

$$\sigma\varepsilon=\varepsilon\sigma \tag{5.3.49}$$

对于谱矩阵

$$S(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-i\omega\tau}G(\tau)d\tau$$

则从(5.3.48)得知

$$S(\omega)=\varepsilon S^T(\omega)\varepsilon \tag{5.3.50}$$

c) 细致平衡含义要扩充的情况

一个马尔可夫过程可能会有几个平稳解。在这种情况下可能存在一种弱形式的细致平衡，即不是(5.3.43)而是

$$p_s^1(\boldsymbol{x},\ t+\tau;\ \boldsymbol{x}',\ t) = p_s^2(\varepsilon\boldsymbol{x}',\ t+\tau;\ \varepsilon\boldsymbol{x},\ t) \qquad (5\ 3.51)$$

成立，其中上标1及2表示两个不同的平稳解。如果有个参变量在时间反演之下是奇的但它不随时间变而变，这种情况就可能存在。例如在一个离心机中总角动量就有这个性质。恒定磁场也是以同样方式发生作用的。

在这种情况下，我们经常把细致平衡条件写成

$$p_s^\lambda(\boldsymbol{x},\ t+\tau;\ \boldsymbol{x}',\ t) = p_s^{\varepsilon\lambda}(\varepsilon\boldsymbol{x}',\ t+\tau;\ \varepsilon\boldsymbol{x},\ t) \qquad (5.3.52)$$

其中 $\boldsymbol{\lambda}$ 是一个常值向量，它在时间反演之下改变为 $\varepsilon\boldsymbol{\lambda}$。按照某种观点来看，这种情况并不表示着细致平衡，因为在一已给定平稳情况下其各种转移并不是细致地平衡着的，或许称(5.3.52)为时间反演不变性更为好些。

在我们考察的其余部分中，所谓细致平衡仍是指(5.3.45)成立的情况，因为从(5.3.52)并不能产生较强的推论。

5.3.5 细致平衡的推论

福克-普朗克方程的细致平衡表达方式已独立地分别由范坎彭[5.7]、乌尔霍恩[5.8]及格雷厄姆-哈肯[5.9]所提出。我们将采用一种稍许更直接和更一般的方式来陈述这些条件，我们想获得关于一个齐次马尔可夫过程的漂移系数、扩散系数及跳概率的充要条件，以便使它具有满足细致平衡的平稳解，我们将证明这些充要条件是由

$$\begin{aligned}
&\text{(i)}\quad W(\boldsymbol{x}|\boldsymbol{x}')p_s(\boldsymbol{x}') = W(\varepsilon\boldsymbol{x}'|\varepsilon\boldsymbol{x})p_s(\boldsymbol{x})\\
&\text{(ii)}\quad \varepsilon_i A_i(\varepsilon\boldsymbol{x})p_s(\boldsymbol{x}) = -A_i(\boldsymbol{x})p_s(\boldsymbol{x})\\
&\qquad\qquad + \sum_j \frac{\partial}{\partial x_j}\,[B_{ij}(\boldsymbol{x})p_s(\boldsymbol{x})]\\
&\text{(iii)}\quad \varepsilon_i\varepsilon_j B_{ij}(\varepsilon\boldsymbol{x}) = B_{ij}(\boldsymbol{x})
\end{aligned} \qquad (5.3.53)$$

所给出的，对福克-普朗克方程这一特殊情况，则只要令跳概率

$W(x|x')$等于零就可得到相应的结果。

必要条件 若对于查普曼-柯尔莫哥洛夫微分方程而不是仅仅局限于福克-普朗克方程来陈述这些条件的话，问题就较为简单。在3.4节我们定义了量$W(x|x')$，$A_i(x)$以及$B_{ij}(x)$（因为是考察齐次过程，所以全部推导过程都与时间无关），根据这一节获得了一个显而易见的结果，即从(5.3.46)得知细致平衡要求

$$W(x|x')p_s(x') = W(\varepsilon x'|\varepsilon x)p_s(x) \tag{5.3.54}$$

现在来考察漂移系数。为简单起见记

$$x = x + \delta \tag{5.3.55}$$

由(5.3.46)，得

$$\int_{|\delta|<K} \mathrm{d}\delta\delta_i p(\varepsilon x + \varepsilon\delta,\ \Delta t|\varepsilon x,\ 0)p_s(x)$$

$$= \int_{|\delta|<K} \mathrm{d}\delta\delta_i p(x,\ \Delta t|x+\delta,\ 0)p_s(x+\delta) \tag{5.3.56}$$

（这里为避免与ε_i相混淆，我们在积分区域中用K代替了ε）用Δt去除两边，再取$\Delta t\to 0$的极限，左边就产生了

$$\varepsilon_i A_i(\varepsilon x)p_s(x) + O(K) \tag{5.3.57}$$

在右边我们可以把被积函数写为

$$p(x+\delta-\delta,\ \Delta t|x+\delta,\ 0)p_s(x+\delta)$$
$$= p(x-\delta,\ \Delta t|x,\ 0)p_s(x)$$
$$+ \sum_j \delta_j \frac{\partial}{\partial x_j}[p(x-\delta,\ \Delta t|x,\ 0)p_s(x)] + O(\delta^2) \tag{5.3.58}$$

所以右边是

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int \mathrm{d}\delta\left\{\delta_i p(x-\delta,\ \Delta t|x,\ 0)p_s(x)\right.$$
$$\left.+\sum_j \delta_i\delta_j \frac{\partial}{\partial x_j}[p(x-\delta,\ \Delta t|x,\ 0)\cdot p_s(x)]\right\} + O(K)$$
$$= -A_i(x)p_s(x) + \sum_j \frac{\partial}{\partial x_j}[B_{ij}(x)p_s(x)] + O(K) \tag{5.3.59}$$

其中我们利用了在3.4节说明过的一个事实，即那些含有δ次比δ^2还要高的δ高阶项都是与K同阶的。令$K\to 0$，就得到

$$\varepsilon_i A_i(\varepsilon x)p_s(x) = -A_i(x)p_s(x) + \sum_j \frac{\partial}{\partial x_j}[B_{ij}(x)p_s(x)] \tag{5.3.60}$$

关于 $B_{ij}(x)$ 的条件可类似地获得，但此时由于主项是 $O(\delta^2)$ 故不会出现如(5.3.60)右边第2项那种项了。我们得到

$$\varepsilon_i\varepsilon_j B_{ij}(\varepsilon x) = B_{ij}(x) \tag{5.3.61}$$

自然，剩下的第3个条件是 $p_s(x)$ 应为微分查普曼-柯尔莫哥洛夫方程的平稳解。这并不是一个平凡的条件，而且一般地说它还是跟其他两个条件独立的。

充分条件 现在来证明(5.3.53)是充分条件。假定这些条件被满足，即 $p_s(x)$ 是微分查普曼-柯尔莫哥洛夫方程的平稳解，并且 $p(x, t|x', 0)$ 是微分查普曼-柯尔莫哥洛夫方程的一个解。我们来考察

$$\hat{p}(x, t|x', 0) \equiv p(\varepsilon x', t|\varepsilon x, 0)p_s(x)/p_s(x') \tag{5.3.62}$$

显然

$$\hat{p}(x, 0|x', 0) = \delta(x - x') = p(x, 0|x', 0) \tag{5.3.63}$$

下面我们把 $\hat{p}$ 代入微分查普曼-柯尔莫哥洛夫方程，并且说明由于 $p(x', t|x, 0)$ 关于变量 x 服从后向微分查普曼-柯尔莫哥洛夫方程，$\hat{p}$ 必是前向微分查普曼-柯尔莫哥洛夫方程的一个解。

这可以明确地做到。为了清晰起见我们要简化记号，用

$$\left.\begin{array}{lll} \hat{p} & \text{取代} & \hat{p}(x, t|x', 0) \\ p_s & \text{取代} & p_s(x) \\ p_s' & \text{取代} & p_s(x') \\ p(x) & \text{取代} & p(x', t|x, 0) \end{array}\right\} \tag{5.3.64}$$

我们一项项地考察。

i) 漂移项：

$$\begin{aligned} -\sum_i \frac{\partial}{\partial x_i}(A_i\hat{p}) &= -\sum_i \frac{\partial}{\partial x_i}(A_i p(\varepsilon x)p_s/p_s') \\ &= -\sum_i \Big\{ \frac{\partial}{\partial x_i}[A_i p_s]\, p(\varepsilon x) \\ &\quad + A_i p_s \frac{\partial}{\partial x_i} p(\varepsilon x) \Big\} \Big/ p_s' \end{aligned} \tag{5.3.65}$$

iii) 扩散项：

$$\frac{1}{2}\sum_{i}\frac{\partial^2}{\partial x_i\partial x_j}(B_{ij}\hat{p}) = \frac{1}{2}\sum_{i,j}\Big[\frac{\partial^2}{\partial x_i\partial x_j}(B_{ij}p_s)p(\varepsilon x) + 2\frac{\partial}{\partial x_i}(B_{ij}p_s)\frac{\partial}{\partial x_j}p(\varepsilon x) + B_{ij}p_s\frac{\partial^2}{\partial x_i\partial x_j}p(\varepsilon x)\Big]\Big/p_s. \tag{5.3.66}$$

iii) 跳跃项：

$$\int dz[W(x|z)\hat{p}(z,\ t|x',\ 0) - W(z|x)\hat{p}(x,\ t|x',\ 0)]$$

$$= \int dz[W(x|z)p_s(z)p(\varepsilon x',\ t|\varepsilon x,\ 0) - W(z|x)p_s(x)p(\varepsilon x',\ t|\varepsilon x)]/p_s \tag{5.3.67}$$

利用 $p_s(x)$ 是平稳微分查普曼-柯尔莫哥洛夫方程的一个解就得到

$$-\sum_{i}\Big[\frac{\partial}{\partial x_i}(A_ip_s) + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i\partial x_j}(B_{ij}p_s)\Big] - \int dz W(z|x)p_s(x)$$

$$= -\int dz W(x|z)p_s(z) \tag{5.3.68}$$

再利用对W的细致平衡条件(5.3.53(i))，上式就

$$= -\int dz W(\varepsilon z|\varepsilon x)p_s(x) \tag{5.3.69}$$

现在代入

$$y = \varepsilon x$$
$$y' = \varepsilon x' \tag{5.3.70}$$

并且把上述三部分合起来之后，考虑到(5.3.68、69)它又

$$= \Big\{-\sum_{i}\varepsilon_iA_i(\varepsilon y)p_s(y)\Big[\frac{\partial}{\partial y_i}p(y)\Big] + \sum_{i,j}\varepsilon_i\varepsilon_j\frac{\partial}{\partial y_j}[B_{ij}(\varepsilon y)p_s(y)]\Big[\frac{\partial}{\partial y_i}p(y)\Big] + \frac{1}{2}\sum_{i,j}\varepsilon_i\varepsilon_jB_{ij}(\varepsilon y)p_s(y)\Big[\frac{\partial^2}{\partial y_i\partial y_j}p(y)\Big]$$

$$+\int dz[W(\varepsilon y|z)p_s(z)p(\varepsilon z, t|y', 0)$$

$$-W(\varepsilon y|z)p_s(z)p(y', t|y, 0)]\Big\}\Big/p_s(y') \qquad (5.3.71)$$

在代入了细致平衡条件(5.3.53)后，上式

$$=\Big\{\sum_i A_i(y)\frac{\partial}{\partial y_i}p(y', t|y, 0)$$

$$+\frac{1}{2}\sum_{i,j}B_{ij}(y)\frac{\partial^2}{\partial y_i\partial y_j}p(y', t|y, 0)$$

$$+\int dz[W(z|y)p(z, t|y', 0)$$

$$-W(z|y)p(y', t|y, 0)\Big\}p_s(y)/p_s(y') \qquad (5.3.72)$$

到这一步，我们就可以把花括号内的项看成为后向微分查普曼-柯尔莫哥洛夫算子(参见第3.6节(3.6.4)式)。由于过程是齐次的，使得

$$p(y', t|y, 0)=p(y, 0|y, -t)$$

我们得知

$$(5.3.72)=\frac{\partial}{\partial t}[p(y', t|y, 0)p_s(y)/p_s(y')]$$

$$=\frac{\partial}{\partial t}\hat{p}(x, t|x', 0) \qquad (5.3.73)$$

这意味着(5.3.62)所定义的 $\hat{p}(x, t|x', 0)$ 满足前向微分查普曼-柯尔莫哥洛夫方程。考虑到 $p(x, t|x', 0)$ 的初值条件，又因为在 $t=0$ 时 $\hat{p}(x, t|x', 0)$ 与 (5.3.63) 式所给的相同而且方程有唯一解，我们就证明了只要细致平衡条件(5.3.53)满足了则细致平衡也必得到满足，因而充分性得证。

注解

(i) 只有偶变量时，如果所有的 ε_i 都是 $+1$ 则条件就变得颇为简单了，此时条件简化为

$$W(x|x')p_s(x') = W(x'|x)p_s(x) \tag{5.3.74}$$

$$A_i(x)p_s(x) = \frac{1}{2}\sum_j \frac{\partial}{\partial x_j}[B_{ij}(x)p_s(x)] \tag{5.3.75}$$

$$B_{ij}(\varepsilon x) = B_{ij}(x) \text{ (*)} \tag{5.3.76}$$

上面最后的式子是平凡的。条件(5.3.75)完全与位势条件(5.3.21)相同，它表示在平稳状态下的概率流 J 不存在。

条件(5.3.74, 75)合在一起可得知 $p_s(x)$ 满足平稳微分查普曼-柯尔莫哥洛夫方程，这不是一般条件(5.3.53)所代表的情况。

(ii) 福克-普朗克方程的情况：范坎彭[5.7]与格雷厄姆-哈肯[5.9]引进了可逆的与不可逆的漂移部分之概念。不可逆的漂移是指

$$D_i(x) = \frac{1}{2}[A_i(x) + \varepsilon_i A_i(\varepsilon x)] \tag{5.3.77}$$

而可逆的漂移则是指

$$I_i(x) = \frac{1}{2}[A_i(x) - \varepsilon_i A_i(\varepsilon x)] \tag{5.3.78}$$

利用下式定义的位势

$$p_s(x) = \exp[-\phi(x)] \tag{5.3.79}$$

就可看出在福克-普朗克方程的情况下细致平衡的条件可以写成

$$\varepsilon_i\varepsilon_j B_{ij}(\varepsilon x) = B_{ij}(x) \tag{5.3.80}$$

$$D_i(x) - \frac{1}{2}\sum_j \frac{\partial}{\partial x_j}[B_{ij}(x)] = -\frac{1}{2}\sum_j B_{ij}(x)\frac{\partial\phi(x)}{\partial x_j} \tag{5.3.81}$$

$$\sum_i\left[\frac{\partial}{\partial x_i}I_i(x) - I_i(x)\frac{\partial\phi(x)}{\partial x_i}\right] = 0 \tag{5.3.82}$$

其中最后的方程是把(5.3.53(i))代入 $p_s(x)$ 满足的平稳福克-普朗克方程后而得到的。与讨论位势条件时的情况一样，我们可以看到(5.3.81)给出了 $\frac{\partial\phi}{\partial x_j}$ 的方程，而且只有当 $D_i(x)$ 及 $B_{ij}(x)$ 具

(*) 原文等式左边误为 $B_{ij}(x)$。——译注

有某些条件时这个方程才被满足。若矩阵 $B_{ij}(x)$ 是可逆的，则这些条件的形式为

$$\frac{\partial \hat{Z}_i}{\partial x_j} = \frac{\partial \hat{Z}_j}{\partial x_i} \tag{5.3.83}$$

其中

$$\hat{Z}_i = \sum_k B_{ik}^{-1}(x)\Big[2D_k(x) - \sum_j \frac{\partial}{\partial x_j} B_{kj}(x)\Big] \tag{5.3.84}$$

并且有

$$p_s(x) = \exp[-\phi(x)] = \exp\left(\int^x dx' \cdot \hat{Z}\right) \tag{5.3.85}$$

这样一来，如同在零概率流的情况下一样，$p_s(x)$ 能够明显地作为一个积分被确定。

(iii) 微分查普曼-柯尔莫哥洛夫方程的前向算子与后向算子之间之联系是由细致平衡所给定的。充分性条件的证明等于说明了只要 $f(x, t)$ 是前向微分查普曼-柯尔莫哥洛夫方程的一个解，则

$$\tilde{f}(x, t) = f(\varepsilon x, -t)/p_s(x) \tag{5.3.86}$$

必是后向微分查普曼-柯尔莫哥洛夫方程的一个解。这个结果在 5.3.7 节中构造本征函数时将会用到。

5.3.6 福克-普朗克方程中的若干细致平衡例子

a) 关于布朗运动的克雷默斯方程[5.10]

考察在涨落环境中的一个粒子运动。它是个一维的运动且粒子的状态由其位置 x 与速度 v 来描写。这就给出了微分方程

$$\frac{dx}{dt} = v \tag{5.3.87}$$

以及

$$m\frac{dv}{dt} = -V'(x) - \beta v + \sqrt{2\beta RT}\,\xi(t) \tag{5.3.88}$$

它实质上是朗之万方程(1.2.4)，不过这里为了简单起见我们记

$$6\pi\eta a = \beta$$

而且 $V(x)$ 是位势，其梯度 $V'(x)$ 产生了一个作用于粒子上的力。在如第 4.1 节中那样假设了物理的涨落力 $\xi(t)$ 要理解为

$$\xi(t)\mathrm{d}t = \mathrm{d}W(t) \tag{5.3.89}$$

之后，就有随机微分方程

$$\mathrm{d}x = v\mathrm{d}t \tag{5.3.90}$$

$$m\mathrm{d}v = -[V'(x) + \beta v]\mathrm{d}t + \sqrt{2\beta kT}\mathrm{d}W(t) \tag{5.3.91}$$

其对应的福克-普朗克方程是

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(vp) + \frac{1}{m}\frac{\partial}{\partial v}\{[V'(x) + \beta v]p\} + \frac{\beta kT}{m^2}\frac{\partial^2 p}{\partial v^2} \tag{5.3.92}$$

若引进几个变量

$$y = x\sqrt{m/kT} \tag{5.3.93}$$

$$u = v\sqrt{m/kT} \tag{5.3.94}$$

$$U(y) = V(x)/kT \tag{5.3.95}$$

$$\gamma = \beta/m \tag{5.3.96}$$

就可以使这个方程稍许简单些，使福克-普朗克方程具有形式

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial y}(up) + \frac{\partial}{\partial u}[U'(y)p] + \gamma\frac{\partial}{\partial u}\left(up + \frac{\partial p}{\partial u}\right) \tag{5.3.97}$$

我们以后称之为克雷默斯(Kramers)方程。

在这里，如 5.3.4 节中说明过的那样，(位置) y 是个偶变量而且 (速度) u 是个奇变量。漂移与扩散系数可以表示为

$$\boldsymbol{A}(y, u) = \begin{bmatrix} u \\ -U'(y) - \gamma u \end{bmatrix} \tag{5.3.98}$$

$$\underset{\sim}{\boldsymbol{B}}(y, u) = \begin{bmatrix} 0 & 0 \\ 0 & 2\gamma \end{bmatrix} \tag{5.3.99}$$

以及

$$\varepsilon\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} y \\ -u \end{bmatrix} \tag{5.3.100}$$

对于所有条件我们则可以逐一验证。

条件(5.3.53(iii))是明显满足的。由于 $\tilde{\boldsymbol{B}}$ 不是可逆的，故条件(5.3.53(iii))是有些退化的。它可以表示为

$$\varepsilon \boldsymbol{A}(y, -u)p_s(y, u) = -\boldsymbol{A}(y, u)p_s(y, u) + \begin{bmatrix} 0 \\ 2\gamma \dfrac{\partial p_s}{\partial u} \end{bmatrix} \tag{5.3.101}$$

或者更完整地写成

$$\begin{bmatrix} -u \\ U'(y) - \gamma u \end{bmatrix} p_s = \begin{bmatrix} -u \\ U'(y) + \gamma u \end{bmatrix} p_s + \begin{bmatrix} 0 \\ 2\gamma \dfrac{\partial p_s}{\partial u} \end{bmatrix} \tag{5.3.102}$$

第一行是个恒等式，而第二行断言

$$-up_s(y, u) = \frac{\partial p_s}{\partial u} \tag{5.3.103}$$

即

$$p_s(y, u) = \exp\left(-\frac{1}{2}u^2\right) f(y) \tag{5.3.104}$$

它意味着只要 $p_s(y, u)$ 表示成(5.3.104)的形式，则细致平衡条件就被满足了。现在我们必须验证(5.3.104)是否确实给出了克雷默斯方程(5.3.97)的一个平稳解。最后一个括号是零值，剩下了

$$0 = -u\frac{\partial f}{\partial y} - U'(y)uf \tag{5.3.105}$$

它意味着

$$f(y) = \mathcal{N}\exp[-U(y)] \tag{5.3.106}$$

以及

$$p_s(y, u) = \mathcal{N}\exp\left[-U(y) - \frac{1}{2}u^2\right] \tag{5.3.107}$$

用原来的变量(x, v)来表示，则成了

$$p_s(x, v) = \mathcal{N}\exp\left[-\frac{V(x)}{kT} - \frac{mv^2}{2kT}\right] \tag{5.3.108}$$

这是统计力学中人们所熟知的玻耳兹曼分布。这里我们要注意分母 kT 来自于(5.3.88)中涨落的设定的系数$\sqrt{2\beta kT}$。这样，我们要用一个宏观的方程式并且附加上一个涨落力，其大小则按照下

列要求来确定，即要求此解是与温度T相对应的玻耳兹曼分布。

我们也已经精确地得到了正确的分布函数。这意味着布朗运动是由形如(5.3.87, 88)的马尔可夫过程来刻划的，这个假设一定是相当有效的。

b) 确定性运动

这里我们令$B_{ij}(x)$及$W(x|x')$为零，使得细致平衡条件仅仅是

$$\varepsilon_i A_i(\varepsilon x) = -A_i(x) \tag{5.3.109}$$

既然现在考察的是刘维方程(见3.5.3节)，则坐标为x的点的运动就由常微分方程

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = A[x(t)] \tag{5.3.110}$$

所刻划。设(5.3.110)的在$t=0$时通过点y的一个解是

$$q[t, y] \tag{5.3.111}$$

它因而满足

$$q[0, y] = y \tag{5.3.112}$$

于是关系式(5.3.109)推出其时间反演解

$$\varepsilon q(-t, \varepsilon y) \tag{5.3.113}$$

也是(5.3.110)的一个解，而且由于

$$\varepsilon q(0, \varepsilon y) = \varepsilon\varepsilon y = y \tag{5.3.114}$$

也就是说，两个解的初值条件是相同的，它们必须是恒同的，即

$$\varepsilon q(-t, \varepsilon y) = q(t, y) \tag{5.3.115}$$

由于稳态下之联合概率可表示为

$$\begin{aligned} p_s(x, t; x', t') &= \int \mathrm{d}y p_s(x, t; x', t'; y, 0) \\ &= \int \mathrm{d}y \delta[x - q(t, y)]\delta[x' - q(t', y)]p_s(y) \end{aligned} \tag{5.3.116}$$

以及

$$p_s(\varepsilon x', -t'; \varepsilon x, -t) = \int \mathrm{d}y \delta[\varepsilon x - q(-t, y)]$$

$$\cdot \delta[\varepsilon x' - q(-t', y)]p_s(y) \tag{5.3.117}$$

把变量 y 改变为 εy 并注意到

$$p_s(y) = p_s(\varepsilon y)$$

和

$$\mathrm{d}\varepsilon y = \mathrm{d}y$$

就使得

$$(5.3.117) = \int \mathrm{d}y \delta[x - \varepsilon q(-t, \varepsilon y)]$$

$$\cdot \delta[x' - \varepsilon q(-t', \varepsilon y)]p_s(y) \tag{5.3.118}$$

再利用(5.3.115)

$$= \int \mathrm{d}y \delta[x - q(t, y)]\delta[x' - q(t', y)]p_s(y) \tag{5.3.119}$$

$$= p_s(x, t'; x', t') \tag{5.3.120}$$

利用 p_s 仅与时间的差值有关这种平稳性质，我们便看到细致平衡是满足的。

不用说，由于原先的证明对于这种确定性系统是有效的，这里的直接证明并不是必需的。进一步讲，既然任何确定性的一阶微分方程系统都可以变换到一个刘维方程，这就使得这种直接证明一般地说也不是必要的。这里之所以包括这部分内容仅仅是出于兴趣。

但是给出隐藏在细致平衡的这种证明后面的简要哲理却是重要的。在一些细致平衡是重要的物理系统中，常常具有令人难以置信的大量变量，其数量至少为 10^{20} 的量级。这些变量(比如，气体中粒子的动量和速度)是出现在分布函数中的变量，而分布函数则由于这些变量遵从像牛顿运动定律这样一类确定性运动方程而服从一个刘维方程。

可以直接证明，对于适当的相互作用形式来说，这些牛顿定律是服从微观可逆性原理的，这个原理意味着这些定律可以纳入(5.3.110)的形式中来，其中 $A(x)$ 服从可逆性条件(5.3.109)。

在这种系统中的宏观可观测量是这些变量(如压力、温度、粒子的密度)的函数,并且在作了适当的变量变换之后可用向量 x 的前几个少量分量加以表示。

这样一来,我们假设 x 可以表示为

$$x = (a, \hat{x}) \tag{5.3.121}$$

其中向量 a 表示宏观可观测量而 $\hat{x}$ 则是其余的量。于是实际上我们感兴趣的是

$$\tilde{p}(a_1, t_1;\ a_2, t_2;\ a_3, t_3;\ \cdots) = \int\int\cdots\int \mathrm{d}\hat{x}_1 \mathrm{d}\hat{x}_2 \cdots p(x_1, t_1;\ x_2, t_2;\ x_3, t_3;\ \cdots) \tag{5.3.122}$$

按照我们上面的推理,由微观可逆性得知 p 并且因而 $\tilde{p}$ 都服从细致平衡条件,但 $\tilde{p}$ 当然不服从刘维方程。如果 $\tilde{p}$ 原来就是或可证明是在某种近似程度上服从一个马尔可夫运动方程,那末对 $\tilde{p}$ 来说我们就一定保持着这种细致平衡性质(形式与 p 所具有的相同)。在此意义下我们可以说细致平衡条件 (5.3.43) 是由运动方程的微观可逆性所导出的。

c) 奥恩斯坦-乌伦贝克过程:线性系统中的翁萨格关系

使人关心细致平衡的绝大部分系统,都可以用奥恩斯坦-乌伦贝克过程去近似,也就是说,这意味着我们假设

$$A_i(x) = \sum_j A_{ij} x_j \tag{5.3.123}$$

$$B_{ij}(x) = B_{ij} \tag{5.3.124}$$

不过细致平衡条件不是不重要的,即还要有

$$\sum_j (\varepsilon_i \varepsilon_j A_{ij} + A_{ij}) x_j = \sum_j B_{ij} \frac{\partial}{\partial x_j} \log p_s(x) \tag{5.3.125}$$

和

$$\varepsilon_i \varepsilon_j B_{ij} = B_{ij} \tag{5.3.126}$$

方程(5.3.125)定性地表明 $p_s(x)$ 是高斯分布,因为 $\log p_s(x)$ 的导数关于 x 是线性的。不仅如此,由于此式的左边未含常数项,这个高斯分布还必然有零均值,因而可写出

$$p_s(x) = \mathcal{N}\exp\left(-\frac{1}{2}x^T\sigma^{-1}x\right) \tag{5.3.127}$$

现在可把(5.3.127)代入平稳福克-普朗克方程中并作整理而得到

$$-\sum_i A_{ii} - \frac{1}{2}\sum_{i,j} B_{ij}\sigma^{-1}_{ij} + \sum_{k,j}\left(\sum_i \sigma^{-1}_{ki}A_{ij} + \frac{1}{2}\sum_{i,l}\sigma^{-1}_{ki}B_{il}\sigma^{-1}_{lj}\right)x_k x_j = 0 \tag{5.3.128}$$

(这里利用了矩阵 σ 的对称性)。若其系数的对称部分是零则二次项就消掉了。这个条件可表示为矩阵的形式

$$\sigma^{-1}A + A^T\sigma^{-1} = -\sigma^{-1}B\sigma^{-1} \tag{5.3.129}$$

或者

$$A\sigma + \sigma A^T = -B \tag{5.3.130}$$

若(5.3.129)被满足,则常数项也消掉了。自然,在代换

$$\begin{aligned} &A \to -A \\ &BB^T \to B \end{aligned} \tag{5.3.131}$$

之下 (5.3.130) 正好就是 4.4.6 节中用随机微分方程技巧所导出的(4.4.51)。现在我们能够以其最精巧的形式表示出细致平衡条件。若用

$$\varepsilon = \mathrm{diag}(\varepsilon_1, \varepsilon_2, \varepsilon_3, \cdots) \tag{5.3.132}$$

定义矩阵 ε,显然

$$\varepsilon^2 = 1 \tag{5.3.133}$$

那末在矩阵记号下条件(5.3.125, 126)变为

$$\varepsilon A\varepsilon + A = -B\sigma^{-1} \tag{5.3.134}$$

$$\varepsilon B\varepsilon = B \tag{5.3.135}$$

位势条件(5.3.83)则完全等价于 σ 的对称性。

正如 5.3.4 节(5.3.49)已指明的那样,细致平衡要求

$$\varepsilon\sigma = \sigma\varepsilon \tag{5.3.136}$$

记住这点,并利用(5.3.130)

$$A\sigma + \sigma A^T = -B$$

和从(5.3.134)得到的

$$\varepsilon A\varepsilon\sigma + A\sigma = -B \tag{5.3.137}$$

就产生

$$\varepsilon A\varepsilon\sigma = \sigma A^T \tag{5.3.138}$$

再利用(5.3.136)，就有

$$\varepsilon(A\sigma) = (A\sigma)^T\varepsilon \tag{5.3.139}$$

这些就是著名的翁萨格关系；请参见文献翁萨格 (Onsager, L.) [5.11]及卡西米尔(Oasimir.H.B.G.) [5.12]。其推导紧密地追随着范坎彭的工作[5.6]。

可以用引进唯象力的方法简单地对这些结论作出解释。这些力被定义为位势

$$\phi = \log[p_s(x)]$$

的梯度：

$$F(x) = -\nabla\phi(x) = \sigma^{-1}x \tag{5.3.140}$$

(物理学中，ϕ/kT 是系统的熵)。因为 $A_i(x)$ 具有线性形式(5.3.123)，$\langle x\rangle$的精确运动方程是

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x\rangle = A\langle x\rangle = A\sigma F(\langle x\rangle) \tag{5.3.141}$$

于是，只要通量

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x\rangle$$

是通过定义为

$$L = A\sigma \tag{5.3.142}$$

的矩阵L而与 $F(\langle x\rangle)$成线性关系的，那末(5.3.139)所说的就是

$$\varepsilon L\varepsilon = L^T \tag{5.3.143}$$

或者

$$\boxed{\begin{aligned} L_{ij} &= L_{ji} \quad (\varepsilon_i \text{与} \varepsilon_j \text{符号相同时}) \\ L_{ij} &= -L_{ji} \quad (\varepsilon_i \text{与} \varepsilon_j \text{符号不同时}) \end{aligned}} \tag{5.3.144}$$

另外还要注意，

$$\left.\begin{aligned} \varepsilon B\varepsilon &= B \\ \varepsilon\sigma\varepsilon &= \sigma \end{aligned}\right\} \tag{5.3.145}$$

意味着只要 ε_i 跟 ε_j 符号相反则 B_{ij} 及 σ_{ij} 为零。

在所有 ε_i 的符号均相同的特殊情况下，我们得到

$$L = L^T \tag{5.3.146}$$

并且会看到，由于 σ 是对称与正定的故它必有实平方根 $\sigma^{1/2}$，从而

$$\hat{A} \equiv \sigma^{1/2} A \sigma^{1/2} \tag{5.3.147}$$

是个对称阵，使得 A 是相似于一对称阵的矩阵。所以 A 的所有本征值都是实值的。

d) 翁萨格关系的意义：涨落耗散定理

翁萨格关系是对一组宏观可观测量而言的，所以提供了细致平衡的一个易观测到的结果。而细致平衡本身如 (b) 中已概述过的则是运动微观方程之可逆性的结果。但是在给定情况下来检验其成立否却要求了解协方差矩阵 σ。

幸运的是此时若是个热力学平衡系统，其细致平衡一定满足，则统计力学令给出平稳分布的形式。其原理类似于朗之万（在 1.2.2 节中）所用过的。下面我们用一个例子来说明尼奎斯公式的推导。

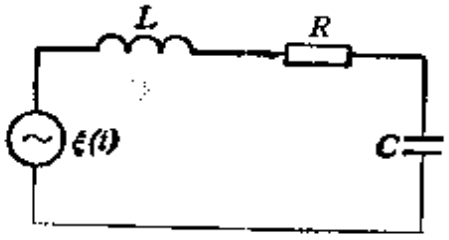


图 5.4 用于导出尼奎斯公式的电路

假设有一电路如图 5.4，假设电路中存在由于系统有非零温度而引起的起伏电压和由于系统与大中性电荷储存器（如离子溶液）相连而引起的起伏电荷。由电荷 q 的守恒性及基尔霍夫电压定律产生了电方程。电荷方程取

$$\frac{\mathrm{d}q}{\mathrm{d}t} = i - \gamma q + \Delta q(t) \tag{5.8.148}$$

的形式，其中我们使电荷储存器上电荷的增加率跟电流 i 减去进入它的泄漏量 γq 再加上由大电容器引起的起伏项 $\Delta q(t)$ 所得的量相等，这些量都可立即算出。

把沿电路的所有电压加起来，并把一个可能发生的起伏电压 $\Delta V(t)$ 也包括在内，就得到基尔霍夫电压定律：

$$\frac{\mathrm{d}i}{\mathrm{d}t} = \frac{1}{L}\left[-\frac{q}{e} - iR + \Delta V(t) \right] \tag{5.3.149}$$

现在我们假设 $\Delta q(t)$ 及 $\Delta V(t)$ 是白噪声. 在最一般情况下可写出

$$\Delta q(t) = b_{11}\xi_1(t) + b_{12}\xi_2(t) \tag{5.3.150}$$

$$\frac{1}{L}\Delta V(t) = b_{21}\xi_1(t) + b_{22}\xi_2(t) \tag{5.3.151}$$

其中 $\xi_1(t)$ 与 $\xi_2(t)$ 是两个不相关的朗之万电源，即

$$\xi_1(t)\mathrm{d}t = \mathrm{d}W_1(t) \tag{5.3.152}$$

$$\xi_2(t)\mathrm{d}t = \mathrm{d}W_2(t) \tag{5.3.153}$$

其中 $W_1(t)$ 与 $W_2(t)$ 是独立的维纳过程.

于是这里有

$$A = -\begin{bmatrix} \gamma & -1 \\ \dfrac{1}{LC} & \dfrac{R}{L} \end{bmatrix} \tag{5.3.154}$$

系统的总能量是

$$E = \frac{L}{2}i^2 + \frac{1}{2L}q^2 \tag{5.3.155}$$

根据统计力学理论，$p_s(q, i)$ 由温度 T 时的玻耳兹曼分布给定，即

$$p_s(q, i) = \mathcal{N}\exp(-E/kT)$$

$$= \mathcal{N}\exp\left(-\frac{Li^2}{2kT} - \frac{q^2}{2CkT}\right) \tag{5.3.156}$$

其中 k 是玻耳兹曼常数，从而协方差矩阵是

$$\sigma = \begin{bmatrix} kCT & 0 \\ 0 & kT/L \end{bmatrix} \tag{5.3.157}$$

现在可以检验翁萨格关系了：

$$A\sigma = -\begin{bmatrix} \gamma kTC & -kT/L \\ kT/L & RkT/L^2 \end{bmatrix} \tag{5.3.158}$$

对于这个系统来说，在时间反演之下其总电荷 q 是偶的而电流 i 是奇的. 于是

$$\varepsilon = \mathrm{diag}(1, -1) \tag{5.3.159}$$

并且显然

$$(A\sigma)_{12} = -(A\sigma)_{21} \tag{5.3.160}$$

只不过是对称性的结果，因而为(5.3.158)所满足。

此外，这里还有

$$B = -(A\sigma + \sigma A^T) = 2kT\begin{bmatrix} \gamma C & 0 \\ 0 & R/L^2 \end{bmatrix} \tag{5.3.161}$$

从而

$$b_{12} = b_{21} = 0 \tag{5.3.162}$$

$$b_{11} = \sqrt{2kT\gamma C} \tag{5.3.163}$$

$$b_{22} = \sqrt{2kTR}/L$$

而且正如物理直观所要求的，有 $B_{12} = B_{21} = 0$，它表明这两个起伏源是由不同原因引起的，应该是相互独立的。

这是一些关于涨落耗散的结果，其涨落的大小 b_{ij} 由耗散项 γ 及 R 所确定。事实上，结果(5.3.163)恰好就是在1.4.4节讨论了的尼奎斯定理。由于线路中的噪声电压是由

$$\Delta V(t) = \sqrt{2kTR}\xi_1(t) \tag{5.3.164}$$

给出的，所以

$$\langle \Delta V(t+\tau)\Delta V(t)\rangle = 2kTR\delta(\tau) \tag{5.3.165}$$

而这就是(1.4.49)形式下的尼奎斯定理。

γ 项与 R 项之所以称为耗散项是由于它们产生了能量的耗散。事实上，确定性地(也就是说在令噪声等于零后)我们有

$$\frac{dE}{dt} = -I^2R - \frac{\gamma q^2}{C} \tag{5.3.166}$$

它明显展现出能量的耗散。

5.3.7 多变量齐次过程情况下的本征函数方法

类似于5.2.5节来进行讨论。假设前向福克-普朗克方程存在一组完备本征函数 $P_\lambda(x)$ 并且后向福克-普朗克方程也存在一组完备本征函数 $Q_\lambda(x)$。于是

$$-\sum_i \partial_i[A_i(x)P_\lambda(x)] + \frac{1}{2}\sum_{i,j}\partial_i\partial_j[B_{ij}(x)P_\lambda(x)]$$

$$= -\lambda P_\lambda(x) \tag{5.3.167}$$

$$\sum A_i(x)\partial_i Q_{\lambda'}(x) + \frac{1}{2}\sum B_{ij}(x)\partial_i\partial_j Q_{\lambda'}(x) = -\lambda' Q_{\lambda'}(x) \tag{5.3.168}$$

不论 $Q_{\lambda'}(x)$ 和 $P_\lambda(x)$ 是满足吸收的还是反射的边界条件，我们都可以用十分类似于 5.2.4 节的方式证明

$$-(\lambda-\lambda')\int_R dx P_\lambda(x)Q_{\lambda'}(x) = 0 \tag{5.3.169}$$

从而 $P_\lambda(x)$ 与 $Q_{\lambda'}(x)$ 构成双正交组。但是只有当具有所有的 $e_i=1$ 的细致平衡成立时在 P_λ 与 $Q_{\lambda'}$ 之间才存在泛函关系。我们假设如果本征值 λ 的谱是离散的则

$$\int dx P_\lambda(x)Q_{\lambda'}(x) = \delta_{\lambda\lambda'} \tag{5.3.170}$$

[当谱是连续的时候，除了反射边界条件或当 $\lambda=0$ 是平稳态的情况外，只要用 $\delta(\lambda-\lambda')$ 代替克罗内克算子 $\delta_{\lambda\lambda'}$ 就行了。那末 $p_s(x)$ 的归一化给出

$$\int dx P_0(x)Q_0(x) = \int dx P_0(x) = 1 \tag{5.3.171}$$

使得此时在谱中也存在一个具有零本征值的离散点。]

若细致平衡成立，由于在 5.3.5 节(III)中已看到

$p(\varepsilon x, -t)/p_s(x)$

是后向福克-普朗克方程的解，所以由解的唯一性可写出

$$Q_\lambda(x) = \eta_\lambda P_\lambda(\varepsilon x)/p_s(x) \tag{5.3.172}$$

在这里 $\eta_\lambda = \pm 1$，并不是那样地不确定，因为若 $\lambda=\lambda'$，就可以把(5.3.170)写成

$$\int dx \frac{P_\lambda(x)P_\lambda(\varepsilon x)}{p_s(x)} = \eta_\lambda \tag{5.3.173}$$

我们虽不能先验地确定 η_λ，但是借适当的归一化可以把它选择为 ± 1。譬如说，若所有的 $e_i=-1$ 且 $P_\lambda(x)$ 是 x 的奇函数，则显然 η_λ 必然是 -1。

i) 只有偶变量的情况：本征值的负值性。若所有的 ε_i 都等于 1，则可写出

$$Q_\lambda(x) = P_\lambda(x)/p_s(x) \tag{5.3.174}$$

而且 η_λ 总可以置为 1。

因此本征函数的展开会有许多地方与单变量场合相同。

本征函数组的完备性是在每个具体情况下要证明的一件事。若所有的 ε_i 都等于 1，则可证明福克-普朗克算子是某个希尔伯特空间中的自共轭半负定算子。

更明确些讲，我们把前向福克-普朗克算子表示为

$$\mathscr{L}(x) = -\sum_i \partial_i A_i(x) + \frac{1}{2}\sum_{i,j}\partial_i\partial_j B_{ij}(x) \tag{5.3.175}$$

而后向福克-普朗克算子表示成

$$\mathscr{L}^*(x) = \sum_i A_i(x)\partial_i + \frac{1}{2}\sum_{i,j} B_{ij}(x)\partial_i\partial_j \tag{5.3.176}$$

这样一来，由对任意函数 $f(x)$ 有

$$\mathscr{L}(x)[f(x)p_s(x)] = p_s(x)\mathscr{L}^*(x)[f(x)] \tag{5.3.177}$$

的事实就可导致下面的事实，即若所有的 $\varepsilon_i = 1$，则借用 $p_s(x)$ 去除的办法就可将前向福克-普朗克方程的解变换为后向福克-普朗克方程的解。若由内积

$$(f, g) = \int_R \mathrm{d}x \frac{f(x)g(x)}{p_s(x)} \tag{5.3.178}$$

来定义一个函数 $f(x)$，$g(x)$，…的希尔伯特空间，则由(5.3.177)知

$$(f, \mathscr{L}g) = \int_R \mathrm{d}x \frac{f(x)}{p_s(x)} \cdot \mathscr{L}(x)\left[\frac{g(x)}{p_s(x)}p_s(x)\right] \tag{5.3.179}$$

$$= \int_R \mathrm{d}x f(x)\mathscr{L}^*(x)\left[\frac{g(x)}{p_s(x)}\right] \tag{5.3.180}$$

再利用分部积分，并用反射或吸收边界条件来去掉曲面项，上式就

$$= \int_R \mathrm{d}x \frac{g(x)}{p_s(x)}\mathscr{L}(x)[f(x)] \tag{5.3.181}$$

这样一来，在此希尔伯特空间中

$$(f, \mathscr{L}g) = (g, \mathscr{L}f)。 \tag{5.3.182}$$

这个条件足以保证 $\mathscr{L}(x)$ 的所有本征值都为实值的。为了证明其

半负定性，可注意对仅有偶变量的情况有[见(5.3.75)]

$$A_i(x)=\sum_j \partial_j[B_{ij}(x)p_s(x)]/2p_s(x) \tag{5.3.183}$$

所以对任何 $p(x)$

$$\begin{aligned}\mathscr{L}(x)p(x)&=\sum_{i,j}\partial_i\Big\{-\frac{\partial_j[B_{ij}(x)p_s(x)]p(x)}{2p_s(x)}\\&\quad+\frac{1}{2}\partial_j[B_{ij}(x)p(x)]\Big\}\\&=\frac{1}{2}\sum_{i,j}\partial_i\{B_{ij}(x)p_s(x)\partial_j[p(x)/p_s(x)]\}\end{aligned} \tag{5.3.184}$$

于是

$$\begin{aligned}(p,\ \mathscr{L}p)=\frac{1}{2}\int \mathrm{d}x p(x)/p_s(x)\\ \cdot\sum_{i,j}\partial_i\{B_{ij}(x)p_s(x)\partial_j[p(x)/p_s(x)]\}\end{aligned}$$

再利用分部积分(而且去掉曲面项后)，由 $B_{ij}(x)$是半正定的，就得到

$$\begin{aligned}(p,\ \mathscr{L}p)=-\int \mathrm{d}x B_{ij}(x)p_s(x)\\ \cdot\partial_i[p(x)/p_s(x)]\partial_j[p(x)/p_s(x)]\leqslant 0\end{aligned} \tag{5.3.185}$$

因此我们得到的结论是对任何本征函数 $P_\lambda(x)$ 来说 λ 都是实值的，并且

$$\lambda\geqslant 0 \tag{5.3.186}$$

(请记住 $-\lambda$ 也是本征值)。

ii) 变分原理。与这个性质相关联的是变分原理。假设选定了任一实值函数 $f(x)$并按本征函数把它展开为

$$f(x)=\sum_\lambda \alpha_\lambda P_\lambda(x) \tag{5.3.187}$$

进一步规定

$$(f,\ f)=\int \mathrm{d}x f(x)^2/p_s(x)=1$$

那末，

$$-(f,\ \mathscr{L}f)=\sum_\lambda \lambda\alpha_\lambda^2 \tag{5.3.188}$$

并且

$$(f,\ f)=\sum_{\lambda}\alpha_{\lambda}^{2}$$

显然，仅当只有 $\lambda=0$ 的项存在，即

$$\begin{aligned}&\alpha_0\neq 0\\&\alpha_\lambda=0\quad \text{对所有}\ \lambda\neq 0\end{aligned}\tag{5.3.189}$$

时 $-(f,\ \mathscr{L}f)$ 才具有极小值零。现在选取 $f(x)$ 与 $P_0(x)$ 正交而使 $\alpha_0=0$，我们看到 $(f,\ \mathscr{L}f)$ 的极小值出现在

$$\alpha_{\lambda_1}=1\tag{5.3.190}$$

且

$$\alpha_\lambda=0\quad \text{对所有其他的}\ \lambda$$

的情况下，其中 λ_1 是第二个最小的本征值。这意味着可在赋予条件

$$\begin{aligned}&(p,\ p)=1\\&(p,\ P_0)=0\end{aligned}\tag{5.3.191}$$

之下用极小化 $-(p,\ \mathscr{L}p)$ 的方法［它可置于 (5.3.185) 的形式之下］得到 $P_{\lambda_1}(x)$。此方法能够给出一个靠试探解估计本征函数的数值方法，在双稳态情况下非常有用。但是它只限于细致平稳成立的情况。

iii) 条件概率。如果假设了本征函数的完备性，则会发现条件概率可表示成

$$p(x,\ t|x_0,\ 0)=\sum_{\lambda}P_\lambda(x)Q_\lambda(x_0)e^{-\lambda t}\tag{5.3.192}$$

iv) 自相关矩阵。如果细致平衡成立，则平稳自相关矩阵［在利用了(5.3.172)后］为

$$\begin{aligned}G(t)=\langle x(t)x^T(0)\rangle=&\sum\eta_\lambda e^{-\lambda t}\left[\int_R dx x P_\lambda(x)\right]\\&\cdot\left[\int_R dx x P_\lambda(x)\right]^T\varepsilon\end{aligned}\tag{5.3.193}$$

它显然满足 5.3.4b 节中导出的条件

$$G(t)=\varepsilon G^T(t)\varepsilon\tag{5.3.194}$$

v) 谱矩阵。谱矩阵是

$$S(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-i\omega t}G(t)\,dt$$

$$=\frac{1}{2\pi}\left\{\int_0^{\infty}e^{-i\omega t}G(t)\,dt+\int_0^{\infty}e^{i\omega t}G^T(t)\,dt\right\} \qquad (5.3.195)$$

若为方便起见定义

$$\boldsymbol{U}_\lambda=\int_R dx\, xP_\lambda(x) \qquad (5.3.196)$$

则

$$\boldsymbol{S}(\omega)=\frac{1}{2\pi}\sum_\lambda\left[\frac{\eta_\lambda}{\lambda+i\omega}\boldsymbol{U}_\lambda\boldsymbol{U}_\lambda^T\varepsilon+\frac{\eta_\lambda}{\lambda-i\omega}\varepsilon\boldsymbol{U}_\lambda\boldsymbol{U}_\lambda^T\right] \qquad (5.3.197)$$

若某个 λ 是复的，则由 $\mathscr{L}(x)$ 取实值我们看到对应于 λ^* 的本征函数是 $[P_\lambda(x)]^*$，$\eta_\lambda=\eta_\lambda^*$ 而且 $[\boldsymbol{U}_\lambda]^*=\boldsymbol{U}_{\lambda^*}$。于是，谱可由(5.3.197)加上含有复本征值的那些项的复共轭项而得到。

在 $\varepsilon=1$ 而且因此 $\eta_\lambda=1$ 的情况下，谱矩阵有较简单的形式

$$\boldsymbol{S}(\omega)=\frac{1}{\pi}\sum_\lambda\frac{\lambda\boldsymbol{U}_\lambda\boldsymbol{U}_\lambda^T}{\lambda^2+\omega^2} \qquad (5.3.198)$$

它显然是个正定矩阵。单个变量 q 若是作为 $\boldsymbol{x}$ 的线性组合而由公式

$$q=\boldsymbol{m}\cdot\boldsymbol{x} \qquad (5.3.199)$$

构成的，则它的谱由

$$S_q(\omega)=\boldsymbol{m}^T\boldsymbol{S}(\omega)\boldsymbol{m}=\frac{1}{\pi}\sum_\lambda\frac{\lambda(\boldsymbol{m}\cdot\boldsymbol{U}_\lambda)^2}{\lambda^2+\omega^2} \qquad (5.3.200)$$

给出，而且是 ω 的严格递减函数。

在 $\varepsilon\neq1$ 的情况下，谱的正定性不再是那样显然了，尽管诸如结果(3.7.19)这类一般化考察证明了它必然是正定的。这里出现了一个重要的差别，它来自于现在 λ 不要求都是实数，并且这意味着会出现

$$1/[\lambda_R^2+(\omega-\lambda_I)^2] \qquad (5.3.201)$$

形式的分母，它们在远离 $\omega=0$ 的谱点上会给出一些峰态。

5.4 区域的首次出口时间(齐次过程)

在5.2.7节我们考察了一维情况的首次出口时间问题，这里来考察多维的类同问题。

这个类同问题就是要计算自具有边界 S 的某区域 R 中出发的一个粒子离开此区域的最早时刻。

如同一维时那样，我们来考察后向福克-普朗克方程在具有吸收边界 S 的条件，即

$$p(x', t|x, 0) = 0 (x \in S) \tag{5.4.1}$$

之下的求解问题。初始位置在 x 的粒子经过时间 t 后仍留在 R 中的概率是

$$G(x, t) = \int_R dx' p(x', t|x, 0) \tag{5.4.2}$$

若 T 是粒子离开 R 的时刻，则

$$\text{Prob}(T \geqslant t) = G(x, t) \tag{5.4.3}$$

由于过程是齐次的，我们看到 $G(x, t)$ 服从后向福克-普朗克方程

$$\partial_t G(x, t) = \sum_i A_i(x) \partial_i G(x, t) + 2 \sum_{i,j} B_{ij}(x) \partial_i \partial_j G(x, t) \tag{5.4.4}$$

(5.4.4)的初始条件产生于：

i) $p(x', 0|x, 0) = \delta(x - x')$ (5.4.5)

而使

$$G(x, 0) = \begin{cases} 1 & x \in R \\ 0 & \text{其他} \end{cases} \tag{5.4.6}$$

ii) 边界条件(5.4.1)要求

$$G(x, t) = 0 \quad x \in S \tag{5.4.7}$$

如同5.2.7节，我们看到用符号 $T(x)$ 所表示的从 x 出发跑出 R 的平均出口时间满足

$$\sum_i A_i(x) \partial_i T(x) + 2 \sum_{i,j} B_{ij}(x) \partial_i \partial_j T(x) = -1 \tag{5.4.8}$$

及边界条件

$$T(x) = 0 \quad x \in S \tag{5.4.9}$$

并且 n 阶矩

$$T_n(x) = \langle T^n \rangle = \int_0^\infty t^{n-1} G(x, t)\mathrm{d}t \tag{5.4.10}$$

满足

$$-nT_{n-1}(x) = \sum_i A_i(x)\partial_i T_n(x) + \sum_{i,j} B_{ij}(x)\partial_i\partial_j T_n(x) \tag{5.4.11}$$

及边界条件

$$T_n(x) = 0 \quad x \in S \tag{5.4.12}$$

包含反射区域。 有时能够把 R 的边界 S 分成两个区域 S_r 及 S_a，使得粒子碰到了 S_r 就被反射而碰到了 S_a 则被吸收。那末按 5.2.4 节的有关推导，$G(x, t)$ 的边界条件为

$$\sum_{ij} n_i B_{ij}(x)\partial_j G(x, t) = 0 \quad (x \in S_r) \tag{5.4.13}$$

$$G(x, t) = 0 \quad (x \in S_a) \tag{5.4.14}$$

并且因此

$$\sum_{ij} n_i B_{ij}(x)\partial_j T_n(x) = 0 \quad (x \in S_r) \tag{5.4.15}$$

$$T_n(x) = 0 \quad (x \in S_a) \tag{5.4.16}$$

5.4.1 平均出口时间问题的解

只有在一维情况或者存在某种特别的对称性可利用时，平均首次通过时间的基本偏微分方程才会简单解出。一些渐近近似可能提供十分有效的结果，但是我们将它安排在第 9 章研究。这里要做的是用一些例子来说明某些方法。

a) 具有旋转对称性的二维奥恩斯坦-乌伦贝克过程

假设一个粒子按照方程

$$\begin{aligned} \mathrm{d}x &= -kx\mathrm{d}t + \sqrt{D}\mathrm{d}W_1(t) \\ \mathrm{d}y &= -ky\mathrm{d}t + \sqrt{D}\mathrm{d}W_2(t) \end{aligned} \tag{5.4.17}$$

作运动。已给定其初始位置是 x_0, y_0，我们要知道它从区域

$$x^2 + y^2 \leqslant a^2 \tag{5.4.18}$$

中跑出的平均出口时间。

只要如 4.4.5 节那样作变元变换，就容易把这个问题简化为关于变量

$$r=\sqrt{x^2+y^2} \tag{5.4.19}$$

的一维问题，即化为

$$\mathrm{d}r=\left(-kr+\frac{1}{2}D/r\right)\mathrm{d}r+\sqrt{D}\mathrm{d}W(t) \tag{5.4.20}$$

并且要知道的是从$(0, a)$中跑出的平均出口时间。这可以在代换

$$\left.\begin{aligned} U(x)&\longrightarrow\frac{1}{2}kr^2-\frac{1}{2}D\log r\\ D&\longrightarrow\frac{1}{2}D\\ x_0&\longrightarrow a\\ a&\longrightarrow\sqrt{x_0^2+y_0^2}\equiv r_0\\ -\infty&\longrightarrow 0 \end{aligned}\right\} \tag{5.4.21}$$

之下使用(5.2.165)来解出。于是有

$$\begin{aligned} T(r_0\to a)=&\frac{1}{D}\int_{r_0}^{a}y^{-1}\exp[ky^2/D]\mathrm{d}y\\ &\cdot\int_0^{r}z\exp(-kz^2/D)\mathrm{d}z \end{aligned} \tag{5.4.22}$$

所以此问题实质上是一维的。这不是经常发生的情况。

b) 应用本征函数的解法

假定我们用本征函数 $Q_\lambda(x)$ 及 $P_\lambda(x)$ 把平均首次通过时间展开为

$$T(x)=\sum t_\lambda Q_\lambda(x) \tag{5.4.23}$$

假定 $P_\lambda(x)$ 与 $Q_\lambda(x)$ 满足我们所考察的特殊出口问题要求的边界条件，从而使得(5.4.23)的 $T(x)$ 满足适当的边界条件。

作展开

$$-1=\sum_\lambda I_\lambda Q_\lambda(x) \tag{5.4.24}$$

其中

$$I_\lambda=-\int_R \mathrm{d}x P_\lambda(x) \tag{5.4.25}$$

把(5.4.23)代入(5.4.8)就得到

$$\lambda t_1 = -I_\lambda \tag{5.4.26}$$

从而使得

$$T(x) = \sum_\lambda \frac{1}{\lambda} Q_\lambda(x) \int_R dx' P_\lambda(x') \tag{5.4.27}$$

这个方法成功与否依赖于对这些本征值的了解，它们满足S上的正确边界条件而且在R上是归一化的。

c) 渐近结果

如果第一个本征值 λ_1 远小于其他的本征值，则级数可以由第一项来近似。这意味着本征函数 $Q_{\lambda_1}^{(*)}$ 十分靠近

$$\sum_i A_i(x)\partial_i f(x) + \frac{1}{2}\sum_{i,j} B_{ij}(x)\partial_i\partial_j f(x) = 0 \tag{5.4.28}$$

的一个解，因为 λ_1 十分小。因此

$$Q_{\lambda_1}(x) \sim K \tag{5.4.29}$$

其中K是一个常数。考虑到 P_λ 与 Q_λ 的双正交性，我们有

$$1 = \int dx P_{\lambda_1}(x) Q_{\lambda_1}(x) \sim K \int dx P_{\lambda_1}(x) \tag{5.4.30}$$

从而

$$T(x) \sim 1/\lambda_1 \tag{5.4.31}$$

这里给出的推导较粗糙，它可用第 9 章的渐近方法加以改进。

d) 本征函数方法的应用

二维布朗运动：设粒子在 xy 平面内角端点分别为 (0, 0)、(0, 1)、(1, 0) 和 (1, 1) 的正方形中移动。这个正方形都是吸收壁。$T(x, y)$服从

$$\frac{D}{2}\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) = -1 \tag{5.4.32}$$

在正方形的各边上满足边界条件 $T = 0$ 的本征函数是

$$P_{n,m}(x, y) = \sin(n\pi x)\sin(m\pi y) \tag{5.4.33}$$

与

$$Q_{n,m}(x, y) = 4\sin(n\pi x)\sin(m\pi y) \tag{5.4.34}$$

(*) 原文在本节多处把Q_{λ_1}, P_{λ_1}误写为Q_1, P_1. ——译注

其中 n, m 是正整数. 本征值是

$$\lambda_{n,m}=\frac{\pi^2 D}{2}(n^2+m^2) \tag{5.4.35}$$

系数

$$\begin{aligned}\int \mathrm{d}x\mathrm{d}y P_{n,m}(x,\ y)&=0 \qquad (n\text{ 或 }m\text{ 为偶数})\\ &=\frac{4}{mn\pi^2} \quad (m\text{ 及 }n\text{ 都为奇数})\end{aligned} \tag{5.4.36}$$

所以

$$T(x,\ y)=\frac{1}{D}\sum_{\substack{n,m\\ \text{都为奇}}}\frac{32}{\pi^2 nm(n^2+m^2)}\sin(n\pi x)\sin(m\pi y) \tag{5.4.37}$$

5.4.2 出口点的分布

这是5.2.8节解决过的问题的多维类同问题. 也就是要求出粒子通过 R 的边界 S 上在 a 点的面积元 $\mathrm{d}S(a)$ 出去的出口概率. 我们假设 S 上全都是吸收的.

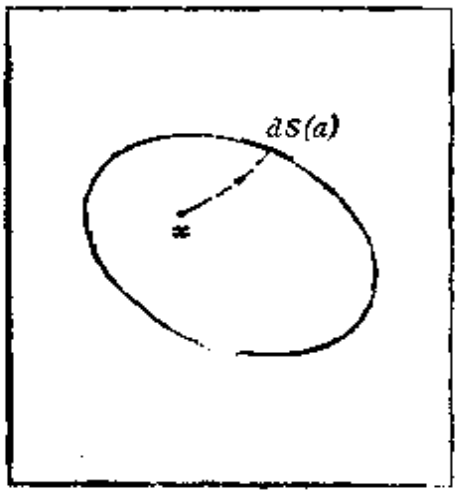


图 5.5 5.4.2节中考虑的曲面与区域

粒子经过时间 t 后才从 $\mathrm{d}S(a)$ 出去的概率是

$$\begin{aligned}&g(a,\ x,\ t)\,|\mathrm{d}S(a)|\\ &\quad=-\int_t^\infty \mathrm{d}t' J(a,\ t'|x,\ 0)\cdot \mathrm{d}S(a)\end{aligned} \tag{5.4.38}$$

类似于5.2.8节推理证明 $g(a,\ x,\ t)$ 服从后向福克-普朗克方程

$$\begin{aligned}&\sum_i A_i(x)\partial_i g(a,\ x,\ t)+\frac{1}{2}\sum_{i,j}B_{ij}(x)\partial_i\partial_j g(a,\ x,\ t)\\ &\qquad=\partial_t g(a,\ x,\ t)\end{aligned} \tag{5.4.39}$$

边界条件由定义得到. 在初始时刻有

$$g(a,\ x,\ 0)=0 \quad x\neq a \text{ 且 } x\in R \tag{5.4.40}$$

而且对任何时刻

$$g(\boldsymbol{a}, \boldsymbol{x}, t) = 0 \quad \boldsymbol{x} \neq \boldsymbol{a} \text{ 且 } \boldsymbol{x} \in S \tag{5.4.41}$$

若 $\boldsymbol{x} = \boldsymbol{a}$，则必然通过 $\mathrm{d}S(\boldsymbol{a})$ 出口，因而

$$g(\boldsymbol{a}, \boldsymbol{a}, t)\mathrm{d}S(\boldsymbol{a}) = 1 \quad \text{对所有 } t \tag{5.4.42}$$

或更有效地表达为

$$g(\boldsymbol{a}, \boldsymbol{x}, t) = \delta_s(\boldsymbol{a} - \boldsymbol{x}) \quad \boldsymbol{x} \in \boldsymbol{S} \text{ 对所有 } t \tag{5.4.43}$$

其中 $\delta_s(\boldsymbol{a} - \boldsymbol{x})$ 是一适当的曲面 δ 函数，使得

$$\int |\mathrm{d}\boldsymbol{S}(\boldsymbol{a})| \delta_s(\boldsymbol{a} - \boldsymbol{x}) = 1 \tag{5.4.44}$$

从 $\mathrm{d}\boldsymbol{S}(\boldsymbol{a})$ 中跑出的出口概率是

$$\pi(\boldsymbol{a}, \boldsymbol{x})|\mathrm{d}\boldsymbol{S}(\boldsymbol{a})| = g(\boldsymbol{a}, \boldsymbol{x}, \infty)|\mathrm{d}\boldsymbol{S}(\boldsymbol{a})| \tag{5.4.45}$$

已知在 $\boldsymbol{a}$ 点出口时的平均出口时间是

$$T(\boldsymbol{a}, \boldsymbol{x}) = \int_0^\infty \mathrm{d}t g(\boldsymbol{a}, \boldsymbol{x}, t) / \pi(\boldsymbol{a}, \boldsymbol{x}) \tag{5.4.46}$$

并且利用 5.2.8 节的相同方法可以证明它必满足

$$\boxed{\sum_i A_i(\boldsymbol{x})\partial_i[\pi(\boldsymbol{a}, \boldsymbol{x})T(\boldsymbol{a}, \boldsymbol{x})] + \frac{1}{2}\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\partial_j[\pi(\boldsymbol{a}, \boldsymbol{x})T(\boldsymbol{a}, \boldsymbol{x})] = -\pi(\boldsymbol{a}, \boldsymbol{x})} \tag{5.4.47}$$

而且边界条件是

$$\pi(\boldsymbol{a}, \boldsymbol{x})T(\boldsymbol{a}, \boldsymbol{x}) = 0, \ \boldsymbol{x} \in S \tag{5.4.48}$$

不仅如此，如果在 $g(\boldsymbol{a}, \boldsymbol{x}, t)$ 的相应福克-普朗克方程中还令 $t \to \infty$，则得到 $\pi(\boldsymbol{a}, \boldsymbol{x})$ 的方程：

$$\boxed{\sum A_i(\boldsymbol{x})\partial_i[\pi(\boldsymbol{a}, \boldsymbol{x})] + \frac{1}{2}\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\partial_j[\pi(\boldsymbol{a}, \boldsymbol{x})] = 0} \tag{5.4.49}$$

(5.4.49) 的边界条件是

$$\pi(\boldsymbol{a}, \boldsymbol{x}) = 0 \quad \boldsymbol{a} \neq \boldsymbol{x} \text{ 且 } \boldsymbol{x} \in \boldsymbol{S} \tag{5.4.50}$$

以及

$$\int |\mathrm{d}\boldsymbol{S}(\boldsymbol{a})| \pi(\boldsymbol{a}, \boldsymbol{x}) = 1 \tag{5.4.51}$$

于是可把它们汇总为

$$\pi(a, x) = \delta_s(a - x) \quad x \in S \tag{5.4.52}$$

其中 $\delta_s(a-x)$ 是边界 S 的曲面 δ 函数。

6. 扩散过程的近似方法

上两章中所刻划的方法是以精确结果为中心的，自然这些结果在适用性上是受到限制的。近似方法则是绝大多数应用的主体，人们总要找到某些方法把一个问题简化为一个确实可解的问题。甚至可以说，应用中的绝大多数工作就在于做各种近似方法的开发。

有两种具有非常重要意义的主要近似方法。第一种是小噪声展开理论，它给出了确定性方程的一些线性化解。由于噪声经常很小，这是个有广泛实际应用的方法。借此方法方程被简化为与时间相依的奥恩斯坦-乌伦贝克过程的序列。大多数是用一阶方程。

另一大类方法则是由绝热消去给出的。在这类方法中不同的时间尺度被当作同一的，而且快变变量被完全消除了，这构成了本章后一半的基本内容。

6.1 小噪声摄动理论

在许多物理与化学系统中，动力学系统中的随机成分产生于通常总是十分小的热起伏。除非作十分细致的测量，否则就难以检测出这种起伏的存在。在这种情况下此系统的按时间演化将是近乎确定性的，而且其起伏将是一种小振动。

鉴于这种想法，我们来考察一个完全可解出的简单线性问题的例子，由随机微分方程

$$dx = -kxdt + \varepsilon dW(t) \tag{6.1.1}$$

所刻划的单变量奥恩斯坦-乌伦贝克过程，其福克-普朗克方程是

$$\partial_t p = \partial_x(kxp) + \frac{1}{2}\varepsilon^2\partial_x^2 p \tag{6.1.2}$$

这些方程的解已在前面的 3.8.4 节和 4.4.4 节中研究过了。这里的 ε 是个小参量，它在确定性地取极限意义下是零。但是上述两种情况下的取极限 $\varepsilon\to 0$ 却是本质不同的。

在随机微分方程(SDE)(6.1.1)情况下，当 $\varepsilon\to 0$ 时微分方程变成非随机的而只剩下关于 t 的一阶项，而且因此取 $\varepsilon\to 0$ 的极限是非奇异的。与之相对照，在福克-普朗克方程(FPE)(6.1.2)的情况下取极限 $\varepsilon\to 0$，则把一个二阶微分方程简化为一个一阶微分方程。这个极限过程是奇异的，而且其任何摄动理论都是奇异摄动理论。

(6.1.1)的解是完全已为人所知的，就是(4.4.26)所给的

$$x_\varepsilon(t) = ce^{-kt} + \varepsilon\int_0^t e^{-k(t-t')}dW(t') \tag{6.1.3}$$

它可表示成

$$x_\varepsilon(t) = x_0(t) + \varepsilon x_1(t) \tag{6.1.4}$$

而且这就是一般解；就是说，我们通常可以把解表示成小参量 ε 的一个幂级数。进一步，其零阶项 $x_0(t)$ 就是令 $\varepsilon=0$ 而得到的方程

$$dx = -kxdt \tag{6.1.5}$$

之解。

但对于福克-普朗克方程 (6.1.2) 来说情况就决非这样简单了。假定初始值 c 是一非随机常数，则精确解就是均值和方差分别由

$$\langle x(t)\rangle = \alpha(t) = ce^{-kt} \tag{6.1.6}$$

$$\mathrm{var}\{X(t)\} = \varepsilon^2\beta(t) = \varepsilon^2(1-e^{-2kt})/2k \tag{6.1.7}$$

所给定的高斯分布，从而使得

$$p_\varepsilon(x,\ t\,|\,c,\ 0) = \frac{1}{\varepsilon}\sqrt{\frac{1}{2\pi\beta(t)}}\ \exp\left\{-\frac{1}{\varepsilon^2}\,\frac{[x-\alpha(t)]^2}{2\beta(t)}\right\} \tag{6.1.8}$$

当 $\varepsilon\to 0$ 时这个条件概率解具有形式

$$p_\varepsilon(x,\ t|c,\ 0)\to\delta[x-\alpha(t)] \tag{6.1.9}$$

它恰好对应着随机微分方程的一阶解，是一个沿着路径 $\alpha(t)=c\exp(-kt)$ 的确定性轨道。但是 p_ε 不能展开为 ε 的简单幂级数。为了实现幂级数展开必须在每一时刻定义一个标度变量

$$y=[x-\alpha(t)]/\varepsilon \tag{6.1.10}$$

使得 y 的概率密度是

$$\bar{p}_\varepsilon(y,\ t|0,\ 0)=p_\varepsilon(x,\ t|c,\ 0)\frac{\mathrm{d}x}{\mathrm{d}y} \tag{6.1.11}$$

$$=\frac{1}{\sqrt{2\pi\beta(t)}}\exp\left[-\frac{y^2}{2\beta(t)}\right] \tag{6.1.12}$$

标度变量 y 的概率密度没有奇异性了；事实上它已经和 ε 无关了。

变换(6.1.10)可以改写成

$$x=\alpha(t)+\varepsilon y \tag{6.1.13}$$

并且可以解释如下：由(6.1.12)知 y 的分布是均值为零而方差为 $\beta(t)$ 的高斯分布。而方程(6.1.13)则是说 x 相对于确定性路径 $\alpha(t)$ 的偏离当 $\varepsilon\to0$ 时是 ε 阶大小的，而且系数是这高斯分布变量 y。这本质上相同于用随机微分方程所得到的结论。

这些结果的一般化形式如下。设有一个用随机微分方程

$$\mathrm{d}x=a(x)\mathrm{d}t+\varepsilon b(x)\mathrm{d}W(t) \tag{6.1.14}$$

描写的系统，则可把解表示成

$$x_\varepsilon(t)=x_0(t)+\varepsilon x_1(t)+\varepsilon^2x_2(t)+\cdots \tag{6.1.15}$$

并且依次地解出 $x_n(t)$。特别地，$x_0(t)$ 是确定性方程

$$\mathrm{d}x=a(x)\mathrm{d}t \tag{6.1.16}$$

的解。另一方面，如果我们是考察福克-普朗克方程

$$\partial_t p=-\partial_x[a(x)p]+\frac{1}{2}\varepsilon^2\partial_x^2[b(x)^2p] \tag{6.1.17}$$

则只要把变量改变到标度变量并记

$$y=[x-x_0(t)]/\varepsilon \tag{6.1.18}$$

$$\bar{p}_\varepsilon(y,\ t)=[p(x,\ t|c,\ 0)]/\varepsilon \tag{6.1.19}$$

就可以写出摄动展开

$$\bar{p}_\varepsilon(y,\ t)=\bar{p}_0(y,\ t)+\varepsilon\bar{p}_1(y,\ t)+\varepsilon^2\bar{p}_2(y,\ t)+\cdots \tag{6.1.20}$$

这里我们将会看到 $\bar{p}_0(y, t)$ 确实是真正的概率密度，也就是说它是正值的且归一的，但高阶项在某些区域中却取负值。所以可以说福克-普朗克摄动理论不是概率论性质的。与此相对照，随机微分方程理论中展开成一些随机变量 $x_n(t)$ 的级数时，这每个变量都具有自己的概率分布。在每一阶段系统都是概率论性质的。

最后来看看它们最值得注意的差别。在随机微分方程摄动理论中的第一项是 $x_0(t)$，它是在(6.1.1) 中令 $\varepsilon = 0$ 而得的随机微分方程之解。与此相对照，(6.1.20)中的第一项 $\bar{p}_0(y, t)$并不是于(6.1.2)中令 $\varepsilon = 0$ 而得的方程的解。一般说来它是 $\bar{p}_\varepsilon(y, t)$的福克-普朗克方程的一个极限形式，是在标度变量 y 的福克-普朗克方程中令 $\varepsilon = 0$ 而得到的。

6.2 随机微分方程的小噪声展开

我们来考察一个形式为

$$\mathrm{d}x = a(x)\mathrm{d}t + \varepsilon b(x)\mathrm{d}W(t) \tag{6.2.1}$$

的随机微分方程，其中 ε 是一个小参量。眼下我们排除了 $a(x)$及 $b(x)$对时间的相依性，这只是出于简化符号及演绎过程的需要，即便不做简化，其结果和方法也完全相同。所以假设 (6.2.1) 的解 $x(t)$可表示为

$$x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \cdots \tag{6.2.2}$$

此外，还假设 a 可以表示为

$$a(x) = a(x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots) \tag{6.2.3}$$

$$= a_0(x_0) + \varepsilon a_1(x_0, x_1) + \varepsilon^2 a_2(x_0, x_1, x_2) + \cdots \tag{6.2.4}$$

(6.2.4)中的这种特殊的函数相依性是重要的，也不难说明，这是因为

$$a(x) = a\left(x_0 + \sum_{m=1}^{\infty} \varepsilon^m x_m\right)$$

$$= \sum_{p=0}^{\infty} \frac{1}{p!} \frac{\mathrm{d}^p a(x_0)}{\mathrm{d}x_0^p} \left(\sum \varepsilon^m x_m\right)^p \tag{6.2.5}$$

正规的论述并不容易，不过可以直接计算 ε 的前几阶幂次而得到

$$a_0(x_0) = a(x_0)$$

$$a_1(x_0, x_1) = x_1 \frac{da(x_0)}{dx_0}$$

$$a_2(x_0, x_1, x_2) = x_2 \frac{da(x_0)}{dx_0} + \frac{1}{2} x_1^2 \frac{d^2a(x_0)}{dx_0^2} \tag{6.2.6}$$

$$a_3(x_0, x_1, x_2, x_3) = x_3 \frac{da(x_0)}{dx_0} + x_1 x_2 \frac{d^2a(x_0)}{dx_0^2} + \frac{1}{6} x_1^3 \frac{d^3a(x_0)}{dx_0^3} \cdots$$

尽管一般地把所有各项都明显写出来是不容易的，但不难看出对 $n \geqslant 1$ 总可以写出

$$a_n(x_0, x_1, \cdots, x_n) = x_n \frac{da(x_0)}{dx_0} + A_n(x_0, x_1, \cdots, x_{n-1}) \tag{6.2.7}$$

其中的 A_n 还可看出是与 x_n 无关的。事实上，不难直接从(6.2.5)看出 ε^n 的系数若是由 $p=1$ 的项所起作用的结果，则它可能只含 x_n，而其他产生 ε^n 项的可能性只能来自于 $m<n$ 的项，但这些项都不含有 x_n。

a_n 的形式(6.2.7)非常重要。若还要求有

$$b(x) = b_0(x_0) + \varepsilon b_1(x_0, x_1) + \varepsilon^2 b_2(x_0, x_1, x_2,) + \cdots \tag{6.2.8}$$

则对它显然也可应用上述所有相同的结论。但是它们在摄动展开中不太重要。

现在我们把展开式 (6.2.2, 7, 8) 代入随机微分方程(6.2.1)中，并把相同 ε 幂次的系数相等起来，就得到由无穷多个随机微分方程构成的组。在这些方程中我们用记号

$$k(x_0) = -\frac{da(x_0)}{dx_0} \tag{6.2.9}$$

来使记号简化，则得到

$$dx_0 = a(x_0)dt \tag{6.2.10a}$$

$$dx_n = [-k(x_0)x_n + A_n(x_0, \cdots, x_{n-1})]dt + b_{n-1}(x_0, \cdots, x_{n-1})dW(t) \tag{6.2.10b}$$

这些方程可以依次地相继解出。方程(6.2.10a)是一(可能是非线性的)常微分方程，假设其满足初值条件的解是已知的。当然有可能对所有的 x_n 去假设独立的而非平凡的初始条件，但这是不必要

的复杂化。最简单的做法是(在置 $t=0$ 为初始时刻后)置

$$x_0(0)=x(0)$$
$$x_n(0)=0 \quad n\geqslant 1 \tag{6.2.11}$$

若假设(6.2.9)的解是由

$$x_0(t)=\alpha(t) \tag{6.2.12}$$

给出的,则对 $n=1$ 方程(6.2.10b)(*)可表示成

$$\boxed{\mathrm{d}x_1=-k[\alpha(t)]x_1\mathrm{d}t+b[\alpha(t)]\mathrm{d}W(t)} \tag{6.2.13}$$

其中我们利用了由(6.2.5)而得知的 A_0 为零以及 $b_0=b$。

这个方程是摄动理论的第一步，它是一个与时间相依的奥恩斯坦-乌伦贝克过程,其解可直接了当地用第4.4.9节方法化为一维而得到。这个解其实只要用积分因子

$$\exp\left\{\int_0^t \mathrm{d}t' k[\alpha(t')]\right\}$$

去乘就可以得到,并且就是

$$x_1(t)=\int_0^t b[\alpha(t')]\exp\left\{-\int_{t'}^t k[\alpha(s)]\mathrm{d}s\right\}\mathrm{d}W(t') \tag{6.2.14}$$

这里初始条件 $x_1(0)=0$ 已包含在式中。

对于许多目的来说这个公式都是十分合用的，而且它就是相当于原来方程关于其确定性解的一个线性化。由于(6.2.10b)具有更复杂的形式,故高阶项更为复杂,但实质上它们也是用相同方法来处理的。为了解出 $x_N(t)$ 的方程我们假设所有的 $x_n(t)$, $n<N$, 都已知,从而把这些解代入后就使得 A_n 及 A_{n-1} 成了 t 的已知(随机)函数。于是(6.2.10b)成了

$$\mathrm{d}x_N=\{-k[\alpha(t)]x_N+A_N(t)\}\mathrm{d}t+b_{N-1}(t)\mathrm{d}W(t) \tag{6.2.15}$$

其解可直接得到或由4.4.9节结果而得到,这便是

$$x_N(t)=\int_0^t [A_N(t')\mathrm{d}t'+b_{N-1}(t')\mathrm{d}W(t')]\cdot\exp\left\{-\int_{t'}^t k[\alpha(s)]\mathrm{d}s\right\} \tag{6.2.16}$$

(*) 原文误为(6.2.10a)。——译注

形式上程序到此已完成了。但此方法的适用范围及其可实行性尚无答案。如同所有的摄动理论那样，随着阶数的增加它的项会迅速变得不可控制。

6.2.1 展开的有效性

这个展式通常不是一个收敛幂级数，由于(6.2.14)说明了$x_1(t)$是个高斯变量，而且是个具有非随机系数的伊藤积分，因而它大于任何给定值的概率都可以是有限值，于是只有包含在(6.2.5)中的各项导数的所有幂级数对所有的变量，不管其多大，都是收敛时，我们才能期望此方法会产生出一个收敛的展式。

事实上，利用4.3.7节中所给的关于依赖于一个参数的那些结果可以证明此展式是个渐近展开。我们用

$$y_n(\varepsilon,\ t)=[x(t)-\sum_{r=0}^{n}\varepsilon^r x_r(t)]/\varepsilon^{n+1} \tag{6.2.17}$$

来定义一个余项，其中$x_r(t)$是具有初始条件(6.2.11)的随机微分方程组(6.2.10)的解。

下面来导出$y_n(t)$满足的方程。我们可写出

$$a[x(t)]=a\Big[\sum_{r=0}^{n}\varepsilon^r x_r(t)+y_n(\varepsilon,\ t)\varepsilon^{n+1}\Big] \tag{6.2.18}$$

并且用

$$\begin{aligned}\hat{a}_{n+1}[X_0,\ X_1,\ \cdots,\ X_n,\ Y,\ \varepsilon]=\varepsilon^{-n-1}\Big\{a\Big[\sum_{r=0}^{n}\varepsilon^r X_r+\varepsilon^{n+1}Y\Big]\\ -\sum_{r=0}^{n}\varepsilon^r a_r(X_0,\ X_1,\ \cdots,\ X_r)\Big\}\end{aligned} \tag{6.2.19}$$

来定义函数$\hat{a}_{n+1}[X_0,\ X_1,\ \cdots,\ X_n,\ Y,\ \varepsilon]$。我们要求对所有给定的$X_0,\ X_1,\ \cdots,\ X_n,\ Y$存在极限

$$\lim_{\varepsilon\to 0}\hat{a}_{n+1}[X_0,\ X_1,\ \cdots,\ X_n,\ Y,\ \varepsilon] \tag{6.2.20}$$

还可类似地定义$\hat{b}_n[X_0,\ X_1,\ \cdots,\ X_n,\ Y,\ \varepsilon]$并赋予它相同的条件。

这个条件不是概率性的，而是表示出函数$a(x)$及$b(x)$所要求

有的解析性质；事实上它要求展开(6.2.4,8)仅仅是渐近展开。

现在可把 $y_n(\varepsilon, t)$ 的微分方程表示为

$$dy_n = \hat{a}_{n+1}[x_0(t), x_1(t), \cdots, x_n(t), y_n, \varepsilon]dt + \hat{b}_n[x_0(t), x_1(t), \cdots, x_n(t), y_n, \varepsilon]dW(t) \quad (6.2.21)$$

因为 $x_r(t)$ 是随机的，dt 及 $dW(t)$ 的系数是随机函数。但是(6.2.20)的要求现在就成了一个几乎必然极限，而且因此蕴含着随机极限

$$\text{st-}\lim_{\varepsilon\to 0} \hat{a}_{n+1}[x_0(t), x_1(t), \cdots, x_n(t), y_n, \varepsilon] = \tilde{a}_{n+1}(t, y_n) \quad (6.2.22)$$

以及

$$\text{st-}\lim_{\varepsilon\to 0} \hat{b}_{n+1}[x_0(t), x_1(t), \cdots, x_n(t), y_n, \varepsilon] = \tilde{b}_n(t, y_n) \quad (6.2.23)$$

的存在。倘若 4.3.7 节的李普希兹条件(ii)及(iii)被满足了，则上述极限的存在就足以使 4.3.7 节中关于随机微分方程(6.2.21)的解对参数 ε 的连续性的结果得到满足。这样一来，$y_n(0, t)$ 就作为随机微分方程

$$dy_n(0, t) = \tilde{a}_{n+1}[t, y_n(0, t)]dt + \tilde{b}_n[t, y_n(0, t)]dW(t) \quad (6.2.24)$$

的一个解是存在的，而由定义(6.2.17)它已说明了

$$x(t) - \sum_{r=0}^{n} \varepsilon^r x_r(t) \sim \varepsilon^{n+1} \quad (6.2.25)$$

因此这个 ε 的幂级数是一个渐近展开。

6.2.2 (齐次过程的)平稳解

平稳解是令 $t\to\infty$ 而得到的。如果过程是已写出的那样为时齐且遍历的，则其初始条件不论是怎样的都没有关系。此时可选 $x_0(0)$ 使得 $a[x_0(0)]$ 为零并且(6.2.10a)的解是

$$x_0(t) = x_0(0) = \alpha \quad (6.2.26)$$

[这里我们简单地用 α 代替 $\alpha(t)$]。

由于在 $t=0$ 处的初始条件，一阶方程(6.1.14)不是个平稳过程。我们必须令 $t\to-\infty$ 或者不在 $t=0$ 处而在 $t=-\infty$ 处设置初始条件。若取后一种办法，则有

$$x_1^s(t)=\int_{-\infty}^{t} b(\alpha)\exp[-(t-t')k(\alpha)]\mathrm{d}W(t') \tag{6.2.27}$$

类似地，可得

$$x_n^s(t)=\int_{-\infty}^{t}[A_n^s(t')\mathrm{d}t'+b_{n-1}^s(t')\mathrm{d}W(t')]\cdot\exp[-(t-t')k(\alpha)] \tag{6.2.28}$$

其中 A_n^s 及 $b^s{}_{n-1}$ 分别表示 A_n 及 b_{n-1} 在各变元用其稳态值代入后得到的值。由(6.2.28)，按 $x_n^s(t)$ 的构造就显然知道它是平稳的。很明显，只有 $k(\alpha)>0$ 时(6.2.27, 28)中的积分才是收敛的，这将意味着确定性过程只有一个稳定的平稳解才可用此方法来产生出平稳解。这是比较明显的事情——因为对一个非稳定状态加上起伏将导致系统偏离其所处状态。

6.2.3 均值、方差与时间相关函数

如果在某种意义下按 ε 展开成级数是有效的话，则对均值与方差来说知道这些级数是很有用的。显然

$$\langle x(t)\rangle=\sum_{n=0}^{\infty}\varepsilon^n\langle x_n(t)\rangle \tag{6.2.29}$$

$$\mathrm{var}\{x(t)\}=\sum_{n=0}^{\infty}\varepsilon^n\sum_{m=0}^{\infty}[\langle x_m(t)x_{n-m}(t)\rangle-\langle x_m(t)\rangle\langle x_{n-m}(t)\rangle] \tag{6.2.30}$$

但是由于假设了确定性的初值条件使 $x_0(t)$ 是确定性的，故所有含 $x_0(t)$ 的项都为零。于是可得到

$$\mathrm{var}\{x(t)\}=\varepsilon^2\mathrm{var}\{x_1(t)\}+2\varepsilon^3\langle x_1(t),x_2(t)\rangle+\varepsilon^4[2\langle x_1(t),x_3(t)\rangle+\mathrm{var}\{x_2(t)\}]+\cdots \tag{6.2.31}$$

并且类似地有

$$\langle x(t),x(s)\rangle=\varepsilon^2\langle x_1(t),x_1(s)\rangle+\varepsilon^3[\langle x_1(t),x_2(s)\rangle$$

$$+\langle x_1(s),\ x_2(t)\rangle] + \varepsilon^4[\langle x_1(t),\ x_3(s)\rangle + \langle x_1(s),\ x_3(t)\rangle$$
$$+\langle x_2(t),\ x_2(s)\rangle] + \cdots \qquad (6.2.32)$$

6.2.4 小噪声摄动理论的失效

a) 例子：立方过程及有关行为

考虑随机微分方程

$$\mathrm{d}x = -x^3\mathrm{d}t + \varepsilon \mathrm{d}W(t) \qquad (6.2.33)$$

不难看出展开条件 (6.2.20) 对 $\mathrm{d}t$ 及 $\mathrm{d}W(t)$的系数来说都是轻而易举地被满足，事实上，对任何有限时刻 t 来说，以由(6.2.16)所给的那些 $x_N(t)$为各项的渐近展开是有效的。但是在 $x=0$ 时显然有

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^3)\Big|_{x=0} = h(0) = 0 \qquad (6.2.34)$$

并且因为 $x=0$ 是确定性方程的平稳解，故平稳解的摄动级数因展开式中的指数化时间因子全为常数而多半是不收敛的。例如，在平稳展式中的一阶项按(6.2.27)是

$$x_1^s(t) = \int_{-\infty}^{t} \mathrm{d}W(t') = W(t) - W(-\infty) \qquad (6.2.35)$$

而它是以概率1取无穷值(是具有无穷大方差的高斯分布变量)。

问题之所在是比较明显的。在 $x=0$ 附近，由(6.2.33)所刻划的过程完全不能由任何奥恩斯坦-乌伦贝克过程来近似。例如，其平稳分布是福克-普朗克方程

$$\partial_t p = \partial_x(x^3 p) + \frac{1}{2}\varepsilon^2\partial_x^2 p \qquad (6.2.36)$$

的平稳解，是由

$$p_s(x) = \mathcal{N}\exp(-x^4/2\varepsilon^2) \qquad (6.2.37)$$

所给定的，而且其各阶矩是

$$\langle x^n\rangle = (2\varepsilon)^{n/4}\Gamma\left(\frac{n+1}{4}\right)\Big/\Gamma\left(\frac{1}{4}\right) \quad (n\text{ 为偶时})$$
$$= 0 \quad (n\text{ 为奇时}) \qquad (6.2.38)$$

这种方差的展开式中最低阶项是与 ε 的一次幂成比例的，而不是

像(6.2.31)中那样与 ε^2 成比例。

在此情况下我们应明确地把(6.2.33)所描述的立方过程看作为一种基本过程。如果我们通过定义

$$x=\sqrt{\varepsilon}\,y$$
$$t=\frac{\tau}{\varepsilon} \tag{6.2.39}$$

来引进新的标度变量并利用

$$\mathrm{d}W(\tau/\varepsilon)=\mathrm{d}W(\tau)/\sqrt{\varepsilon} \tag{6.2.40}$$

那末此立方过程就可以表现为无参量的形式

$$\mathrm{d}y=-y^3\mathrm{d}\tau+\mathrm{d}W(\tau) \tag{6.2.41}$$

把(6.2.41)的解看作为一已知量，则可写出

$$x(t)=\sqrt{\varepsilon}\,y(\varepsilon t) \tag{6.2.42}$$

从而使$\varepsilon\to 0$的极限过程同$\sqrt{\varepsilon}$趋于零的过程一样，而且也是以一较慢的时间尺度来进行的。这类标度结果是许多临界现象的基础。

对于在 $x=0$ 附近 $a(x)$ 的行为像 x^3 的情况来说，一个成功的摄动理论首先应该包含着变量代换(6.2.39)，然后是一类摄动理论，它们与已作略述的相似，但是其中的零阶解为立方过程。所以我们假设可写出

$$a(x)=-x^3c(x) \tag{6.2.43}$$

其中 $c(x)$ 是一光滑函数，$c(0)\neq 0$。于是利用代换(6.2.39, 40)可改写随机微分方程为

$$\mathrm{d}y=-y^3c\,(y\sqrt{\varepsilon})\mathrm{d}\tau+b(y\sqrt{\varepsilon})\mathrm{d}W(\tau) \tag{6.2.44}$$

若展开 $y(t)$、$c(y\sqrt{\varepsilon})$ 及 $b(y\sqrt{\varepsilon})$ 为 $\sqrt{\varepsilon}$ 的级数，就得到了一个摄动理论。只要把 y 表示为

$$y(t)=\sum_{n=0}^{\infty}\varepsilon^{n/2}y_n(t) \tag{6.2.45}$$

则作为首二项我们有

$$\mathrm{d}y_0=-y_0^3c(0)\mathrm{d}\tau+b(0)\mathrm{d}W(\tau) \tag{6.2.46}$$

$$\mathrm{d}y_1=-y_1\left[3y_0^2c(0)+y_0^3\frac{\mathrm{d}c}{\mathrm{d}x}(0)\right]d\tau+\left[\frac{\mathrm{d}b}{\mathrm{d}x}(0)y_0\right]\mathrm{d}W(\tau) \tag{6.2.47}$$

我们看到 y_1 的方程实际上是具有随机系数的时间相依奥恩斯坦-乌伦贝克过程。这样，从原则上讲，只要立方过程已知，其余都容易计算。但实际中关于立方过程事实上并不知道很多。这类摄动不是非常切实可行的。

b) ν 阶过程(ν 为奇数)

若代以考虑随机微分方程

$$\mathrm{d}x = -x^{\nu}\mathrm{d}t + \varepsilon \mathrm{d}W(t) \tag{6.2.48}$$

则只有当 ν 是奇数时才会找到一个稳定的确定性状态。在此情况下我们就作变量变换变到标度变量

$$x = y\varepsilon^{2/(1+\nu)} \tag{6.2.49}$$

$$t = \tau\varepsilon^{2(1-\nu)/(1+\nu)} \tag{6.2.50}$$

然后采用十分类似于立方过程用过的步骤来处理。

c) 双稳态系统

假定有个由

$$\mathrm{d}x = (x - x^3)\mathrm{d}t + \varepsilon \mathrm{d}W(t) \tag{6.2.51}$$

来描述的系统，它有三个确定性的平稳态，分别位于 $x=0, \pm 1$。在 $x=0$ 时状态是确定性不稳定的，并且由摄动理论可直接看出：由于摄动级数积分(6.2.27, 28) 中的指数因子有着不断增加的变元，所以由它不能产生平稳解。

确定性微分方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x - x^3 \tag{6.2.52}$$

之解 $x_0(t)$按其 $t\to\infty$ 时的行为分成三类，即

i) $x_0(t) < 0 \Rightarrow x_0(t) \to -1$

ii) $x_0(0) = 0 \Rightarrow$对所有 t, $x_0(t) = 0$

iii) $x_0(t) > 0 \Rightarrow x_0(t) \to 1$,

这样一来，按照初始条件的不同我们可得到两种不同的渐近展开，它们的平稳极限表示着关于两种不同的确定性平稳态的涨落。在这些解中，至少没有以任何明显的方式存在着有关从 $x=1$ 的一支跃迁到 $x=-1$ 一支的或相反方向跃迁的信息。因为它没有给出

平稳态的从头到尾行为的完整图景，在此意义上说，这个渐近展开是失效的。在第9章中我们将会看到这个结果是由于对以从一支到另一支的跃迁为特征的行为作渐近展开时，是以 $\exp\left(-\frac{1}{\varepsilon^2}\right)$ 为其典型的大小量级的，而当 $\varepsilon\to 0$ 时它以比 ε 的任何阶次都要快的速度收敛于0，因而无法表示成 ε 的幂级数展开式。

6.3 福克-普朗克方程的小噪声展开

正如6.1节中提到过的，福克-普朗克方程的小噪声展开是一个奇异的展开，它牵涉到引入标度变量。现在让我们来考察一下这是如何实现的。

我们来考虑福克-普朗克方程

$$\partial_t p = -\partial_x[A(x)p] + \frac{1}{2}\varepsilon^2\partial_x^2[B(x)p] \tag{6.3.1}$$

假设其确定性方程的解是 $\alpha(t)$ 而使得

$$d_t\alpha(t) = A[\alpha(t)] \tag{6.3.2}$$

再引入下列新变量 (y, s)

$$y = [x-\alpha(t)]/\varepsilon \tag{6.3.3}$$

$$s = t \tag{6.3.4}$$

以及 $\hat{p}(y, s) = \varepsilon p(x, t)$ (6.3.5)

注意到

$$\partial_t\hat{p}(y, s) = \frac{\partial\hat{p}}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial\hat{p}}{\partial s}\frac{\partial s}{\partial t} = -\frac{-\dot{\alpha}(t)}{\varepsilon}\frac{\partial\hat{p}}{\partial y} + \frac{\partial\hat{p}}{\partial s} \tag{6.3.6}$$

$$\partial_x\hat{p}(y, s) = \frac{1}{\varepsilon}\frac{\partial\hat{p}}{\partial y} \tag{6.3.7}$$

把它们代入(6.3.1)后，利用 $\alpha(t)$ 的运动方程(6.3.2)，就得到

$$\frac{\partial\hat{p}}{\partial s} = -\frac{\partial}{\partial y}\left\{\frac{A[\alpha(s)+\varepsilon y]-A[\alpha(s)]}{\varepsilon}\hat{p}\right\} + \frac{1}{2}\frac{\partial^2}{\partial y^2}\{B[\alpha(s)+\varepsilon y]\hat{p}\} \tag{6.3.8}$$

现在我们面临的是要构造 $\hat{p}$ 的 ε 幂级数展开。假设 A 及 B 具有形

式

$$A[\alpha(s)+\varepsilon y]=\sum_{n=0}^{\infty}\tilde{A}_n(s)\varepsilon^n y^n \tag{6.3.9}$$

$$B[\alpha(s)+\varepsilon y]=\sum_{n=0}^{\infty}\tilde{B}_n(s)\varepsilon^n y^n \tag{6.3.10}$$

并且 $\hat{p}$ 已展开成 ε 的幂级数：

$$\hat{p}=\sum_{n=0}^{\infty}\hat{p}_n\varepsilon^n \tag{6.3.11}$$

把这些展开式代入福克-普朗克方程(6.3.8)，令两边的相应系数相等，就得到

$$\frac{\partial\hat{p}_0}{\partial s}=-\tilde{A}_1(s)\frac{\partial}{\partial y}(y\hat{p}_0)+\frac{1}{2}B_0(s)\frac{\partial^2\hat{p}_0}{\partial y^2} \tag{6.3.12}$$

$$\frac{\partial\hat{p}_1}{\partial s}=-\frac{\partial}{\partial y}[\tilde{A}_1(s)y\hat{p}_1+\tilde{A}_2(s)y^2\hat{p}_0]+\frac{1}{2}\frac{\partial^2}{\partial y^2}[\tilde{B}_0(s)\hat{p}_1+\tilde{B}_1(s)y\hat{p}_0] \tag{6.3.13}$$

以及一般地

$$\frac{\partial\hat{p}_r}{\partial s}=-\frac{\partial}{\partial y}\left[\sum_{m=0}^{r}y^{r-m+1}\tilde{A}_{r-m+1}(s)\hat{p}_m\right]+\frac{1}{2}\frac{\partial^2}{\partial y^2}\left[\sum_{m=0}^{r}y^{r-m}\tilde{B}_{r-m}(s)\hat{p}_m\right] \tag{6.3.14}$$

只有 $\hat{p}_0$ 的方程是个福克-普朗克方程，而且如6.1节已提到过的，只有 $\hat{p}_0$ 是个概率分布。在这个有层次性系统中第一个方程(6.3.12)是与时间相依的奥恩斯坦-乌伦贝克过程，它恰好对应着随机微分方程的有层次性系统中的第一个方程(6.2.13)所表示的过程。但此后就不再有这种对应关系。

关于 $\hat{p}_r$ 的边界条件却存在着一些技术性困难。这是因为 x 到 y 的变换是与时间相依的，而且原来关于 x 变量的固定了位置的边界却对应着关于 y 变量的变动的边界。进一步，在 $x=a$ 处的边界对应于在

$$y=\frac{[a-\alpha(s)]}{\varepsilon} \tag{6.3.15}$$

处的边界，而当 $\varepsilon\to 0$ 时它就趋于 $\pm\infty$。除了当 $a=\pm\infty$ 而使得 y

边界也是在 $\pm\infty$ 处并因而是一常值的情况外，似乎还不存在任何已知的技巧来处理这类边界。所以我们假设变量 y 的边界条件与变量 ω 的具有相同形式。

如果情况是边界在无穷远处，则变换(6.3.3)的结果是把奇异的摄动问题(6.3.1)(在此问题中 $\varepsilon\to 0$ 的极限产生出一个低阶的方程)变到一个正常的摄动问题(6.3.8)，此方程的系数关于 ε 是光滑的，并且 $\varepsilon\to 0$ 所得到的极限是个 2 阶方程。展开方法的有效性将取决于系数的形式。

6.3.1 矩方程及自相关函数

有层次性的(6.3.14)不很容易处理，但它产生了用摄动计算矩的比较直捷了当的方法。假设边界在无穷远处而使得可作分部积分并弃去曲面项，则可定义

$$\langle[y(t)]^n\rangle = \sum_{r=0}^{\infty}\varepsilon^r M_r^n(t) \tag{6.3.16}$$

于是显然有

$$M_r^n(t) = \int \mathrm{d}y y^n \hat{p}_r(y,\ t) \tag{6.3.17}$$

由(6.3.12~14)，易借分部积分得知

$$\frac{\mathrm{d}M_r^n(t)}{\mathrm{d}t} = \sum_{m=0}^{r}\left[n\tilde{A}_{r-m+1}(t)M_m^{n+r-m}(t) + \frac{n(n-1)}{2}\tilde{B}_{r-m}(t)M_m^{n+r-m-2}(t)\right] \tag{6.3.18}$$

若 $M_m^p(t)$ 对所有的 $m<r$ 及 $p<n+r$ 都已知则 $M_r^n(t)$ 就可解出，所以这个有层次性的方程体系是封闭的。

只要写出均值 $M_r^1(t)$ 的以及二次均值 $M_r^2(t)$ 的开头很少几个方程，就会发现有

$$\frac{\mathrm{d}M_0^1(t)}{\mathrm{d}t} = \tilde{A}_1(t)M_0^1(t) \tag{6.3.19}$$

$$\frac{\mathrm{d}M_1^1(t)}{\mathrm{d}t} = \tilde{A}_1(t)M_1^1(t) + \tilde{A}_2(t)M_0^2(t) \tag{6.3.20}$$

$$\frac{dM_2^1(t)}{dt} = \tilde{A}_1(t)M_2^1(t) + \tilde{A}_2(t)M_1^2(t) + \tilde{A}_3(t)M_0^3(t) \tag{6.3.21}$$

$$\frac{dM_0^2(t)}{dt} = 2\tilde{A}_1(t)M_0^2(t) + \tilde{B}_0(t) \tag{6.3.22}$$

$$\frac{dM_1^1(t)}{dt} = 2\tilde{A}_1(t)M_1^2(t) + 2\tilde{A}_2(t)M_0^3(t) + \tilde{B}_1(t)M_0^1(t) \tag{6.3.23}$$

$$\frac{dM_0^3(t)}{dt} = 3\tilde{A}_1(t)M_0^3(t) + 3\tilde{B}_0(t)M_0^1(t) \tag{6.3.24}$$

在导出最后二个方程的过程中我们考虑到了

$$M_r^0(t) = \int dy \hat{p}_r(y, t) \tag{6.3.25}$$

并利用了

$$\int dy \hat{p}(y, t) = 1 = \sum_r e^r M_r^0(t) \tag{6.3.26}$$

从而可看出

$$M_0^0(t) = 1 \tag{6.3.27}$$

$$M_r^0(t) = 0, \quad r \neq 1 \tag{6.3.28}$$

这些方程是非齐次线性常微分方程，它们可以在有层次性的体系中由较低阶的方程开始计算。

a) 平稳矩

这些矩是令 $t \to \infty$ 并置(6.3.18)式的左边为零而得到的。(所有系数 $\tilde{A}$, $\tilde{B}$ 等等都被取为与时间无关了。)

由(6.3.19～22)我们发现

$$\begin{aligned}
&M_0^1(\infty) = 0 \\
&M_0^2(\infty) = -\frac{1}{2}\tilde{B}_0/\tilde{A}_1 \\
&M_0^3(\infty) = 0 \\
&M_1^1(\infty) = -\tilde{A}_2 M_0^2(\infty)/\tilde{A}_1 = \frac{1}{2}\tilde{A}_2\tilde{B}_0/(\tilde{A}_1)^2 \\
&M_1^2(\infty) = -\tilde{A}_2 M_0^3(\infty)/\tilde{A}_1 = 0 \\
&M_2^1(\infty) = -[\tilde{A}_2 M_1^2(\infty) + \tilde{A}_3 M_0^3(\infty)]/\tilde{A}_1 = 0
\end{aligned} \tag{6.3.29}$$

这样一来，平稳均值、平稳方差等等就是

$$\langle x\rangle_s = \alpha + \varepsilon[M_0^1(\infty) + \varepsilon M_1^1(\infty) + \varepsilon^2 M_2^1(\infty)]$$
$$= \alpha + \frac{1}{2}\varepsilon^2 \tilde{A}_2 \tilde{B}_0/(\tilde{A}_1)^2 \qquad (6.3.30)$$

$$\mathrm{var}\langle x\rangle_s = \langle x^2\rangle_s - \langle x\rangle_s^2$$
$$= \langle(\alpha+\varepsilon y)^2\rangle_s - \langle \alpha+\varepsilon y\rangle_s^2 = \varepsilon^2 \mathrm{var}\langle y\rangle_s$$
$$= -\frac{\varepsilon^2}{2}\tilde{B}_0/\tilde{A}_1 \quad (\text{到 } \varepsilon^2 \text{ 阶}) \qquad (6.3.31)$$

这种步骤显然可继续下去到任何高阶。当然，在单变量系统中平稳分布可以精确地算出大小，并且其各阶矩可以由积分得到。但是在多变量系统中这就不是总可以做到的。尽管如此，上述方法的多变量推广却总是可以实现的。

b) 平稳自相关函数

在平稳状态下 x 的自相关函数与 y 的自相关函数简单地由

$$\langle x(t)x(0)\rangle_s = \alpha^2 + \varepsilon^2\langle y(t)y(0)\rangle_s \qquad (6.3.32)$$

相关联着，并且容易发现 $\langle y(t)y(0)\rangle_s$ 的有层次性的方程体系。这只要注意

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle y(t)^n y(0)\rangle_s = \left\langle \left\{ \frac{A[\alpha+\varepsilon y(t)] - A(\alpha)}{\varepsilon} n y(t)^{n-1} + \frac{1}{2}n(n-1)B[\alpha+\varepsilon y(t)]y(t)^{n-2} \right\} y(0) \right\rangle_s \qquad (6.3.33)$$

它可由 $p(y, t|y_0, t_0)$ 的福克-普朗克方程 (6.3.1) 和分部积分导出，或者由对应的随机微分方程用伊藤公式导出。

利用 $\tilde{A}_r$, $\tilde{B}_r$ 的定义(6.3.9, 10)并把 A、B 展开成级数，我们可得

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle y(t)^n y(0)\rangle_s = \sum_{q=0}^{\infty}\left[\varepsilon^q[n\tilde{A}_{q+1}\langle y(t)^{q+n}y(0)\rangle_s + \frac{n(n-1)}{2}\tilde{B}_q\langle y(t)^{q+n-2}y(0)\rangle_s\right] \qquad (6.3.34)$$

这些方程本身也构成了一个新层次性体系，它们可以简明地解出

为 ε 的幂级数. 通常我们最感兴趣的是$\langle y(t)y(0)\rangle_s$，只要知道了所有的$\langle y(t)^n y(0)\rangle_s$，$p\leqslant q+1$，它就可以计算到 ε^q 阶. 我们有初始条件

$$\langle y(0)^n y(0)\rangle_s = \langle y^{n+1}\rangle_s \tag{6.3.35}$$

而且平稳矩可由平稳分布算出，或者利用前面讲过的办法来得到.

6.3.2 例子

考虑福克-普朗克方程

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}[(x+x^3)p] + \frac{\varepsilon^2}{6}\frac{\partial^2 p}{\partial x^2} \tag{6.3.36}$$

此时我们有[在平稳态时 $\alpha(t)=0$]

$$\begin{aligned} &A_1 = -1 \\ &A_2 = 0 \\ &A_3 = -1 \\ &A_q = 0 \quad (q>3) \\ &B_q = \frac{1}{3}\delta_{q,0} \\ &\alpha = 0 \end{aligned} \tag{6.3.37}$$

利用(6.3.30, 31)则有

$$\begin{aligned} &\langle x\rangle_s = 0 \\ &\mathrm{var}\{x\}_s = \varepsilon^2/6 \end{aligned} \tag{6.3.38}$$

为方便起见我们使用记号

$$c_n(t) = \langle y^n(t)y(0)\rangle_s \tag{6.3.39}$$

而使得 c_1 及 c_3 的方程为

$$\begin{bmatrix} \dfrac{dc_1}{dt} \\ \dfrac{dc_3}{dt} \end{bmatrix} = \begin{bmatrix} -1 & -\varepsilon^2 \\ 1 & -3 \end{bmatrix}\begin{bmatrix} c_1 \\ c_3 \end{bmatrix} \tag{6.3.40}$$

[由于 $B(x)$是常值而 $A(x)$是 x 的奇函数，故 c_{2n} 的方程可由 c_{2n+1} 的方程退耦].

精确地解出(6.3.40)比用摄动求解简单些。矩阵的本征值是

$$\lambda_1 = -2 + \sqrt{1-e^2}$$
$$\lambda_2 = -2 - \sqrt{1-e^2} \tag{6.3.41}$$

其相应的本征向量是

$$u_1 = \begin{bmatrix} 1+\sqrt{1-e^2} \\ 1 \end{bmatrix}$$
$$u_2 = \begin{bmatrix} 1-\sqrt{1-e^2} \\ 1 \end{bmatrix} \tag{6.3.42}$$

于是(6.3.40)之解可表示为

$$\begin{bmatrix} c_1(t) \\ c_2(t) \end{bmatrix} = \alpha_1 e^{-\lambda_1 t} u_1 + \alpha_2 e^{-\lambda_2 t} u_2 \quad (t>0) \tag{6.3.43}$$

初始条件是

$$c_1(0) = \langle y^2 \rangle_s$$
$$c_2(0) = \langle y^4 \rangle_s \tag{6.3.44}$$

利用延伸到 M_0^4 的有层次性矩方程体系 (6.3.18)，就可计算出 $\langle y^4 \rangle_s$。我们发现

$$M_0^4 = -\frac{3}{2}\frac{\tilde{B}_0 M_0^2}{\tilde{A}_1} = \frac{1}{12} \tag{6.3.45}$$

于是 $\langle y^4 \rangle = \frac{1}{12}$。所以

$$\begin{bmatrix} \frac{1}{6} \\ \frac{1}{12} \end{bmatrix} = \alpha_1 u_1 + \alpha_2 u_2 \tag{6.3.46}$$

其解是

$$\alpha_1 = \frac{1}{24}(1+\sqrt{1-e^2})/\sqrt{1-e^2}$$
$$\alpha_2 = \frac{1}{24}(-1+\sqrt{1-e^2})/\sqrt{1-e^2} \tag{6.3.47}$$

到 e 的二阶为止(消去许多项后的)相关函数是

$$c_1(t) = \frac{1}{6}e^{-|\lambda_1|t} \tag{6.3.48}$$

要注意本征值 λ_1 与 λ_2 是依赖于 ε^2 的，任何用摄动方法解出系统(6.3.40)的努力都将涉及把 $\exp(\lambda_1 t)$ 及 $\exp(\lambda_2 t)$ 展开成 ε^2 的幂级数，并且将引起诸如 $t^N\exp(-2t)$ 这种项的出现，但它们不应是自相关函数长时期行为的一种精确表现。

x 的相关函数是

$$\langle x(t)x(0)\rangle_s = \varepsilon^2 c_1(t) \tag{6.3.49}$$

且谱是

$$S(\omega) = \varepsilon^2\int_{-\infty}^{\infty} \mathrm{d}t\mathrm{e}^{-i\omega t}c_1(t)/2\pi \tag{6.3.50}$$

$$= \frac{\varepsilon^2}{12\pi}\left[\frac{1}{\lambda_1^2+\omega^2}\right] \tag{6.3.51}$$

6.3.3 平稳分布的渐近方法

对于任意的福克-普朗克方程

$$\partial_t p = -\sum_i \partial_i A_i(x)p + \frac{1}{2}\varepsilon^2\sum_{i,j}\partial_i\partial_j B_{ij}(x)p \tag{6.3.52}$$

若令

$$p_s(x) = \exp[-\phi(x)/\varepsilon^2] \tag{6.3.53}$$

则可产生此平稳解的一个渐近展开。我们看到，用 p_s 来表示时有

$$\left[\sum_i A_i(x)\partial_i\phi + \frac{1}{2}\sum_{i,j}B_{ij}(x)\partial_i\phi\partial_j\phi\right] + \varepsilon^2\left[-\sum_i\partial_i A_i(x) + \sum_{i,j}\partial_i B_{ij}\partial_j\phi + \sum_{i,j}\partial_i\partial_j B_{ij}(x)\right] = 0 \tag{6.3.54}$$

它的第一项是 ε^0 阶的项，是个哈密顿-雅可比方程。此结果的主要意义在于 $\phi(x)$ 原则上可以渐近展开为

$$\phi(x) = \sum_n \varepsilon^n\phi_n(x) \tag{6.3.55}$$

其中 $\phi_0(x)$ 满足

$$\sum_i A_i(x)\partial_i\phi_0 + \frac{1}{2}\sum_{i,j}B_{ij}(x)\partial_i\phi_0\partial_j\phi_0 = 0 \tag{6.3.56}$$

格雷厄姆及特尔(Graham and Tel) [6.8, 9] 最近说明了一般情况下可怎样解出方程(6.3.56)。其主要结果是这些解本身虽然是连续的，但除某些特殊情况外，其导函数却一般是有无穷多个不连

续点的，这些解与满足位势条件的福克-普朗克方程的情况密切相关。

6.4 快变变量的绝热消去

对一个动力学系统来说，非常常见的情况是它可以用随机微分方程来描写，这种方法具有广泛而不同的响应时间，并且其在十分短的时间内的行为是不重要的。这种情况的最自然的例子出现在布朗运动中。此时通常只观察布朗粒子的位置，但通常为不可观测的量也含于基本方程之中，这样，郎之万方程(1.2.4)可以改为位置与速度的方程如

$$\frac{\mathrm{d}x}{\mathrm{d}t}=v \tag{6.4.1}$$

$$m\frac{\mathrm{d}v}{\mathrm{d}t}=-\beta v+\sqrt{2k\beta T}\,\xi(t) \tag{6.4.2}$$

如果把这个方程理解为伊藤随机微分方程，则 4.4.6 节中已给出了解法。不过更简单的做法是先积分(6.4.2)而给出解

$$v(t)=v(0)\exp(-\beta t/m)+\frac{\sqrt{2k\beta T}}{m}\int_0^t\exp[-\beta(t-t')/m]\xi(t')\,\mathrm{d}t' \tag{6.4.3}$$

现在我们来考察摩擦系数 β 不小而质量 m 十分小的情况。此时对于使得

$$t\gg m/\beta\equiv\tau \tag{6.4.4}$$

的时刻 t，上式右边第一项中的指数因子可以忽略而且积分下限也可以延伸到 $-\infty$ 而不产生重大误差。因而

$$v(t)\to\frac{\sqrt{2k\beta T}}{m}\int_{-\infty}^t\exp[-(t-t')/\tau]\xi(t')\mathrm{d}t' \tag{6.4.5}$$

这里的 τ 称为弛豫时间，因为它确定了弛豫到 (6.4.5) 的时间尺度。

现在定义

$$\eta(t,\tau)=\tau^{-1}\int_{-\infty}^t\exp[-(t-t')/\tau]\mathrm{d}W(t') \tag{6.4.6}$$

由 4.4.4 节知道这是个平稳奥恩斯坦-乌伦贝克过程。其相关函

数是

$$\langle \eta(t,\ \tau)\eta(t',\ \tau)\rangle = \frac{1}{2\tau}\exp(-|t-t'|/\tau) \tag{6.4.7}$$

$$\xrightarrow[\tau\to 0]{} \delta(t-t') \tag{6.4.8}$$

我们看到 $\tau\to 0$ 的极限相应于白噪声极限，相关函数变成了 δ 函数。

于是可把(6.4.1)写成

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sqrt{\frac{2kT}{\beta}}\eta(t,\ \tau) \tag{6.4.9}$$

并且在极限过程 $\tau\to 0$ 下应变成

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sqrt{\frac{2kT}{\beta}}\xi(t) \tag{6.4.10}$$

另外一种而且是更为显而易见得多的获得上式的方法是宣称(6.4.2)中 $m\to 0$ 的极限过程就对应于令其左边等于0，所以有

$$v(t) = \sqrt{\frac{2kT}{\beta}}\xi(t) \tag{6.4.11}$$

上述推理是十分有启发性的但完全是不严格的，而且没有提供任何有系统的近似方法观念，而近似方法应该多半是按照一个小维数的参数所作的某种渐近展开．不仅如此，它甚至似乎不是任何种类的补充方法，诸如直接基于随机微分方程而作展开的方法——至少就作者所知，还没有人成功地开发出这种方案。

关于分布函数 $p(x,\ v,\ t)$ 的等价于 (6.4.12) 的福克-普朗克方程是

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(vp) + \frac{\partial}{\partial v}\left(\frac{vp}{\tau}\right) + \frac{kT}{\tau^2\beta}\frac{\partial^2 p}{\partial v^2} \tag{6.4.12}$$

若用

$$\hat{p}(x,\ t) = \int_{-\infty}^{\infty}\mathrm{d}v\,p(x,\ v,\ t) \tag{6.4.13}$$

来定义位置分布函数 $\hat{p}(x,\ t)$，则可以期望，相应于"简化了的"郎之万方程(6.4.10)，$\hat{p}(x,\ t)$所满足的福克-普朗克方程是

$$\frac{\partial \hat{p}}{\partial t} = \frac{kT}{\beta}\frac{\partial^2 \hat{p}}{\partial x^2} \tag{6.4.14}$$

我们要寻找一个以某种摄动方式从 (6.4.12) 导出 (6.4.14) 的方法，以便获得按适当小参数的幂次表示的高阶修正项。

更一般地，我们可以考虑在一位势中的布朗运动，其郎之万方程是(见5.3.6节)

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v \ ,$$
$$m\frac{\mathrm{d}v}{\mathrm{d}t} = -\beta v - V'(x) + \sqrt{2\beta kT}\,\xi(t) \tag{6.4.15}$$

β 变大的极限过程就要导致第二个方程非常迅速地弛豫到一个拟稳态状态，此时 $\frac{\mathrm{d}v}{\mathrm{d}t}\to 0$。因此，我们假设对足够大的 β，

$$v = -\frac{1}{\beta}[V'(x) - \sqrt{2\beta kT}\,\xi(t)] \tag{6.4.16}$$

把它代入(6.4.15)，就得到

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\beta^{-1}V'(x) + \sqrt{\frac{2kT}{\beta}}\,\xi(t) \tag{6.4.17}$$

它对应着 $\hat{p}(x)$ 所满足的福克-普朗克方程：

$$\boxed{\frac{\partial \hat{p}}{\partial t} = \beta^{-1}\left\{\frac{\partial}{\partial x}\left[V'(x)\hat{p} + kT\frac{\partial \hat{p}}{\partial x}\right]\right\}} \tag{6.4.18}$$

这就是所谓斯莫路乔夫斯基方程。

在此情况下我们已经消除了快变变量 v，已假设了这个变量会很快地弛豫到(6.4.16)式给定之值。

这个步骤就是所有绝热消去法程序的原型，它曾被当作哈肯的支配原理之基础而使用过[6.1]。其基本的物理假设是很大的 β (或者一般地，很短的弛豫时间)迫使含着大值 β 的方程所支配的变量(例如 v)弛豫到由假设慢变变量(在此情况下就是 x)为常数而得到的那个值。这样，这种快变变量就由慢变变量有效地支配了。

出人意料的是，斯莫路乔夫斯基方程的严格推导和对其作修正的估计问题直到最近才被解决。布林克曼(Brinkman)[6.2]首先对此作了研究。他仅估计了对(6.4.18)所作修正的数量级，而没有给出直到最低阶的所有修正项。最早的正确解答是斯特拉托

诺维奇给出的[文献6.3，第4章11.1节]。威伦斯基(G.Wilemski)[6.4]和泰特莱尔(U. M. Titulaer)[6.5]也独立地给出了正确的处理。

我们将在下面几节中就斯莫路乔夫斯基方程的推导及其修正问题提出一个系统而相当一般的理论，然后对更为一般的绝热消去问题进行考察。使用的是经改写的投影算子方法，这个方法已在统计物理、量子光学和有关领域中使用多年。这些方法可以直接按时间域的方式加以表述，但是我们发现更为方便的是利用拉普拉斯变换，这个方法就是原先由威伦斯基所用的方法。其表达方式类似于帕普尼科劳[6.6] (G. C. Papanicolaou)，他给出了应用于某些问题的严格基础。但是这里我们所用的证明大多是符号的形式演绎。

6.4.1 利用算子及投影算子的抽象表述

我们来考察福克-普朗克方程 (6.4.12) 的一种重新标度过的形式，即在5.3.6节中导出的(5.3.97)。它可以写为

$$\frac{\partial p}{\partial t} = (\gamma L_1 + L_2)\, p \tag{6.4.19}$$

的形式，其中 L_1 及 L_2 是由

$$L_1 = \frac{\partial}{\partial u}\left(u + \frac{\partial}{\partial u}\right) \tag{6.4.20}$$

$$L_2 = -\frac{\partial}{\partial y} u + U'(y)\,\frac{\partial}{\partial u} \tag{6.4.21}$$

所给出的微分算子。我们想对 y 的分布函数

$$\hat{p}(y, t) = \int \mathrm{d}u\, p(u, y, t) \tag{6.4.22}$$

来导出它的方程，要求这个方程当 γ 趋于十分大时也应有效。

我们预期(6.4.19)的一个近似解可由 $\hat{p}(y, t)$ 乘以

$$\frac{\partial p}{\partial t} = L_1 p \tag{6.4.23}$$

的平稳解，即

$$(2\pi)^{-1/2}\exp\left(-\frac{1}{2}u^2\right) \tag{6.4.24}$$

来得到。其理由在于对于很大的 γ，速度分布会迅速地被热化，或者更粗略地说，相比于 L_1 可以忽略 L_2，并且其解是一个 y 的函数乘以(6.4.23)的一个解，它将在一个量级为 γ^{-1} 的时间内到达一个十分小的平稳解。我们可以用定义投影算子 P 为

$$(Pf)(u, y) = (2\pi)^{-1/2}\exp\left(-\frac{1}{2}u^2\right)\int du' f(u', y) \tag{6.4.25}$$

来把此解表示出来，这里的 $f(u, y)$ 是任意函数。读者容易验证

$$P^2 = P \tag{6.4.26}$$

用所有 u 及 y 的函数组成的向量空间，P 就成了一个算子，它把任何上述向量投影到可表示为

$$g(u, y) = (2\pi)^{-1/2}\exp\left(-\frac{1}{2}u^2\right)\hat{g}(y) \tag{6.4.27}$$

的所有向量组成的向量空间中，其中 $\hat{g}(y)$ 是任意的 y 的函数。但(6.4.27)形式的函数是

$$L_1 g = 0 \tag{6.4.28}$$

的解全体，这也就是说，P 所投影于其中的空间是 L_1 的零空间。我们还要注意在此情况下有

$$\boxed{P = \lim_{t\to\infty}[\exp(L_1 t)]} \tag{6.4.29}$$

为证明它，可以像 5.2.5 节那样按 L_1 的本征函数 $P_\lambda(u)$ 来展开 u 与 y 的任意函数：

$$f(u, y) = \sum_\lambda A_\lambda(y) P_\lambda(u) \tag{6.4.30}$$

其中

$$A_\lambda(y) = \int du Q_\lambda(u) f(u, y) \tag{6.4.31}$$

于是

$$\lim_{t\to\infty}[\exp(L_1 t) f(u, y)] = \sum_\lambda A_\lambda(y)\lim_{t\to\infty} e^{-\lambda t} P_\lambda(u) \tag{6.4.32}$$

$$= P_0(u)\int du' Q_0(u') f(u', y) \tag{6.4.33}$$

注意对这个(奥恩斯坦-乌伦贝克)过程

$$P_0(u) = (2\pi)^{-1/2}\exp\left(-\frac{1}{2}u^2\right) \tag{6.4.34}$$

$$Q_0(u) = 1 \tag{6.4.35}$$

则可看出(6.4.29)得证。

不论是在这种还是其他一切情况下，我们还有个本质性的关系式

$$\boxed{PL_2P = 0} \tag{6.4.36}$$

这是因为考察 $PL_2Pf(u, y)$ 时，由 L_2 的定义可看出 $L_2Pf(u, y)$ 与

$$u\exp\left(-\frac{1}{2}u^2\right) \propto P_1(u) \tag{6.4.37}$$

成比例，而且由直接代入验证或者注意到 $P_1(u)$ 不在 L_1 的零空间中而得知

$$PP_1(u) = 0 \tag{6.4.38}$$

现在来定义

$$v = Pp \tag{6.4.39}$$

$$w = (1-P)p \tag{6.4.40}$$

使得

$$p = v + w$$

而且 v 在 L_1 的零空间中而 w 不在 L_1 的零空间中。

我们现在可以指出，由(6.4.29)有

$$PL_1 = L_1P = 0 \tag{6.4.41}$$

而且由原来的方程有

$$\begin{aligned}\frac{\partial v}{\partial t} &= P(\gamma L_1 + L_2)p \\ &= P(\gamma L_1 + L_2)[Pp + (1-P)p] \\ &= PL_2(1-P)p\end{aligned}$$

[其中利用了(6.4.41)及(6.4.36)]，这导致

$$\boxed{\frac{\partial v}{\partial t} = PL_2w} \tag{6.4.42}$$

类似地，

$$\frac{\partial w}{\partial t} = (1-P)(\gamma L_1 + L_2)p$$
$$= (1-P)(\gamma L_1 + L_2)[Pp + (1-P)p]$$
$$= \gamma L_1(1-P)p + (1-P)L_2(1-P)p + (1-P)L_2Pp \tag{6.4.43}$$

再利用 $PL_2P = 0$，便得到

$$\boxed{\frac{\partial w}{\partial t} = \gamma L_1 w + (1-P)L_2 w + L_2 v} \tag{6.4.44}$$

6.4.2 使用拉普拉斯变换的解法

基本方程 (6.4.42, 44) 可以用多种迭代法求解，但是既然它是线性方程，用拉普拉斯变换法来求解自然是十分恰当的，而且会顺当地产生出摄动展开。

时间的任何函数 $f(t)$ 之拉普拉斯变换是由

$$\tilde{f}(s) = \int_0^\infty e^{-st} f(t)\,dt \tag{6.4.45}$$

来定义的，而且对算子及抽象向量也很容易以相同方式来定义它。利用

$$\int_0^\infty e^{-st}\frac{df}{dt}dt = s\tilde{f}(s) - f(0) \tag{6.4.46}$$

我们把基本方程变成了

$$s\tilde{v}(s) = PL_2\tilde{w}(s) + v(0) \tag{6.4.47}$$
$$s\tilde{w}(s) = [\gamma L_1 + (1-P)L_2]\tilde{w}(s) + L_2\tilde{v}(s) + w(0)$$

这些方程是线性的算子方程，它们的形式解可直捷了当地得到。为简单起见让我们首先假定

$$w(0) = 0 \tag{6.4.48}$$

它意味着假定初始分布具有形式

$$p(u, y, 0) = (2\pi)^{-1/2}\exp\left(-\frac{1}{2}u^2\right)\hat{p}(y, 0) \tag{6.4.49}$$

这就是说假定了速度 u 的初始热化。于是，形式地有

$$\tilde{v}(s) = [s - \gamma L_1 - (1-P)L_2]^{-1} L_2 \tilde{v}(s) \tag{6.4.50}$$

并且因此

$$\boxed{s\tilde{v}(s) = PL_2[s - \gamma L_1 - (1-P)L_2]^{-1} L_2 \tilde{v}(s) + v(0)} \tag{6.4.51}$$

到此为止，至少形式地我们已经有了问题的完全解答了。对于任何有限的 s，我们可取 γ 的大值极限而发现

$$s\tilde{v}(s) \simeq -\gamma^{-1} PL_2 L_1^{-1} L_2 \tilde{v}(s) + v(0) \tag{6.4.52}$$

要注意 L_1^{-1} 不总是存在的。但是我们由(6.4.36)知道

$$PL_2\tilde{v}(s) = PL_2 P\tilde{p}(s) = 0 \tag{6.4.53}$$

因此 $L_2\tilde{v}(s)$ 不含 L_1 的零空间中的分量，并且由此 $L_1^{-1}L_2\tilde{v}(s)$ 总是存在的。

现在来看一看(6.4.52)在克雷默斯方程的情况下是什么样子的。(6.4.52)等价于微分方程

$$\boxed{\frac{\partial v}{\partial t} = -\gamma^{-1} PL_2 L_1^{-1} L_2 v} \tag{6.4.54}$$

请注意

$$L_2 v = \left[-\frac{\partial}{\partial y} u + U'(y) \frac{\partial}{\partial u} \right] P_0(u) \int du' p(u', y, t) \tag{6.4.55}$$

[$P_0(u)$是由(6.4.34)所定义的]。

对于这个问题来说，引进奥恩斯坦-乌伦贝克过程的本征函数[第5.2.6 c节]是有用的。对于参数

$$\frac{1}{2} D = k = 1 \tag{6.4.56}$$

它们具有形式

$$P_n(u) = (2\pi)^{-1/2} \exp\left(-\frac{1}{2}u^2\right) Q_n(u) \tag{6.4.57}$$

其中

$$Q_n(u) = (2^n n!)^{-\frac{1}{2}} H_n(u/\sqrt{2}) \tag{6.4.58}$$

利用

$$L_1 P_n(u) = -nP_n(u) \tag{6.4.59}$$

及厄密多项式的递推公式

$$xH_n(x) = \frac{1}{2}H_{n+1}(x) + nH_{n-1}(x) \tag{6.4.60}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}[\mathrm{e}^{-x^2}H_n(x)] = -\mathrm{e}^{-x^2}H_{n+1}(x) \tag{6.4.61}$$

则会看到

$$L_2 v = -\left[U'(y) + \frac{\partial}{\partial y}\right]P_1(u)\hat{p}(y) \tag{6.4.62}$$

从而使得利用(6.4.59)后有

$$L_1^{-1}L_2 v = \left[U'(y) + \frac{\partial}{\partial y}\right]P_1(u)\hat{p}(y) \tag{6.4.63}$$

现在再用一次 L_2，并再利用关系式(6.4.60, 61)，就会发现

$$L_2 P_1(u) = -[\sqrt{2}P_2(u) + P_0(u)]\frac{\partial}{\partial y} - \sqrt{2}P_2(u)U'(y) \tag{6.4.64}$$

$$PL_2L_1^{-1}L_2 v = -\frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right]\hat{p}(y)P_0(u) \tag{6.4.65}$$

并且运动方程(6.4.54)[在消去因子 $P_0(u)$ 后]具有形式

$$\boxed{\frac{\partial \hat{p}}{\partial t} = \gamma^{-1}\frac{\partial}{\partial y}\left[U'(y)\hat{p} + \frac{\partial \hat{p}}{\partial y}\right]} \tag{6.4.66}$$

它正是我们在 6.4 节中用单纯的消去法导出的斯莫路乔夫斯基方程。

6.4.3 短期行为

应该注意，在(6.4.52)中取 $\gamma \to \infty$ 的极限时意味着 s 是有限值的，或者粗略地说，

$$\gamma \gg s \tag{6.4.67}$$

所以，拉普拉斯变换法之解仅在

$$s \ll \gamma \tag{6.4.68}$$

成立时才是适用的。或者对于这个解来说，它将意味着有

$$t \gg \gamma^{-1} \tag{6.4.69}$$

我们定义

$$s_1 = s\gamma^{-1} \tag{6.4.70}$$

则(6.4.51)变成

$$\gamma s_1 \tilde{v} = PL_2[s_1\gamma - L_1\gamma - (1-P)L_2]^{-1}L_2\tilde{v} + v(0) \tag{6.4.71}$$

取 $\gamma \to \infty$ 的极限，则

$$\gamma_1 s_1 \tilde{v} \simeq \gamma^{-1} PL_2(s - L_1)^{-1}L_2\tilde{v} + v(0) \tag{6.4.72}$$

利用 $L_2\tilde{v}$ 与 $P_1(u)$ 成比例的事实(6.4.52)，我们看到

$$\gamma s_1 \tilde{v} = \gamma^{-1}(s_1 + 1)^{-1} PL_2^2\tilde{v} + v(0) \tag{6.4.73}$$

再改回到变元 s 并予整理，就有

$$s\tilde{v} = \gamma^{-1}\left(\frac{s}{\gamma} + 1\right)^{-1} PL_2^2\tilde{v} + v(0) \tag{6.4.74}$$

而它是等价于

$$\frac{\partial v}{\partial t} = \int_0^t \mathrm{d}t' \exp[\gamma(t'-t)] PL_2^2 v(t') \tag{6.4.75}$$

的。另一方面，我们可改写(6.4.74)成

$$\frac{1}{\gamma}[s^2\tilde{v} - sv(0)] + [s\tilde{v} - v(0)] = \gamma^{-1}PL_2^2 v \tag{6.4.76}$$

的形式。利用(6.4.46)并同 6.4.2 节的做法，就知道上式等价于 $\hat{p}$ 的方程：

$$\boxed{\frac{1}{\gamma}\frac{\partial^2\hat{p}}{\partial t^2} + \frac{\partial\hat{p}}{\partial t} = \gamma^{-1}\frac{\partial}{\partial y}\left[U'(y)\hat{p} + \frac{\partial\hat{p}}{\partial y}\right]} \tag{6.4.77}$$

因为 $\int_0^\infty \mathrm{e}^{-st} f''(t)\mathrm{d}t = s^2\tilde{f}(s) - sf(0) - f'(0)$

并且在(6.4.76)式的第一个括号中没有出现常数项，所以上式蕴含了初始条件

$$\frac{\partial\hat{p}}{\partial t}(0) = 0$$

我们还可类似地改写(6.4.75)或积分(6.4.77)式来得到

$$\boxed{\frac{\partial\hat{p}}{\partial t} = \frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right]\int_0^t \mathrm{d}t' \exp[\gamma(t'-t)]\hat{p}(t')} \tag{6.4.78}$$

方程(6.4.77, 78)说明了一个从(6.4.78)中可明显看到的非马尔可夫特性，它表明预测 $\hat{p}(t+\Delta t)$ 要求知道 $0\leqslant t'\leqslant t$ 期间的 $\hat{p}(t')$。但是核 $\exp[\gamma(t'-t)]$ 只有当 $|t'-t|\sim\gamma^{-1}$ 时才显著地区别于零而且在一个远比它长得多的时间范围内(6.4.78)可由斯莫路乔夫斯基方程(6.4.66)作为近似。形式上，我们可在(6.4.78)中利用分部积分来实现它，

$$\int_0^t \mathrm{d}t'\exp\left[\gamma(t'-t)\right]\hat{p}(t') = \frac{\hat{p}(t)-\mathrm{e}^{-\gamma t}\hat{p}(0)}{\gamma} - \gamma^{-1}\int_0^t \exp[\gamma(t'-t)]\frac{\partial\hat{p}}{\partial t}\,\mathrm{d}t' \tag{6.4.79}$$

最后一项是 γ^{-2} 阶的，所以忽略它之后就会发现斯莫路乔夫斯基方程已被

$$\boxed{\frac{\partial\hat{p}}{\partial t} = \frac{\partial}{\partial y}\left[U'(y)+\frac{\partial}{\partial y}\right][\hat{p}(t)-\mathrm{e}^{-\gamma t}\hat{p}(0)]} \tag{6.4.80}$$

所代替了。这个方程在精确到 γ 的最低阶程度上对于所有的时刻，即从十分短的时刻($<\gamma^{-1}$)到十分长的时刻($\gg\gamma^{-1}$)，都等价于(6.4.78)。它说明了表征性的"记忆时间"γ^{-1}，这是此方程在接近斯莫路乔夫斯基的形式以前所经过的时间。

于是在精确到这种程度上，此过程可由一马尔可夫过程来近似，但在同样的精确度上又存在可能不是马尔可夫过程的另外表示方式(6.4.77, 78)。显然，在任何更高精确性程度上，我们必然有一个非马尔可夫过程的表示。

6.4.4 边界条件

我们考虑一个竖在 $x=a$ 处的壁，粒子限制在 $y\leqslant a$ 的范围内。$u>0$ 与 $u<0$ 时的行为是不同的。

从随机微分方程

$$\begin{aligned}\mathrm{d}y &= u\mathrm{d}t\\ \mathrm{d}u &= -[U'(y)+\gamma u]\mathrm{d}t+\sqrt{2\gamma}\,\mathrm{d}W(t)\end{aligned} \tag{6.4.81}$$

我们看出

对 $u>0$, $y=a$ 是一出口边界.

对 $u<0$, $y=a$ 是一入口边界.

这是因为 $u>0$ 时一个在 $y=a$ 处的粒子必然要跑到 $y>a$ 或被吸收；类似地，若 $u<0$，则边界左边的粒子就永不能达到 $y=a$. 按惯例，我们把 $y=a$ 按如下方式刻画为吸收的或反射的：

(i) 吸收壁：当 $u>0$ 时粒子被吸收，没有 $u<0$ 的粒子

$$\Longrightarrow p(u, a, t)=0 \quad u>0 \qquad =0 \quad u<0 \tag{6.4.82}$$

第一个条件是通常的吸收边界条件，如在5.2.1节中所导出的那样，第二个条件表示着一个事实，即任何处于 $u<0$ 的粒子在 $y=a$ 时就立刻跑到 $y<a$，而且没有任何新粒子会引进来。

吸收壁条件显然意味着

$$\hat{p}(a, t)=0 \tag{6.4.83}$$

这就是单变元福克-普朗克方程通常的吸收壁条件。

(ii) 反射壁：物理学上在 $y=a$ 处的反射意味着有个具有速度 u 的到达 $y=a$ 的粒子，并且它会立刻以不同的但是为负的速度反射回来。若假设

$$u \longrightarrow -u$$

那末如5.2.1节及5.3.2节中那样，就有一个"周期性边界条件"，这意味着

$$p(u, a, t)=p(-u, a, t) \tag{6.4.84}$$

而且离开 (u, a) 的流的法向分量等于进入 $(-u, a)$ 的流的法向分量。但是由于对于该方程

$$J=\left\{-up, [\gamma u+U'(y)]p+\gamma\frac{\partial p}{\partial u}\right\} \tag{6.4.85}$$

并且法向分量就是沿着 y 方向的分量，我们看到(6.4.84)与(6.4.85)是等价的。

这个边界条件十分自然地导致了斯莫路乔夫斯基方程的边界条件。

由于(6.4.84)意味着只有偶数阶本征函数 $P_n(u)$ 才会在 $p(u,$

$a, t)$的展式中出现，因此

$$(1-P)p(u, a, t) = w(u, a, t) \tag{6.4.86}$$

只含有偶的本征函数，而且其拉普拉斯变换 $\tilde{w}(u, a, s)$ 也这样。但是由(6.4.50)我们看到 γ^{-1} 的最低阶有

$$\tilde{w}(u, a, s) \simeq (-\gamma^{-1}L^{-1}L_2\tilde{v})(u, a, s) \tag{6.4.87}$$

而且利用(6.4.63)后

$$= -P_1(u)\gamma^{-1}\left[U'(y)\hat{p}(y) + \frac{\partial\hat{p}(y)}{\partial y}\right]_{y=a} \tag{6.4.88}$$

由于它与奇本征函数 $P_1(u)$ 成比例而为零，故我们导出了

$$\left[U'(y)\hat{p}(y) + \frac{\partial\hat{p}(y)}{\partial y}\right]_{y=a} = 0 \tag{6.4.89}$$

它就是斯莫路乔夫斯基方程正确的反射壁边界条件。

不难证明对 6.4.3 节中的方程类似地可以导出相同的边界条件。

6.4.5 系统的摄动分析

我们回到(6.4.17)上来，但为简单起见仍规定 $w(0) = 0$。于是有

$$\tilde{w}(s) = [s - \gamma L_1 - (1-P)L_2]^{-1}L_2\tilde{v}(s) \tag{6.4.90}$$

以及

$$s\tilde{v}(s) = PL_2[s - \gamma L_1 - (1-P)L_2]^{-1}L_2\tilde{v}(s) + v(0) \tag{6.4.91}$$

我们可以直捷了当地展开(6.4.91)中的逆算子为 γ 的幂。但是必须决定展开式中 s 大小的量级。由前面几节我们看到可定义

$$s_1 = s\gamma^{-1}, \quad \tilde{v} \longrightarrow \gamma^{-1}\tilde{v} \tag{6.4.92}$$

由此产生出可明显展示在 γ^{-1} 量级大小时间内的初始时间行为的展开式。

另外的办法是定义

$$s_2 = s\gamma, \quad \tilde{v} \longrightarrow \gamma\tilde{v} \tag{6.4.93}$$

我们将发现由它可产生一个只展示长时间行为的展开式。

在代入(6.4.93)后，我们发现(6.4.91)成了

$$s_2\tilde{v} = -PL_2[L_1 + (1-P)L_2\gamma^{-1} - s_2\gamma^{-2}]^{-1}L_2\tilde{v} + v(0) \tag{6.4.94}$$

而且现在可以写成

$$s_2\tilde{v} = -PL_2L_1^{-1}L_2\tilde{v} + v(0) \quad (\gamma\to\infty) \tag{6.4.95}$$

但反之，若没有作这种代换的话，$\gamma\to\infty$ 时的极限就只产生出

$$s\tilde{v} = v(0) \tag{6.4.96}$$

事实上这与(6.4.52)一致，而实际上我们并未在获得后者时取极限 $\gamma\to\infty$。

在代换 $s_1 = s\gamma^{-1}$ 下所得到的

$$s_1\tilde{v} = \gamma^{-2}PL_2[s_1 - L_1 - (1-P)L_2\gamma^{-1}]L_2\tilde{v}(s) + v(0) \tag{6.4.97}$$

在 $\gamma\to\infty$ 时并不趋于一个极限，但它的确产生了一个类似于普通摄动问题的展开式，而且正如我们已看到的，它表现了短期行为。

a) 长时间摄动理论

我们可以展开(6.4.94)到 γ^{-2} 阶而具有

$$s_2\tilde{v} = [A + B\gamma^{-1} + (C + Ds_2)\gamma^{-2}]\tilde{v} + v(0) \tag{6.4.98}$$

的形式，其中

$$\begin{aligned} A &= -PL_2L_1^{-1}L_2 \\ B &= PL_2L_1^{-1}(1-P)L_2L_1^{-1}L_2 \\ C &= -PL_2L_1^{-1}(1-P)L_2L_1^{-1}(1-P)L_2L_1^{-1}L_2 \\ D &= PL_2L_1^{-2}L_2 \end{aligned} \tag{6.4.99}$$

(6.4.98)经过整理，就得到

$$s_2(1-\gamma^{-2}D)\tilde{v} = [A + B\gamma^{-1} + C\gamma^{-2}]\tilde{v} + v(0) \tag{6.4.100}$$

或者精确到 γ^{-2} 阶，

$$s_2\tilde{v} = [A + B\gamma^{-1} + (C + DA)\gamma^{-2}]\tilde{v} + (1+\gamma^{-2}D)v(0) \tag{6.4.101}$$

这给出了 v 的运动方程之拉普拉斯变换，在此式中初始条件不是 $v(0)$ 而是 $(1+\gamma^{-2}D)v(0)$。注意，这个方程关于时间是一阶的，这是因为 $s_2^n(n>1)$ 不出现。当然这对于展开到 γ 的高阶展式来说是办不到的事。

b) 应用于布朗运动

对于布朗运动我们已经计算过 (6.4.99) 的算子 A；它是由(6.4.65)给出的

$$\boxed{A = -PL_2L_1^{-1}L_2 = \frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right]} \tag{6.4.102}$$

其他算子也可类似地算出来。

例如，由(6.4.63, 64)就有

$$L_2L_1^{-1}L_2 = -\left\{P_0(u)\frac{\partial}{\partial y} + \sqrt{2}\,P_2(u)\left[U'(y) + \frac{\partial}{\partial y}\right]\right\}\left\{U'(y) + \frac{\partial}{\partial y}\right\}$$

只要乘以$(1-P)$就消掉了含 P_0 的项，并且在用 L_1^{-1} 乘它以后再乘 $-\frac{1}{2}$。这就有了

$$L_1^{-1}(1-P)L_2L_1^{-1}L_2 = \sqrt{2}\,P_2(u)\left[U'(y) + \frac{\partial}{\partial y}\right]^2$$

如果用 L_2 去乘时利用厄密多项式的递推关系式(6.4.60, 61)，则可导出

$$\begin{aligned} L_2P_2(u) &= \left[-\frac{\partial}{\partial y}u + U'(y)\frac{\partial}{\partial u}\right]P_2(u) \\ &= -\sqrt{3}\,U'(y)P_3(u) - \frac{\partial}{\partial y}\left[\sqrt{3}\,P_3(u) \right. \\ &\quad \left. + \sqrt{2}\,P_1(u)\right] \end{aligned} \tag{6.4.103}$$

最后，因为 $P_0(u)$ 不出现，所以再乘以 P 后就把所有的项化为零了。因而

$$\boxed{B = PL_2L_1^{-1}(1-P)L_2L_1^{-1}L_2 = 0} \tag{6.4.104}$$

C 与 DA 的计算也可类似地进行。

我们发现有

$$C = \frac{\partial^2}{\partial y^2}\left[U'(y) + \frac{\partial}{\partial y}\right]\left[U'(y) + \frac{\partial}{\partial y}\right] \tag{6.4.105}$$

和

$$DA = -\frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right]\frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right] \qquad (6.4.106)$$

从而

$$C + DA = \frac{\partial}{\partial y}U''(y)\left[U'(y) + \frac{\partial}{\partial y}\right] \qquad (6.4.107)$$

而且(6.4.101)等价于下列微分方程

$$\boxed{\frac{\partial p}{\partial t} = \gamma^{-1}\frac{\partial}{\partial y}[1 + \gamma^{-2}U''(y)]\left[U'(y)p + \frac{\partial p}{\partial y}\right]} \qquad (6.4.108)$$

其初始条件为

$$\lim_{t\to 0}\hat{p}(y, t) = \left\{1 - \gamma^{-1}\frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right]\right\}\hat{p}(y, 0) \qquad (6.4.109)$$

这种初始条件的改变是一种“分层”现象的反映。方程(6.4.108)对 $t \gg \gamma^{-1}$ 有效，被称为修正的斯莫路乔夫斯基方程。

精确解应该考虑到 $t \sim \gamma^{-1}$ 时间以前的行为，此时会出现 $\exp(-\gamma t)$ 类型的项。图形上就如图 6.1 所示。改变了的初始条件是考虑到了靠近 $t \sim \gamma^{-1}$ 时初始分层的影响。

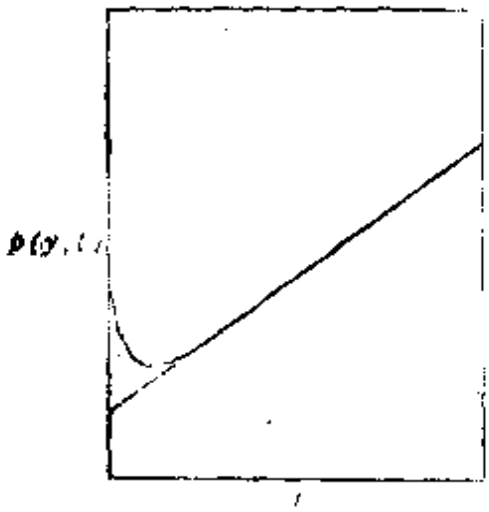


图 6.1 边界点上的分层现象
精确解(实线)靠近左边的边界时迅速变化。除了这边界点附近外，近似(虚线)是很好的，于是近似的适当边界条件就用一个较小的值给出，它就是虚线与边界相交处所取的值

c) 边界条件

对边界条件的高阶补充是不能用本节的方法来实现的，这是因为在接近边界的地方关于变元 y 出现了一个快速变化层而使算子 $\frac{\partial}{\partial y}$ 为有界的假设变成不合理。关于这个问题泰特莱尔及其合作者们已取得了有重大意义的进展 [6.10～6.12]。假设边界是在 $y = 0$ 处，则可把变换 $z = \gamma y$ 及 $\tau = \gamma t$ 代入克雷默斯方程(6.4.19)以得到

$$\frac{\partial P}{\partial t}=\left[\frac{\partial}{\partial u}\left(u+\frac{\partial}{\partial u}\right)-u\frac{\partial}{\partial u}\right]P+\frac{1}{\gamma}U'(z/\gamma)\frac{\partial P}{\partial u} \quad (6.4.110)$$

从而使得其零阶问题的解就是此方程独立于 γ 的那部分的解：对于零阶而言势是无关的。至今，只有(6.4.110)的平稳解被认真研究过。可以证明当 $\gamma\to\infty$ 时(6.4.110)的平稳解可写成

$$P(u,z)=\psi_0^2(u,z)+d_0^2\psi_0(u,z)+\sum_{n=1}^{\infty}d_n^2\psi_n(u,z) \quad (6.4.111)$$

的形式，其中

$$\psi_0(u,z)=(2\pi)^{-1/2}\exp\left(-\frac{1}{2}u^2\right) \quad (6.4.112)$$

$$\psi_1(u,z)=(2\pi)^{-1/2}(z-u)\exp\left(-\frac{1}{2}u^2\right) \quad (6.4.113)$$

而且 $\psi_n(u,z)$ 是与厄密多项式有关的某种复杂函数。确定系数 d_n^2 的问题并不是显而易解的，读者在[6.12]中可获得一种处理方法。人们发现此解在 $z=0$ 处具有无穷可导性，对值小的 z 来说，它还具有 $a+bz^{1/3}$ 的形式。

6.5 白噪声过程作为非白噪声过程的极限

关于真实噪声与白噪声之间的关系我们在前面的 1.4.4 节和 4.1 节中已谈过了。我们所关心的是微分方程

$$\frac{dx}{dt}=a(x)+b(x)\alpha_0(t) \quad (6.5.1)$$

的极限，其中 $\alpha_0(t)$ 是具有某种非零相关时间的随机源。我们将证明，只要 $\alpha_0(t)$ 是马尔可夫过程，那末在将它变成一个 δ 相关过程的极限过程之下，微分方程将变成具有相同系数的斯特拉托诺维奇随机微分方程，也就是说，成了

$$(S)dx=a(x)dt+b(x)dW(t) \quad (6.5.2)$$

它等价于伊藤型方程

$$dx=\left[a(x)+\frac{1}{2}b(x)b'(x)\right]dt+b(x)dW(t) \quad (6.5.3)$$

为了得到 δ 相关的马尔可夫过程这个极限，我们需要取 γ 增大时

$$\alpha_0(t)=\gamma\alpha(\gamma^2t) \quad (6.5.4)$$

的极限，其中 $\alpha(\tau)$ 是具有

$$\langle \alpha(\tau) \rangle = 0 \tag{6.5.5}$$

$$\langle \alpha(\tau)\alpha(0) \rangle_s = g(\tau) \tag{6.5.6}$$

的平稳随机过程。此时

$$\langle \alpha_0(t) \rangle = 0$$

$$\langle \alpha_0(t)\alpha_0(0) \rangle_s = \gamma^2 g(\gamma^2 t) \tag{6.5.7}$$

在 $\gamma \to \infty$ 的极限之下，相关函数要变成 δ 函数。更确切地说，我们假设

$$\int_{-\infty}^{\infty} g(\tau)\mathrm{d}\tau = 1 \tag{6.5.8}$$

而且

$$\int_{-\infty}^{\infty} |\tau| g(\tau)\mathrm{d}\tau = \tau_c \tag{6.5.9}$$

定义着过程 $\alpha(t)$ 的相关时间。[若 $g(\tau)$ 是指数型的，则在(6.5.9)中定义的 τ_c 就要求

$$g(\tau) \propto \exp(-\tau/\tau_c)$$

这是与1.4.4节及3.7.1节中的用法一致的。]

那末显然 $\alpha_0(t)$ 的相关时间是 τ_c/γ^2，当 $\gamma \to \infty$ 时它就变成零；进而还有

$$\lim_{\gamma \to \infty} \langle \alpha_0(t)\alpha_0(0) \rangle_s = 0 \quad (t \neq 0) \tag{6.5.10}$$

但是无论怎样我们总有

$$\int_{-\infty}^{\infty} \langle \alpha_0(t)\alpha_0(0) \rangle_s \mathrm{d}t = \int_{-\infty}^{\infty} g(\tau)\mathrm{d}\tau = 1 \tag{6.5.11}$$

而使我们可写出

$$\lim_{\gamma \to \infty} \langle \alpha_0(t)\alpha_0(0) \rangle_s = \delta(t) \tag{6.5.12}$$

因此当 $\gamma \to \infty$ 时 $\alpha_0(t)$ 的极限正好对应着归一化了的白噪声。高阶相关函数曾被设想也是重要的，但结果却并非如此。

对于 $\alpha(\tau)$ 是马尔可夫扩散过程且其福克-普朗克方程是

$$\frac{\partial p(\alpha)}{\partial t} = -\frac{\partial}{\partial \alpha}[A(\alpha)p(\alpha)] + \frac{1}{2}\frac{\partial^2}{\partial \alpha^2}[B(\alpha)p(\alpha)] \tag{6.5.13}$$

的情况，我们来给出一个证明。假设意味着(x, α)的福克-普朗克方程是

$$\frac{\partial p(x, \alpha)}{\partial t} = (\gamma^2 L_1 + \gamma L_2 + L_3)\, p(x, \alpha) \tag{6.5.14}$$

其中

$$\begin{aligned} L_1 &= -\frac{\partial}{\partial x} A(\alpha) + \frac{1}{2} \frac{\partial^2}{\partial \alpha^2} B(\alpha) \\ L_2 &= -\frac{\partial}{\partial x} b(x)\alpha \\ L_3 &= -\frac{\partial}{\partial x} a(x) \end{aligned} \tag{6.5.15}$$

其渐近分析可类似于6.4.1，6.4.3节去做，只不过要考虑L_3算子而在方法上稍作修改。类似于1.4.1节，我们可以在x及α的函数全体组成的空间上用

$$(Pf)(x, \alpha) = p_s(\alpha) \int d\alpha' f(x, \alpha') \tag{6.5.16}$$

来定义一个投影算子P，其中$p_s(\alpha)$是

$$L_1 p_s(\alpha) = 0 \tag{6.5.17}$$

的解。

我们假设α的平稳分布均值$\langle\alpha\rangle_s$是零。这意味着投影算子P满足实质性条件

$$PL_2P = 0 \tag{6.5.18}$$

这是因为

$$\begin{aligned} &(PL_2Pf)(x, \alpha) \\ &= p_s(\alpha) \int d\alpha' \left[-\frac{\partial}{\partial x} b(x)\alpha' p_s(\alpha') \right] \int d\alpha' f(x, \alpha') \\ &= -p_s(\alpha) \langle\alpha\rangle_s \frac{\partial}{\partial x} b(x) \int d\alpha' f(x, \alpha') = 0 \end{aligned} \tag{6.5.19}$$

的缘故。同样明显的是

$$PL_3 = L_3P \tag{6.5.20}$$

而且与前面一样，

$$PL_1 = L_1P = 0 \tag{6.5.21}$$

与前面相同，定义了

$$v = Pp \tag{6.5.22}$$

$$w = (1-P)p \tag{6.5.23}$$

并且用符号 $\tilde{v}$ 及 $\tilde{w}$ 表示其相应的拉普拉斯变换后，就得到

$$\begin{aligned} s\tilde{v}(s) &= P(\gamma^2 L_1 + \gamma L_2 + L_3)\tilde{p}(s) + v(0) \\ &= \gamma P L_2[P\tilde{p}(s) + (1-P)\tilde{p}(s)] + L_3 P\tilde{p}(s) + v(0) \end{aligned} \tag{6.5.24}$$

从而

$$\boxed{s\tilde{v}(s) = -\gamma P L_2 \tilde{w}(s) + L_3\tilde{v}(s) + v(0)} \tag{6.5.25}$$

而且类似地有

$$\boxed{s\tilde{w}(s) = [\gamma^2 L_1 + \gamma(1-P)L_2 + L_3]\tilde{w}(s) + \gamma L_2\tilde{v}(s) + w(0)} \tag{6.5.26}$$

它们之所以与(6.4.47)不同只是由于 (6.5.25) 中存在着 $L_3\tilde{v}(s)$ 项以及(6.5.26)中有 $L_3\tilde{w}$ 项。我们再假设 $w(0)=0$，它意味着 $\alpha(t)$ 是平稳马尔可夫过程，则

$$\begin{aligned} s\tilde{v}(s) = L_3\tilde{v}(s) &- \gamma P L_2[-s + \gamma^2 L_1 + \gamma(1-P)L_2 \\ &+ L_3]^{-1}\gamma L_2\tilde{v}(s) + v(0) \end{aligned} \tag{6.5.27}$$

现在取 $\gamma \to \infty$ 的极限，就给出了

$$\boxed{s\tilde{v}(s) \simeq (L_3 - PL_2L_1^{-1}L_2)\tilde{v}(s) + v(0)} \tag{6.5.28}$$

下面来计算 $PL_2L_1^{-1}L_2\tilde{v}$。记

$$\tilde{v}(s) = \tilde{p}(x)p_s(\alpha) \tag{6.5.29}$$

则

$$\begin{aligned} PL_2L_1^{-1}L_2\tilde{v} = p_s(\alpha)\int d\alpha' \left[-\frac{\partial}{\partial x}b(x)\alpha'\right] \\ \cdot L_1^{-1}\left[-\frac{\partial}{\partial x}b(x)\alpha'\right]p_s(\alpha')\tilde{p}(x) \end{aligned} \tag{6.5.30}$$

下面要计算

$$\int d\alpha \alpha L_1^{-1}\alpha p_s(\alpha) \equiv -D \tag{6.5.31}$$

并且为此需要 L_1^{-1} 的一个方便的表示。考虑到

$$\int_0^t \exp(L_1 t')dt' = L_1^{-1}\exp(L_1 t) - L_1^{-1} \tag{6.5.32}$$

并利用(6.4.29)，有

$$\int_0^\infty \exp(L_1 t)\mathrm{d}t = -L_1^{-1}(1-P) \tag{6.5.33}$$

由假设，

$$P\alpha p_s(\alpha) = p_s(\alpha)\int \mathrm{d}\alpha' \alpha' p_s(\alpha') = p_s(\alpha)\langle\alpha\rangle_s = 0 \tag{6.5.34}$$

我们有

$$D = \int \mathrm{d}\alpha \alpha \int_0^\infty \exp(L_1 t)\alpha p_s(\alpha)\mathrm{d}t \tag{6.5.35}$$

我们注意到

$$\exp(L_1 t)\alpha p_s(\alpha) \tag{6.5.36}$$

是福克-普朗克方程

$$\partial_t f = L_1 f \tag{6.5.37}$$

在初始条件

$$f(\alpha, 0) = \alpha p_s(\alpha) \tag{6.5.38}$$

下的解，所以

$$\exp(L_1 t)\alpha p_s(\alpha) = \int \mathrm{d}\alpha' p(\alpha, t|\alpha', 0)\alpha' p_s(\alpha') \tag{6.5.39}$$

而且把它代入(6.5.35)后得到

$$D = \int_0^\infty \mathrm{d}t \int \mathrm{d}\alpha \mathrm{d}\alpha' \alpha \alpha' p(\alpha', t|\alpha, 0) p_s(\alpha) \tag{6.5.40}$$

也就是说

$$\boxed{D = \int_0^\infty \mathrm{d}t \langle \alpha(t)\alpha(0)\rangle_s} \tag{6.5.41}$$

而且由(6.5.8)以及相关函数的对称性

$$D = \frac{1}{2} \tag{6.5.42}$$

在利用(6.5.42)式作为D之值后，就得到

$$-PL_2 L_1^{-1} L_2 \bar{v} = \frac{1}{2} p_s(\alpha) \frac{\partial}{\partial x}\left[b(x) \frac{\partial}{\partial x} b(x) \hat{p}(x) \right] \tag{6.5.43}$$

从而相应于(6.5.28)，

$$\hat{p}(x, t) = \int \mathrm{d}\alpha p(x, \alpha) \tag{6.5.44}$$

的微分方程是

$$\frac{\partial \hat{p}}{\partial t} = -\frac{\partial}{\partial x}\hat{p}(x) + \frac{1}{2}\frac{\partial}{\partial x}b(x)\frac{\partial}{\partial x}b(x)\hat{p}(x) \tag{6.5.45}$$

自然，这就是对应于斯特拉托诺维奇形式的

$$(S)\,\mathrm{d}x = a(x)\mathrm{d}t + b(x)\mathrm{d}W(t) \tag{6.5.46}$$

或如前面已证实了的，对应于伊藤型形式的

$$\mathrm{d}x = \left[a(x) + \frac{1}{2}b'(x)b(x)\right]\mathrm{d}t + b(x)\mathrm{d}W(t) \tag{6.5.47}$$

之福克-普朗克方程。

6.5.1 结果的一般性

从证明一眼可看出，我们对 $\alpha(t)$ 之全部要求就在于要构造一个具有零均值和具有形式为

$$\frac{\partial p(\alpha)}{\partial t} = L_1 p(\alpha) \tag{6.5.48}$$

的发展方程的平稳马尔可夫过程，其中 L_1 是个线性算子。这对于任何类型的马尔可夫过程来说都是可能做到的，特别地，譬如说，当 $\alpha(t)$ 为仅取 $\pm a$ 值的随机电报过程时就是如此。当取 $\gamma \to \infty$ 的极限时，所得到的仍然是个福克-普朗克方程。这是中心极限定理的一种表现。因为当 $\gamma \to \infty$ 时作用明显的高斯白噪声是由许多个别分量叠加组成的，并且其最后结果仍然是一个明显有效的高斯噪声。事实上，帕普尼科劳和科勒 (G. C. Papanicolauou & W. Kohler) [6.7] 已经严格证明即使 $\alpha(t)$ 是一非马尔可夫过程最后结果仍然如此，只要它是"强混合的"，粗略地说，这个条件意味着其所有的相关函数在一个较大的时差内很快地衰减下去。

6.5.2 更一般的涨落方程

注意若不在(6.5.1)简单地用 $\gamma\alpha(t/\gamma^2)$ 作为 $\alpha_0(t)$ 的定义，则可以用更一般些的形式

$$\alpha_0(t, x) = \gamma\psi[x, \alpha(t/\gamma^2)] \tag{6.5.49}$$

并且由于对 x 的所有相依性可以包含于 ψ 中，我们可以只考虑

$b(x)=1$ 的情况。类似于前面 $\langle\alpha\rangle_s=0$ 的假设，我们假设

$$\int d\alpha\psi(x,\ \alpha)p\ (\alpha)=0 \tag{6.5.50}$$

于是 D 成了与 x 相依，并且我们必须用算式

$$D(x)=\int_0^\infty dt\langle\psi[x,\ \alpha(t)]\psi[x,\ \alpha(0)]\rangle \tag{6.5.51}$$

和

$$E(x)=\int_0^\infty dt\left\langle\frac{\partial\psi}{\partial x}[x,\ \alpha(t)]\psi[x,\ \alpha(0)]\right\rangle \tag{6.5.52}$$

福克-普朗克方程也成了

$$\frac{\partial\hat{p}}{\partial t}=-\frac{\partial}{\partial x}[\{a(x)+E(x)\}\hat{p}]+\frac{\partial^2}{\partial x^2}[D(x)\hat{p}] \tag{6.5.53}$$

在这种形式下我们已经与斯特拉托诺维奇[文献6.3，方程(4.4.39)]所导出的相一致了。

6.5.3 非齐次系统

如果用

$$\frac{dx}{dt}=a(x,\ t)+b(x,\ t)\alpha_0(t) \tag{6.5.54}$$

代替(6.5.1)，则不能简单地用拉普拉斯变换方法。我们可借下面的手法来避开这个困难。引入一个附加变量 τ 使方程变成

$$dx=[a(x,\ \tau)+\gamma b(x,\ \tau)\alpha]dt \tag{6.5.55}$$

$$d\alpha=\gamma^2A(\alpha)dt+\gamma\sqrt{B(\alpha)}\,dW(t) \tag{6.5.56}$$

$$d\tau=dt \tag{6.5.57}$$

最后的方程约束了 t 是与 τ 相同的，但是系统现在成了关于变量 $(x,\ \alpha,\ \tau)$ 的齐次马尔可夫过程了。的确，任何非齐次的马尔可夫过程都可以用这种手法写成一个齐次的马尔可夫过程。

现在福克-普朗克方程是

$$\frac{\partial p}{\partial t}=\gamma^2L_1+\gamma L_2+L_3 \tag{6.5.58}$$

其中

$$L_1 = -\frac{\partial}{\partial \alpha} A(\alpha) + \frac{1}{2}\frac{\partial^2}{\partial \alpha^2} B(\alpha) \tag{6.5.59}$$

$$L_2 = -\frac{\partial}{\partial x} b(x, \tau)\alpha \tag{6.5.60}$$

$$L_3 = -\frac{\partial}{\partial t} - \frac{\partial}{\partial x} a(x, \tau) \tag{6.5.61}$$

利用相同于前面的步骤，我们可得到

$$\frac{\partial \hat{p}}{\partial t} = \left[-\frac{\partial}{\partial \tau} - \frac{\partial}{\partial x} a(x, \tau) + \frac{1}{2}\frac{\partial}{\partial x} b(x, \tau)\frac{\partial}{\partial x} b(x, \tau)\right]\hat{p} \tag{6.5.62}$$

它导致

$$d\tau = dt \tag{6.5.63}$$

而使得用 t 消去 τ 后得到

$$\boxed{\frac{\partial \hat{p}}{\partial t} = \left[-\frac{\partial}{\partial x} a(x, t) + \frac{1}{2}\frac{\partial}{\partial x} b(x, t)\frac{\partial}{\partial x} b(x, t)\right]\hat{p}} \tag{6.5.64}$$

它与(6.5.45)完全类同.

6.5.4 算子 L_1 与时间相依的后果

若另外还假设 A、B 也依赖于时间，使得

$$L_1 = -\frac{\partial}{\partial \alpha} A(\alpha, \tau) + \frac{1}{2}\frac{\partial^2}{\partial x^2} B(\alpha, \tau) \tag{6.5.65}$$

我们则看到 P 是 τ 的函数，因而就不能与 L_3 交换. 于是

$$PL_3 \neq L_3 P \tag{6.5.66}$$

尽管这样，我们还是可以对付它的. 如前面一样定义 $\tilde{v}(s)$ 及 $\tilde{w}(s)$，就有

$$s\tilde{v}(s) = P(\gamma L_2 + L_3)\tilde{w}(s) + PL_3\tilde{v}(s) + v(0) \tag{6.5.67}$$

$$s\tilde{w}(s) = [\gamma^2 L_1 + \gamma(1-P)L_2 + (1-P)L_3]\tilde{w}(s) + \gamma L_2\tilde{v}(s) + (1-P)L_3\tilde{v}(s) \tag{6.5.68}$$

从而

$$s\tilde{v}(s) = PL_3\tilde{v}(s) + P(\gamma L_2 + L_3)\lfloor s - \gamma^2 L_1 - \gamma(1-P)L_2$$

$$+(1-P)L_3]^{-1}[\gamma L_2+(1-P)L_3]\tilde{v}(s)+v(0) \tag{6.5.69}$$

由于 L_2 被 γ 所乘而 L_3 没有，所以对大值 γ 的极限过程来说

$$s\tilde{v}(s)\simeq(PL_3-PL_2L_1^{-1}L_2)\tilde{v}(s)+v(0) \tag{6.5.70}$$

在这种情况下我们将不假设可归范自相关函数到一常数。$-PL_2L_1^{-1}L_2$ 这项给出了

$$\frac{\partial}{\partial x}b(x,\tau)\frac{\partial}{\partial x}b(x,\tau)\int_0^\infty dt\langle\alpha_\tau(t)\alpha_\tau(0)\rangle \tag{6.5.71}$$

其中 $\alpha_\tau(t)$ 表示一随机变量，其福克-普朗克方程为

$$\frac{\partial p}{\partial t}=\left[\frac{\partial}{\partial\alpha}A(\alpha,\tau)+\frac{1}{2}\frac{\partial^2}{\partial x^2}B(\alpha,\tau)\right]p \tag{6.5.72}$$

于是，取 $\gamma\to\infty$ 的极限就有效地使 α 的随机运动比因 A 与 B 与时间相依而使 α 也与时间相依所带来的运动不知快多少。定义

$$D(\tau)=\int_0^\infty dt\langle\alpha_\tau(t)\alpha_\tau(0)\rangle \tag{6.5.73}$$

则经前面那样消去 τ 后就会得到

$$\frac{\partial p}{\partial t}=\left[-\frac{\partial}{\partial x}a(x,t)+D(t)\frac{\partial}{\partial x}b(x,t)\frac{\partial}{\partial x}b(x,t)\right]p \tag{6.5.74}$$

6.6 快变变量的绝热消去：一般情况

我们现在要考察两个变量 x 和 α 的一般情况，此时，这两个变量是以一种相互影响的方式耦合在一起的。这是个类似于导出具非零 $V'(x)$ 的斯莫路乔夫斯基方程的问题，而上一节是 $V'(x)$ 为零时相同问题的推广。

这种类型的最一般化问题是如此复杂和难以控制，以至到了不可理解的地步。为了介绍涉及的概念，我们首先考虑一个线性的化学系统，然后再展开一般性理论。

6.6.1 例子：短寿命中间化合物的消去

我们考察化学反应系统

$$X \underset{k}{\overset{\gamma}{\rightleftharpoons}} Y \underset{\gamma}{\overset{k}{\rightleftharpoons}} A \tag{6.6.1}$$

这个例子，其中X与Y是化学物种，其数量是变化的，而A则用某种方法给定了数量．此系统的确定性速率方程是

$$\frac{dx}{dt} = -x + \gamma y \tag{6.6.2a}$$

$$\frac{dy}{dt} = -2\gamma y + x + a \tag{6.6.2b}$$

在这里，x, y, a是X, Y, A的浓度．为简单起见，速率常数选定为$k=1$.

物理的情况往往是Y为短寿命的中间态，它会按时间常数γ^{-1}衰变为X和A．因此在γ取大值的范围内Y的寿命变得更短，其浓度则成可忽略．这是令人感兴趣的情况．此时导致我们在(6.6.2b)中要解$\frac{dy}{dt}=0$而有

$$y = (x+a)/2\gamma \tag{6.6.3}$$

而且在(6.6.2a)中代入它后得到

$$\frac{dx}{dt} = \frac{x}{2} - \frac{a}{2} \tag{6.6.4}$$

在随机的类同情况下，由于加在(6.6.2)上的白噪声是相关的，而且y的平稳分布又依赖于γ，相应的步骤复杂了．更确切地说，相应于(6.3.2)的随机微分方程通常选择为(见7.6.1节)

$$\begin{aligned} dx &= (-x+\gamma y)dt + eB_{11}dW_1(t) + eB_{12}dW_2(t) \\ dy &= (-2\gamma y + x + a)dt + eB_{21}dW_1(t) + eB_{22}dW_2(t) \end{aligned} \tag{6.6.5}$$

其中矩阵B满足

$$BB^T = \begin{bmatrix} 2a & -2a \\ -2a & 4a \end{bmatrix} \tag{6.6.6}$$

这里的e是个参数，实质上它是反应系统体积倒数的平方根，而且通常取小值，虽然下面并不用到这个事实．

变量y的均值由(6.6.3)所给出，而且在取极限过程中逐渐消失为零．我们希望消去这个变量y．这只有应用我们已经开发出

的一种观念才有可能做到，这就是要看我们想消除的变量在与 γ 无关的平稳状态时是否具有一分布函数。于是，我们必须定义一个是 y 与 x 的函数的新变量，它具有这种所希望的性质。

相应于(6.6.5)的福克-普朗克方程是

$$\frac{\partial p}{\partial t}=\left[\frac{\partial}{\partial x}(x-\gamma y)+\frac{\partial^2}{\partial x^2}\varepsilon^2 a-2\frac{\partial^2}{\partial x\partial y}\varepsilon^2 a\right.$$
$$\left.+\frac{\partial}{\partial y}(2\gamma y-x-a)+2\varepsilon^2 a\frac{\partial^2}{\partial y^2}\right]p \qquad (6.6.7)$$

看来有理由用

$$z=2\gamma y-x-a \qquad (6.6.8)$$

来定义新变量 z，它与 y 和其平稳值之差成比例。于是正式地用

$$\begin{aligned} x_1&=x & x&=x_1\\ z&=2\gamma y-x-a & y&=(z+x_1+a)/2\gamma \end{aligned} \qquad (6.6.9)$$

来定义一对新变量 (x_1, z)，使得可以用

$$\begin{aligned} \frac{\partial}{\partial x}&=\frac{\partial}{\partial x_1}-\frac{\partial}{\partial z}\\ \frac{\partial}{\partial y}&=2\gamma\frac{\partial}{\partial z} \end{aligned} \qquad (6.6.10)$$

来变换其福克-普朗克方程从而得到

$$\frac{\partial p}{\partial t}=\left[\frac{\partial}{\partial x_1}\left[\frac{x_1-a}{2}-\frac{z}{2}\right]+\varepsilon^2 a\frac{\partial^2}{\partial x_1^2}\right.$$
$$+\frac{\partial^2}{\partial x_1\partial z}(-2\varepsilon^2 a-4\gamma\varepsilon^2 a)+\frac{\partial}{\partial z}\left(2\gamma z-\frac{x-a}{2}+\frac{z}{2}\right)$$
$$\left.+\frac{\partial^2}{\partial z^2}(8\varepsilon^2\gamma^2 a+\varepsilon^2 a+4\gamma\varepsilon^2 a)\right]p \qquad (6.6.11)$$

取 $\gamma\to\infty$ 的极限并未给出 z 的福克-普朗克算子，只给出了它与一给定了的算子成比例。我们看到 z 的漂移项和扩散项分别与 γ 及 γ^2 成比例。

不过作代换

$$\alpha=z\gamma^{-\frac{1}{2}} \qquad (6.6.12)$$

则改变了这种情况。用 α 来表示时，漂移系数与扩散系数变成都与 γ 成比例，并且可看到(现在用 x 代替 x_1)

$$\frac{\partial p}{\partial t} = [\gamma L_1 + \gamma^{\frac{1}{2}} L_2(\gamma) + L_3] p \tag{6.6.13}$$

其中

$$L_1 = 2\frac{\partial}{\partial \alpha}\alpha + 8\varepsilon^2 a \frac{\partial^2}{\partial \alpha^2} \tag{6.6.14}$$

$$\begin{aligned} L_2(\gamma) = {} & \frac{\partial}{\partial x}\left[-4\varepsilon^2 a \frac{\partial}{\partial \alpha} - \frac{1}{2}\alpha\right] + 4\gamma^{-\frac{1}{2}}\varepsilon^2 a \frac{\partial^2}{\partial \alpha^2} \\ & - \gamma^{-1}\frac{\partial}{\partial \alpha}\left[\frac{x-a}{2}\right] + \left\{-\gamma^{-\frac{1}{2}}\frac{\partial}{\partial \alpha}\alpha - \gamma^{-1} 2\varepsilon^2 a \frac{\partial^2}{\partial x \partial \alpha}\right. \\ & \left. + \gamma^{-3/2}\varepsilon^2 a \frac{\partial^2}{\partial \alpha^2}\right\} \end{aligned} \tag{6.6.15}$$

$$L_3 = \frac{\partial}{\partial x}\left(\frac{x-a}{2}\right) + \varepsilon^2 a \frac{\partial^2}{\partial x^2} \tag{6.6.16}$$

注意(6.6.15)第一行的第一项给出了 $\gamma \to \infty$ 时 $L_2(\gamma)$ 的极限值。L_2 的唯一重要性质在于 $PL_2P = 0$。照常定义 P 为

$$Pf(x, \alpha) = p_s(\alpha) \int d\alpha' f(x, \alpha') \tag{6.6.17}$$

其中 $p_s(\alpha)$ 是 L_1 的平稳解，则可看到对任何以 $\frac{\partial}{\partial \alpha}$ 开始的算子[例如 $L_2(\gamma)$ 中与 γ 有关的部分]，只要它可丢掉边界项，就有

$$P\frac{\partial f}{\partial \alpha}(x, \alpha) = p_s(\alpha) \int d\alpha' \frac{\partial f}{\partial \alpha'}(x, \alpha') = 0 \tag{6.6.18}$$

因此 $L_2(\gamma)$ 中与 γ 有关的部分满足 $PL_2P = 0$。不仅如此，由(6.6.14)显然 $\langle \alpha \rangle_s = 0$，所以我们发现其独立于 γ 的那部分也满足这个条件。于是有

$$PL_2(\gamma)P = 0 \tag{6.6.19}$$

尽管这样，但仍值得说明一下：$L_2(\gamma)$ 中与 γ 有关的部分包含了看来是更适用于 L_1 的那些项，这就是未含 x 的任何导数的那些项。不过，把这些项移到 L_2 中就把 L_1 变成与 γ 无关了。这样一来，P 就与 γ 无关而且极限更为明显。

现在步骤十分简捷。照常定义

$$\begin{aligned} P\tilde{p}(s) &= \tilde{v}(s) \\ (1-P)\tilde{p}(s) &= \tilde{w}(s) \end{aligned} \tag{6.6.20}$$

并且假设 $w(0)=0$ 后，就有

$$s\tilde{v}(s)=P[\gamma L_1+\gamma^{1/2}L_2(\gamma)+L_3][\tilde{v}(s)+\tilde{w}(s)]+v(0) \tag{6.6.21}$$

再利用

$$\begin{aligned} &PL_1=L_1P=0\\ &PL_2P=0\\ &PL_3=L_3P \end{aligned} \tag{6.6.22}$$

就得到

$$s\tilde{v}(s)=P\gamma^{1/2}L_2(\gamma)\tilde{w}(s)+L_3\tilde{v}(s)+v(0) \tag{6.6.23}$$

以及类似地

$$s\tilde{w}(s)=[\gamma L_1+\gamma^{1/2}(1-P)L_2(\gamma)+L_3]\tilde{w}(s)+\gamma^{1/2}L_2(\gamma)\tilde{v}(s) \tag{6.6.24}$$

从而

$$\begin{aligned} s\tilde{v}(s)=\{&L_3+\gamma PL_2(\gamma)[s-\gamma L_1-\gamma^{1/2}(1-P)L_2(\gamma)\\ &-L_3]^{-1}L_2\}\tilde{v}(s)+v(0) \end{aligned} \tag{6.6.25}$$

现在取 γ 的大值极限，就得到

$$\boxed{s\tilde{v}(s)\simeq(L_3-PL_2L_1^{-1}L_2)\tilde{v}(s)+v(0)} \tag{6.6.26}$$

其中

$$L_2=\lim_{\gamma\to\infty}L_2(\gamma)=\frac{\partial}{\partial x}\left(-4e^2a\frac{\partial}{\partial x}-\frac{1}{2}\alpha\right) \tag{6.6.27}$$

方程(6.6.26)与(6.5.28)形式完全相同，而且实际上由(6.6.7)正式导出它的过程也几乎与后者的相同。但是由于这里存在着含 $\frac{\partial}{\partial\alpha}$ 的项而使 $PL_2L_1^{-1}L_2$ 的计算稍有不同。首先，注意到由于 $P\frac{\partial}{\partial\alpha}=0$，我们可写出

$$\begin{aligned} -PL_2L_1^{-1}L_2v=&-p_s(\alpha)\int d\alpha'\left(-\frac{1}{2}\alpha'\frac{\partial}{\partial x}\right)L_1^{-1}\\ &\cdot\frac{\partial}{\partial x}\left(-4e^2a\frac{\partial}{\partial\alpha'}-\frac{1}{2}\alpha'\right)p_s(\alpha')p(x) \end{aligned} \tag{6.6.28}$$

并且由 $p_s(\alpha)$ 的定义知它满足 $L_1 p_s(\alpha)=0$，所以由(6.6.14)看出

$$\frac{\partial}{\partial\alpha}p_s(\alpha) = -\alpha p_s(\alpha)/4\varepsilon^2 a \tag{6.6.29}$$

而使

$$-PL_2L_1^{-1}L_2 v = \frac{1}{4}p_s(\alpha)\frac{\partial^2}{\partial\alpha^2}\int d\alpha'\alpha' L_1^{-1}\alpha' p_s(\alpha')\hat{p}(x) \tag{6.6.30}$$

$$= -\frac{1}{4}p_s(\alpha)\frac{\partial^2\hat{p}(x)}{\partial x^2}\int_0^\infty dt\langle\alpha(t)w(0)\rangle, \tag{6.6.31}$$

其中我们用了6.5节中给出的推理以便用相关函数来表示解答。这里的 L_1 是 $k=2$, $D=16\varepsilon^2 a$ 的奥恩斯坦-乌伦贝克过程（见3.8.4节）的生成算子，所以由(3.8.2)式，

$$-PL_2L_1^{-1}L_2 v = -\frac{1}{4}p_s(\alpha)\frac{\partial^2\hat{p}(x)}{\partial x^2}4\varepsilon^2 a\int_0^\infty dt e^{-2t}$$

$$= -\frac{1}{2}\varepsilon^2 a\frac{\partial^2\hat{p}(x)}{\partial x^2}p_s(\alpha) \tag{6.6.32}$$

因此，由(6.6.26)知其有效的福克-普朗克方程是

$$\boxed{\frac{\partial\hat{p}}{\partial t} = \frac{\partial}{\partial x}\frac{x-a}{2}\hat{p}(x) + \frac{1}{2}\varepsilon^2 a\frac{\partial^2\hat{p}}{\partial x^2}} \tag{6.6.33}$$

注解

(i) 这完全是个从反应

$$X \underset{k/2}{\overset{k/2}{\rightleftharpoons}} A \tag{6.6.34}$$

($k=1$)中预料得到的方程(见7.5.3节)。它之所以是可预期的，是因为一般原理已告诉我们浓度涨落的平稳方差是由

$$\mathrm{var}\{x(t)\}_s = \varepsilon^2\langle x\rangle, \tag{6.6.35}$$

来给定的。

(ii) 请注意绝热消去法的净效果是简化了 $\frac{\partial^2}{\partial x^2}$ 的系数。这是原来方程中 x 中的与 y 中的噪声项相关的结果。

(iii) 这个结果不同于通常的绝热消去法，在那里，被消去的变量中的噪声项是重要的，而这里并非如此．以后对它们还要作简短讨论．

6.6.2 哈肯模型中的绝热消去法

哈肯为说明绝热消去引入了一个简单的模型[文献6.1，第7.2节]．这个模型的确定性版本是一对耦合方程，它们可表示为

$$\dot{x} = -\varepsilon x - axa \tag{6.6.36}$$

$$\dot{\alpha} = -\kappa\alpha + bx^2 \tag{6.6.37}$$

我们假定 κ 充分大时可和前面一样，用(6.6.37)中以 x 表示的平稳解来代替 α，以得到

$$\alpha = \frac{b}{\kappa}x^2 \tag{6.6.38}$$

$$\dot{x} = -\varepsilon x - \frac{ab}{\kappa}x^3 \tag{6.6.39}$$

模型的真实目的是要得到(6.6.39)右边的立方形式．

在转换到随机系统的情况时，我们发现有各种可供选择的可能性．使绝热消去有效的通常条件是

$$\varepsilon \ll \kappa \tag{6.6.40}$$

在一个随机版本中，所有其他的参量都同样在起作用，而且实际上至少可以用三种具有不同特性的答案方式来实现条件(6.6.40)．

让我们把(6.6.36, 37)的随机化版本记为

$$\begin{aligned} \mathrm{d}x &= -(\varepsilon x + axa)\mathrm{d}t + C\mathrm{d}W_1(t) \\ \mathrm{d}\alpha &= (-\kappa\alpha + bx^2)\mathrm{d}t + D\mathrm{d}W_2(t) \end{aligned} \tag{6.6.41}$$

并且为简单起见，这里假定 C 与 D 是常数，且 $W_1(t)$ 与 $W_2(t)$ 彼此独立．

福克-普朗克方程是

$$\frac{\partial p}{\partial t} = \left[\frac{\partial}{\partial x}(\varepsilon x + ax(\alpha) + \frac{1}{2}C^2\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial\alpha}(\kappa\alpha - bx^2) + \frac{1}{2}D^2\frac{\partial^2}{\partial\alpha^2}\right]p \tag{6.6.42}$$

我们希望消去 α。为此方便的做法是用

$$\beta = \alpha - \frac{b}{\kappa}x^2 \tag{6.6.43}$$

定义一个新变量 β，使得对给定的 x 来说 β 具有零均值。借这个新变量我们可把福克-普朗克方程写为

$$\frac{\partial p}{\partial t} = (L_1^0 + L_2^0 + L_3^0)p \tag{6.6.44}$$

$$L_1^0 = \frac{\partial}{\partial \beta}\kappa\beta + \frac{D^2}{2}\frac{\partial^2}{\partial \beta^2} \tag{6.6.45}$$

$$\begin{aligned} L_2^0 = {} & \frac{\partial}{\partial \beta}a\beta x - \frac{2bx}{\kappa}\frac{\partial}{\partial \beta}\left(\varepsilon x - \frac{ab}{\kappa}x^3 + ax\beta\right) \\ & - C^2\left(\frac{bx}{\kappa}\frac{\partial^2}{\partial x \partial \beta} + \frac{\partial^2}{\partial x \partial \beta}\frac{bx}{\kappa}\right) + \frac{2bx^2c^2}{\kappa^2}\frac{\partial^2}{\partial \beta^2} \end{aligned} \tag{6.6.46}$$

$$L_3^0 = \left[\frac{\partial}{\partial x}\left(\varepsilon x + \frac{ab}{\kappa}x^3\right) + \frac{c^2}{2}\frac{\partial^2}{\partial x^2}\right] \tag{6.6.47}$$

在这些变量的表示下，$\varepsilon \to 0$ 的极限过程并不有趣，因为只不过得到了 $\varepsilon = 0$ 时的相同系统。由于 L_1 未被一个大值参量去乘，不可能有消去发生。

为了使极限过程 $\varepsilon \to 0$ 确凿地意味着(6.6.39)是一有效的极限形式，必须存在一个 A 使得

$$\frac{ab}{\kappa} = \varepsilon A \quad \text{当 } \varepsilon \to 0 \tag{6.6.48}$$

为了使这个极限是确定性地可识别的，它应该不被噪声所掩没掉，所以我们也必须有

$$\frac{C^2}{2} = \varepsilon\beta \quad \text{当 } \varepsilon \to 0 \tag{6.6.49}$$

这意味着当 $\varepsilon \to 0$ 时

$$L_3^0 \longrightarrow \varepsilon\left[\frac{\partial}{\partial x}(x + Ax^3) + B\frac{\partial^2}{\partial x^2}\right] \tag{6.6.50}$$

但是对 L_2^0 来说却有两种不同可能。为使 L_1^0 独立于 ε，必须 κ 独立于 ε，这是合理的。于是极限(6.6.48)必然是靠乘积 ab 是与 ε 成比例而得到的。下面来考察几种可能性。

a) 无噪声支配：a 与 ε 成比例的情况

我们假设可写出

$$a = \varepsilon\tilde{a} \tag{6.6.51}$$

于是看到 L_1^0 与 ε 无关而 L_2^0 及 L_3^0 与 ε 成比例。若用

$$\tau = \varepsilon t \tag{6.6.52}$$

重新给出时间标度，则

$$\frac{\partial p}{\partial \tau} = \left(\frac{1}{\varepsilon}L_1 + L_2 + L_3\right)p \tag{6.6.53}$$

其中

$$\begin{aligned} L_1 &= L_1^0 \\ L_2 &= L_2^0/\varepsilon \\ L_3 &= L_3^0/\varepsilon \end{aligned} \tag{6.6.54}$$

由于 $\varepsilon \to 0$ 时 L_2 并不趋于无穷大，显然，用通常的消去程序到最低阶可写出

$$\boxed{\frac{\partial \hat{p}}{\partial t} = L_3\hat{p} = \left[\frac{\partial}{\partial x}(x + Ax^3) + B\frac{\partial^2}{\partial x^2}\right]\hat{p}} \tag{6.6.55}$$

这恰好相当于在 x 的方程中忽略 α 的涨落而只取其确定性之值后绝热地消去 α。我把它称为“无噪声支配”（用哈肯所用的术语来说），因为 α 被 x 支配着，并且在 x 的方程中没有噪声的作用，这是通常的支配方式，正如哈肯所考察过的。

b) 有噪声支配，a 与 $\varepsilon^{\frac{1}{2}}$ 成比例的情况

若换种方式，假设 a 与 b 都与 $\varepsilon^{\frac{1}{2}}$ 成比例，则可写出

$$\begin{aligned} a &= \tilde{a}\varepsilon^{\frac{1}{2}} \\ b &= \tilde{b}\varepsilon^{\frac{1}{2}} \end{aligned} \tag{6.6.56}$$

其中

$$\tilde{a}\tilde{b} = \kappa A \tag{6.6.57}$$

此时 L_1^0 仍为常值，L_3^0 与 ε 成比例并且

$$L_2^0 = \varepsilon^{\frac{1}{2}}L_2 + \varepsilon \text{ 的高阶项} \tag{6.6.58}$$

其中

$$L_2 = \tilde{a}\beta \frac{\partial}{\partial x} x \tag{6.6.59}$$

这样一来，极限方程就是

$$\frac{\partial \hat{p}}{\partial t} = (L_3 - PL_2L_1^{-1}L_2)\,\hat{p} \tag{6.6.60}$$

$PL_2L_1^{-1}L_2$ 这项如同前面那样可解出；我们发现

$$-PL_2L_1^{-1}L_2 = \tilde{a}^2 \frac{D^2}{2\kappa^2} \frac{\partial}{\partial x} x \frac{\partial}{\partial x} x \tag{6.6.62}$$

所以

$$\boxed{\frac{\partial \hat{p}}{\partial \tau} = \left\{ \frac{\partial}{\partial x} \left[x\left(1 - \frac{\tilde{a}^2 D^2}{2\kappa^2}\right) + Ax^3 \right] + \frac{\partial^2}{\partial x^2} \left[B + \frac{\tilde{a}^2 D^2 x^2}{2\kappa^2} \right] \right\} \hat{p}} \tag{6.6.62}$$

我称它为“有噪声支配”，因为支配在最后方程中的存在是靠加上噪声(并且影响到漂移项，虽然这仅在伊藤形式的表示下才显示出来，而在斯特拉托诺维奇形式下似乎没有另外的漂移项)使人感觉到的。

c) 一般的情况

因为我们假设 $ab \propto \varepsilon$，可以看到(6.6.46)中的第二项总是与 $\varepsilon^p (p>1)$ 成比例的，所以它是可忽略的(只要 b 是有界的话)。于是，在 L_2^0 中只有第一项是重要的。这就导致若

$$a = \varepsilon^r \tilde{a}$$

则有下列几种可能性：

$r > \frac{1}{2}$ 时，L_2 没有作用，极限方程为(6.6.55)；

$r = \frac{1}{2}$ 时，极限方程是(6.6.62)，是有噪声支配；

$r < \frac{1}{2}$ 时，$PL_2L_1^{-1}L_2$ 这项成了 ε^{2r-1} 阶次的而 $\to \infty$，并且是支配项。

对于 $r < \frac{1}{2}$，方程渐近为

$$\frac{\partial \hat{p}}{\partial \tau} = \varepsilon^{2r-1}\tilde{a}^2 \frac{D^2}{2K^2}\left(\frac{\partial}{\partial x} x \frac{\partial}{\partial x} x\right)\hat{p} \tag{6.6.63}$$

这是些十分明显的差别，它们全可合并到一个公式中，即总的有

$$\boxed{\frac{\partial \hat{p}}{\partial \tau} = \left[\frac{\partial}{\partial x}(x + Ax^3) + \frac{\partial^2}{\partial x^2} B + \varepsilon^{2r-1}\tilde{a}^2 \frac{D^2}{2K^2} \frac{\partial}{\partial x} x \frac{\partial}{\partial x} x\right]\hat{p}} \tag{6.6.64}$$

在应用绝热消去技巧时，一般说来，我们当然应该特别注意保证所取系统中对所有常数中的那些小参数的正确相依关系。

6.6.3 快变变量的绝热消去：非线性情况

这里我们要考察一对变元 x 和 α 的一般情况，其中这两个变元是相互影响地耦合在一起的，尽管 α 的时间尺度比 x 的快得多。

让我们考虑一个由一对随机微分方程所刻画的系统：

$$\mathrm{d}x = [a(x) + b(x)\alpha]\mathrm{d}t + c(x)\mathrm{d}W_1(t) \tag{6.6.65}$$

$$\mathrm{d}\alpha = \gamma^2[A(\alpha) - f(x)]\mathrm{d}t + \gamma\sqrt{2B(\alpha)}\,\mathrm{d}W_2(t) \tag{6.6.66}$$

若简单地按 6.4 节的演绎去做，我们立刻会碰到麻烦。因为在极限下，由于假设对大值的 α(6.6.66)总使得 $\mathrm{d}\alpha/\mathrm{d}t = 0$，而应该令

$$A(\alpha) - f(t) = -\sqrt{\frac{2B(\alpha)}{\gamma^2}}\,\frac{\mathrm{d}W_2(t)}{\mathrm{d}t} \tag{6.6.67}$$

但这样一来，用(6.6.67)解出 α 为 x 的函数时，一般说，得到的是 x 的某种复杂的非线性函数，而且，$\mathrm{d}W_2(t)/\mathrm{d}t$ 的含义也是不可理解的。不过，只要 $B(\alpha)$ 是零，则可定义 $u_0(x)$ 为

$$A[u_0(x)] = f(x) \tag{6.6.68}$$

并代入(6.6.65)而得到

$$\mathrm{d}x = [a(x) + b(x)u_0(x)]\mathrm{d}t + c(x)\mathrm{d}W_1(t) \tag{6.6.69}$$

下面我们要设计更好些的程序，以上面的方法为基础，而且也可考虑到(6.6.66)中涨落的影响。

等价于(6.6.65, 66)的福克-普朗克方程是

$$\frac{\partial p}{\partial t} = (\gamma^2 L_1 + L_2 + L_3)p \tag{6.6.70}$$

其中

$$L_1 = \frac{\partial}{\partial \alpha}[f(x) - A(\alpha)] + \frac{\partial^2}{\partial \alpha^2} B(\alpha) \tag{6.6.71}$$

而 L_2 及 L_3 要选择得使 $PL_2P = 0$ 得以满足，这放在后面再确定。第一步我们照常选 P 是到 L_1 的零空间上的投影算子。我们用 $p_x(\alpha)$ 记平稳解，也就是方程

$$L_1 p_x(\alpha) = 0 \tag{6.6.72}$$

的解，因为 L_1 明显地通过 $f(x)$ 来依赖于 x，这就使得 $p_x(\alpha)$ 明显地依赖于 x。投影算子 P 是用对任何函数 $F(x, \alpha)$

$$(PF)(x, \alpha) = p_x(\alpha) \int d\alpha' F(x, \alpha') \tag{6.6.73}$$

来定义的。

现在定义函数 $u(x)$ 为

$$u(x) = \int d\alpha \alpha p_x(\alpha) \equiv \langle \alpha \rangle_x \tag{6.6.74}$$

则可定义

$$L_2 = -\frac{\partial}{\partial x}\{b(x)[\alpha - u(x)]\} \tag{6.6.75}$$

$$L_3 = -\frac{\partial}{\partial x}[a(x) + b(x)u(x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)]^2 \tag{6.6.76}$$

以致它们加起来时 $\left\{\frac{\partial}{\partial x}\right\} b(x)u(x)$ 项消掉了。这样一来，(6.6.70) 就是相应于 (6.6.65, 66) 的合适的福克-普朗克方程。

因为 $\int \alpha p_x(\alpha) d\alpha = u(x)$，现在可得

$$PL_2PF = -p_x(\alpha) \int d\alpha' \frac{\partial}{\partial x}\{b(x)[\alpha' - u(x)]\} p_x(\alpha') \int d\alpha'' F(x, \alpha'') = 0 \tag{6.6.77}$$

$$PL_1 = L_1P = 0 \tag{6.6.78}$$

是自然成立的，但是

$$PL_3 \neq L_3P \tag{6.6.79}$$

现在来完成通常的步骤，照例记

$$\begin{aligned} P\tilde{p}(s) &= \tilde{v}(s) \\ (I-P)\tilde{p}(s) &= \tilde{w}(s) \end{aligned} \tag{6.6.80}$$

并假设 $w(0)=0$，则看到

$$\begin{aligned} s\tilde{v}(s) &= P(L_2+L_3)\tilde{w}(s) + PL_3\tilde{v}(s) + v(0) \\ s\tilde{w}(s) &= [\gamma^2 L_1 + (1-P)L_2 + (1-P)L_3]\tilde{w}(s) + L_2\tilde{v}(s) \\ &\quad + (1-P)L_3\tilde{v}(s) \end{aligned} \tag{6.6.81}$$

从而

$$\begin{aligned} s\tilde{v}(s) = PL_3\tilde{v}(s) &+ P(L_2+L_3)[s-\gamma^2 L_1 - (1-P)L_2 \\ &- (1-P)L_3]^{-1}[L_2 + (1-P)L_3]\tilde{v}(s) + v(0) \end{aligned} \tag{6.6.82}$$

到二阶为止，则有

$$\boxed{\begin{aligned} s\tilde{v}(s) \simeq \{&(PL_3 - \gamma^{-2}P(L_2+L_3)L_1^{-1}[L_2 \\ &+ (1-P)L_3]\}\tilde{v}(s) + v(0) \end{aligned}} \tag{6.6.83}$$

$PL_3\tilde{v}(s)$ 这项是最重要的，并且产生确定性的绝热消去结果。记

$$v(t) = p_x(\alpha)\hat{p}(x)$$

则有

$$\begin{aligned} PL_3 v(t) = p_x(\alpha) \int d\alpha' \Big\{ &- \frac{\partial}{\partial x}[a(x) + b(x)u(x)] \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)^2] \Big\} p_x(\alpha')\hat{p}(x) \end{aligned} \tag{6.6.84}$$

并且因为 $\int d\alpha p_x(\alpha) = 1$，

$$\begin{aligned} PL_3 v(t) = p_x(\alpha) \Big\{ &- \frac{\partial}{\partial x}[a(x) + b(x)u(x)] \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)]^2 \Big\}\hat{p}(x) \end{aligned} \tag{6.6.85}$$

从而最低阶微分方程是

$$\begin{aligned} \frac{\partial \hat{p}(x)}{\partial t} = &- \frac{\partial}{\partial x}[a(x) + b(x)u(x)]\hat{p}(x) \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2}c(x)^2\hat{p}(x) \end{aligned} \tag{6.6.86}$$

它等价于随机微分方程

$$\boxed{\mathrm{d}x = [a(x) + b(x)u(x)]\mathrm{d}t + c(x)\mathrm{d}W(t)} \tag{6.6.87}$$

到此为止，运动方程包含着源出于 α 的方程的非涨落项。但是它完全忽略所有涨落的结果是由(6.6.69)所给定的，它十分类似于(6.6.87)，只不过用 $u_0(x)$ 代替了 $u(x)$。我们还同时期待的是在平稳状态下 α 的平均值应该类似于 $u_0(x)$，它不与之相同，而且只有当噪声项 $B(\alpha)$ 小的时候这种类似性才很强。

二阶修正

把(6.6.83)计算到 γ^{-2} 阶是可能的。初看上去，在 L_3 中二阶导数的存在似乎表明计算到 γ^{-2} 阶就会出现四阶导数，因为 L_3 出现了两次。但是我们可以证明四阶导数不存在。

考虑表示式

$$P(L_2+L_3)L_1^{-1}[L_2+(1-P)L_3]\tilde{v}(s) \tag{6.6.88}$$

我们知道

$$\text{i)}\ \ P\tilde{v}(s) = \tilde{v}(s)$$

$$\text{ii)}\ \ (1-P)L_2P\tilde{v}(s) = L_2P\tilde{v}(s) \tag{6.6.89}$$

这里用了 $PL_2P = 0$。

于是(6.6.88)成了

$$\begin{aligned} &P(L_2+L_3)L_1^{-1}(1-P)(L_2+L_3)P\tilde{v}(s) \\ &= P\{PL_2+[P,\ L_3]+L_3P\}(1-P)L_1^{-1}(1-P)\{L_2P \\ &\quad +[L_3,\ P]+PL_3\}\tilde{v}(s) \end{aligned} \tag{6.6.90}$$

其中交换子 $[A, B]$ 定义为

$$[A,\ B] = AB - BA \tag{6.6.91}$$

我们已注意到 L_1^{-1} 与 $(1-P)$ 可交换而且在(6.6.90)中可利用 $(1-P)^2 = (1-P)$ 以便在 L_1^{-1} 之前面插入另一个 $(1-P)$。因为 $P^2 = P$，我们也已在整个表示式的前面加上了另一个 P。现在用一下

$$P(1-P) = (1-P)P = 0$$

(6.6.90)就成了

$$P\{PL_2+[P,\ L_3]\}L_1^{-1}(1-P)\{L_2+[L_3,\ P]\}\tilde{v}(s) \tag{6.6.92}$$

下面我们要计算$[P, L_3]$：

$$(PL_3f)(x, \alpha) = p_x(\alpha)\left\{-\frac{\partial}{\partial x}[a(x)+b(x)u(x)]\right.$$
$$\left.+\frac{1}{2}\frac{\partial^2}{\partial x^2}[a(x)^2]\right\}\int f(x, \alpha')\mathrm{d}\alpha' \quad (6.6.93)$$

而且

$$(L_3Pf)(x, \alpha) = \left\{-\frac{\partial}{\partial x}[a(x)+b(x)u(x)]\right.$$
$$\left.+\frac{1}{2}\frac{\partial^2}{\partial x^2}[a(x)^2]\right\}p_x(\alpha)\int \mathrm{d}\alpha' f(x, \alpha') \quad (6.6.94)$$

两式相减，并定义

$$r_x(\alpha) = \frac{\partial p_x(\alpha)}{\partial x}\Big/ p_x(\alpha) \quad (6.6.95)$$

$$s_x(\alpha) = \frac{\partial^2 p_x(\alpha)}{\partial x^2}\Big/ p_x(\alpha) \quad (6.6.96)$$

就得到

$$([P, L_3]f)(x, \alpha) = r_x(\alpha)[a(x)+b(x)u(x)]Pf(x, \alpha)$$
$$-\frac{1}{2}s_x(\alpha)c(x)^2Pf(x, \alpha)$$
$$-r_x(\alpha)P\frac{\partial}{\partial x}[c(x)^2f(x, \alpha)] \quad (6.6.97)$$

因为我们感兴趣的情况是$f(x, \alpha)$为v，即

$$f(x, \alpha) = p_x(\alpha)\hat{p}(x) \quad (6.6.98)$$

最后一项还可以进一步简化。此时

$$P\frac{\partial}{\partial x}c(x)^2p_x(\alpha)\hat{p}(x) \quad (6.6.99)$$

$$= p_x(\alpha)\frac{\partial}{\partial x}c(x)^2\int \mathrm{d}\alpha' p_x(\alpha')\hat{p}(x) \quad (6.6.100)$$

$$= p_x(\alpha)\frac{\partial}{\partial x}c(x)^2\hat{p}(x) \quad (6.6.101)$$

不仅如此，我们还可证明

$$P[P, L_3] = 0 \quad (6.6.102)$$

这是因为由

$$\int \mathrm{d}\alpha p_x(\alpha) = 1 \tag{6.6.103}$$

得到

$$\int \mathrm{d}\alpha r_x(\alpha) p_x(\alpha) = \int \mathrm{d}\alpha s_x(\alpha) p_x(\alpha) = 0 \tag{6.6.104}$$

而它对说明(6.6.102)是充分的。因此我们可把 (6.6.92) 改写为

$$PL_2L_1^{-1}\{L_2 + [L_3, P]\}\tilde{v}(s) \tag{6.6.105}$$

的形式，并得到

$$\begin{aligned}[L_3, P]\tilde{v}(s) = -p_x(\alpha)\Big\{ & r_x(\alpha)[a(x) + b(x)u(x)] \\ & - \frac{1}{2}s_x(\alpha)c(x)^2\Big\}\hat{p}(x) \\ & + p_x(\alpha)r_x(\alpha)\frac{\partial}{\partial x}[c(x)^2\hat{p}(x)]\end{aligned} \tag{6.6.106}$$

6.6.4 任意非线性耦合的例子

我们考察一对方程

$$\begin{aligned}\mathrm{d}x &= \gamma b(x)\mathrm{d}t \\ \mathrm{d}\alpha &= -\gamma^2 A(x, \alpha, \gamma)\mathrm{d}t + \gamma\sqrt{2B(x, \alpha, \gamma)}\,\mathrm{d}W(t)\end{aligned} \tag{6.6.107}$$

并假设存在下列极限和渐近展开

$$\begin{aligned}A(x, \alpha, \gamma) &\sim \sum_{n=0}^{\infty} A_n(x, \alpha)\gamma^{-n} \\ B(x, \alpha, \gamma) &\sim \sum_{n=0}^{\infty} B_n(x, \alpha)\gamma^{-n}\end{aligned} \tag{6.6.108}$$

这些展开式意味着在给定 x 后 α 具有一个渐近平稳分布，它由

$$p_s(\alpha, x) = \lim_{\gamma\to\infty} p_s(\alpha, x, \gamma) \tag{6.6.109}$$

$$p_s(\alpha, x) \propto B_0(x, \alpha)^{-1}\exp\left\{\int \mathrm{d}\alpha[A_0(x, \alpha)/B_0(x, \alpha)]\right\} \tag{6.6.110}$$

给出。我们还假设 $A_0(x, \alpha)$ 与 $B_0(x, \alpha)$ 满足

$$\langle\alpha(x)\rangle_s = \int \mathrm{d}\alpha\alpha p_s(\alpha, x) = 0 \tag{6.6.111}$$

从而由(6.6.108)导致对有限的 γ

$$\langle \alpha(x,\ \gamma)\rangle_s = \int \mathrm{d}\alpha \alpha p_s(\alpha,\ x,\ \gamma) \sim \alpha_0(x)\gamma^{-1} \tag{6.6.112}$$

其中 $\alpha_0(x)$ 可由(6.6.108)确定。

现在定义新变量

$$\beta = \alpha - \frac{1}{\gamma}\alpha_0(x) \tag{6.6.113}$$

$$x_1 = x$$

用这些新变量来表示福克-普朗克算子，然后把 x_1 改回到 x，它就成了(通常，雅可比是常数)

$$\begin{aligned} L = &-\frac{\partial}{\partial x}\alpha_0(x)b(x) - \gamma\beta\frac{\partial}{\partial x}b(x) \\ &+\frac{1}{\gamma}\alpha_0'(x)b(x)\frac{\partial}{\partial\beta} + \frac{\partial}{\partial\beta}\beta\alpha_0'(x)b(x) \\ &+\gamma^2\left[\frac{\partial}{\partial\beta}A\left(\frac{\alpha_0(x)}{\gamma} + \beta,\ x,\ \gamma\right)\right. \\ &\left.+\frac{\partial^2}{\partial\beta^2}B\left(\frac{\alpha_0(x)}{\gamma} + \beta,\ x,\ \gamma\right)\right] \end{aligned} \tag{6.6.114}$$

再利用渐近展开(6.6.108)，它就可以写为

$$L = L_3 + \gamma L_2(\gamma) + \gamma^2 L_1 \tag{6.6.115}$$

其中

$$L_3 = -\frac{\partial}{\partial x}\alpha_0(x)b(x) \tag{6.6.116}$$

$$L_1 = \frac{\partial}{\partial\beta}A_0(\beta,\ x) + \frac{\partial^2}{\partial\beta^2}B_0(\beta,\ x) \tag{6.6.117}$$

$$L_2(\gamma) = L_2 + O(\gamma^{-1}) \tag{6.6.118}$$

$$\begin{aligned} L_2 = &-\beta\frac{\partial}{\partial x}b(x) - \frac{\partial}{\partial\beta}\left[\frac{\partial A_0(\beta,\ x)}{\partial\beta}\alpha_0(x) + A_1(\beta,\ x)\right] \\ &-\frac{\partial^2}{\partial\beta^2}\left[\frac{\partial B_0(\beta,\ x)}{\partial\beta}\alpha_0(x) + B_1(\beta,\ x)\right] \end{aligned} \tag{6.6.119}$$

我们注意到 L_3 与 L_1 不是可交换的，但是与 6.5.4 节一样，这并不影响极限结果，即

$$\frac{\partial p}{\partial t} = (L_3 - PL_2L_1^{-1}L_2)p \tag{6.6.120}$$

计算 $PL_2L_1^{-1}L_2$ 项的方法是明确的，但是并不容易。我们注意到含 $\frac{\partial}{\partial\beta}$ 的那些项在 P 作用之后都变成零了。按照 $p_s(\alpha, x)$ 的显式，我们可用

$$\frac{\partial}{\partial\beta}\left\{\left[\frac{\partial A_0(\beta, x)}{\partial\beta}\alpha_0(x)+A_1(\beta,x)\right]p_s(\beta, x)\right\}$$
$$+\left\{\frac{\partial^2}{\partial\beta^2}\left[\frac{\partial B_0(\beta, x)}{\partial\beta}\alpha_0(x)+B_1(\beta, x)\right]\right\}p_s(\beta, x)$$
$$=G(\beta, x)p_s(\beta, x) \tag{6.6.121}$$

来定义 $G(\beta, x)$，则会看到

$$PL_2L_1^{-1}L_2\hat{p}=\left[\frac{\partial}{\partial x}b(x)D(x)\frac{\partial}{\partial x}b(x)+\frac{\partial}{\partial x}b(x)E(x)\right]\hat{p} \tag{6.6.122}$$

其中

$$D(x)=\int_0^\infty dt\langle\beta(t), \beta(0)|x\rangle$$
$$E(x)=\int_0^\infty dt\langle\beta(t), G(\beta, x)|x\rangle \tag{6.6.123}$$

并且$\langle\cdots|x\rangle$表示关于 $p_s(\beta, x)$ 的平均。这是个比较强的绝热消去结果，在这里，任意非线性消去都可处理，而且还考虑到了所产生的有限噪声。计算比上一节要简单些，因为这里含 L_3 的那些项比含 L_2 的那些项阶次要低。

7. 主方程与跳过程

在一个系统中包含着许多粒子或个体(动物、细菌等等)的情况极常见，使得用跳过程作出的刻画可以做到令人信服的地步。在这种情况下我们发现，正如在1.1节已提到的，在适当的极限意义下，会产生宏观确定性的规律，围绕着这个规律，过程的随机特性则产生着一个涨落部分。不过，确定性运动及涨落都直接产生于以个体的跳变或迁移来表示的同一刻画。就这点来说，用跳过程(及其相应的主方程)来表示的刻画是非常令人满意的。

与之相对照，我们似乎也可以近似地用随机微分方程来建立这种系统的模型，其中的确定性运动与涨落具有完全独立的来源。但在这种模型中分别对涨落和确定性运动作独立无关的刻画是令人不解的，而且为了获得有关涨落的一些信息还必须要有一些涨落耗散变元。就此而言，主方程方法则是一种完善得多的刻画方式。

但是宏观上确定性规律的存在就是一个意义十分重大的成果，而且本章中我们将说明常常存在一个极限过程，使一个主方程的解(借助于一个刻画系统大小的大参量 Ω) 可渐近地用一个(是某确定性微分方程之解的) 确定性部分加上可用某个随机微分方程刻画的涨落部分来近似，此随机微分方程的系数则是由原来的主方程所给定。在3.8.3节我们考察泊松过程这个十分简单的跳过程时，已经注意到了这种渐近展开，我们在7.2节还要详细地研究。

这些展开的结果是发展出一些较简单的法则来(在渐近逼近意义下)写出等价于主方程的福克-普朗克方程，而且事实上在实

践中常常十分简单，只要写下适当的近似福克-普朗克方程而无需陈述主方程本身。有好几种方法表述一阶的近似福克-普朗克方程，它们都是等价的。但是按 Ω^{-1} 的幂次作系统展开的方法还只有一种，这就是范坎彭的系统大小展开。

本章还包含了泊松表示理论的一个概要，这是作者及其合作者们所发明的方法，它对某类主方程建立了福克-普朗克方程与主方程的精确相等。在这种特殊情况下，系统大小展开则作为泊松表示下的福克-普朗克方程的小噪声展开而出现。

7.1 生灭主方程——单变量情况

所有生灭系统的一维原型由一些个体的群体 X 组成。群体中可能出现的个体数称为 x，是个非负整数。我们要考察条件概率 $P(x, t|x', t')$ 及其相应主方程。诞生与死亡的概念通常是指在一给定事件中 X 只有有限个个体产生(生下来)或毁灭(死亡)。最简单的情况是 X 在一次只能以某一与时间无关的概率生出或死掉一个，从而其转移概率 $W(x|x', t)$ 可以写为

$$W(x|x', t) = t^+(x')\delta_{x,x'+1} + t^-(x')\delta_{x,x'-1} \tag{7.1.1}$$

这样，有二个过程存在，

$$x \to x+1; \quad t^+(x) = \text{每单位时间内的转移概率} \tag{7.1.2}$$

$$x \to x-1 \quad t^-(x) = \text{每单位时间内的转移概率} \tag{7.1.3}$$

一般主方程(3.5.5)则取形式

$$\begin{aligned}\partial_t P(x, t|x', t') = {} & t^+(x-1)P(x-1, t|x', t') \\ & + t^-(x+1)P(x+1, t|x', t') \\ & - [t^+(x) + t^-(x)]P(x, t|x', t')\end{aligned} \tag{7.1.4}$$

除了是与时间无关的情况外，这个方程无一般解法。

7.1.1 平稳解

我们可以把平稳解 $P_s(x)$ 的方程写为

$$0 = J(x+1) - J(x) \tag{7.1.5}$$

其中

$$J(x) = t^-(x)P_s(x) - t^+(x-1)P_s(x-1) \tag{7.1.6}$$

注意到 x 为非负整数的事实，不会有负数个个体。这就要求

(i) $t^-(0) = 0$：若不存在任何个体，则有个体死亡的概率是零 (7.1.7)

(ii) $P(x, t|x', t') = 0$ 对 $x<0$ 或 $x'<0$ (7.1.8)

这意味着

$$J(0) = t^-(0)P_s(0) - t^+(-1)P_s(-1) = 0 \tag{7.1.9}$$

现在对(7.1.5)求和，就使得

$$0 = \sum_{z=0}^{x-1} [J(z+1) - J(z)] = J(x) - J(0) \tag{7.1.10}$$

因此

$$J(x) = 0 \tag{7.1.11}$$

并且由此

$$P_s(x) = \frac{t^+(x-1)}{t^-(x)} P_s(x-1) \tag{7.1.12}$$

所以

$$\boxed{P_s(x) = P_s(0) \prod_{z=1}^{x} \frac{t^+(z-1)}{t^-(z)}} \tag{7.1.13}$$

a) 细致平衡解释

条件 $J(x) = 0$ 可以看作为细致平衡要求，其中 x 是个偶变量。这是因为它显然是细致平衡条件(5.3.74)的形式，这里具体化为下列形式：

$$P(x, \tau|x', 0)P_s(x') = P(x', \tau|x, 0)P_s(x) \tag{7.1.14}$$

令 $x' = \pm 1$，取 $\tau \to 0$ 的极限并注意从定义(3.4.1)得到

$$W(x|x', t) = \lim_{\tau \to 0} P(x, t+\tau|x', t)/\tau \tag{7.1.15}$$

就不难证明这个条件的必要性。

b) 变化速率方程

我们注意 x 的均值满足

$$\partial_t \langle x(t) \rangle = \partial_t \sum_{x=0}^{\infty} xP(x, t|x', t') \tag{7.1.16}$$

$$= \sum_{x=0}^{\infty} x[t^+(x-1)P(x-1,\ t|x',\ t') - t^+(x)P(x,\ t|x',\ t') + t^-(x+1)P(x+1,\ t|x',\ t') - t^-(x)P(x+1,\ t|x',\ t')] \tag{7.1.17}$$

$$= \sum_{x=0}^{\infty} [(x-1)t^+(x) - xt^+(x) + (x-1)t^-(x) - xt^-(x)]P(x,\ t|x',\ t') \tag{7.1.18}$$

也就是说，

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\langle x(t)\rangle = \langle t^+[x(t)]\rangle - \langle t^-[x(t)]\rangle} \tag{7.1.19}$$

相应的确定性方程则应由略去涨落而得到，即

$$\frac{\mathrm{d}x}{\mathrm{d}t} = t^+(x) - t^-(x) \tag{7.1.20}$$

注意当

$$t^+(x) = t^-(x) \tag{7.1.21}$$

时就出现了确定性的平稳态。与此相对应，要注意 $P_s(x)$ 的极大值则是

$$P_s(x)/P_s(x-1) \simeq 1 \tag{7.1.22}$$

时出现的，按(7.1.12)，它对应于

$$t^+(x-1) = t^-(x) \tag{7.1.23}$$

由于变量 x 仅取整数值，对于充分大的 x 来说，(7.1.21) 与 (7.1.23)实质上是相同的。

于是，对应于(7.1.23)的峰值是对应于(7.1.21)的确定性稳态之平稳随机类似值。

7.1.2 例子：化学反应 $X \rightleftharpoons A$

我们要处理的化学反应 $X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} A$ 假定了 A 是一给定了的浓度。这样，我们假设

$$t^+(x) = k_2 a \tag{7.1.24}$$

$$t^-(x) = k_1 x \tag{7.1.25}$$

而使主方程有简单的形式[其中简记 $P(x, t|x', t')$ 为 $P(x, t)$]

$$\partial_t P(x, t) = k_2 a P(x-1, t) + k_1(x+1)P(x+1, t) - (k_1 x + k_2 a)P(x, t) \tag{7.1.26}$$

a) 母函数

为解此方程，我们引进母函数(见 1.4.1 节，3.8.2 节)

$$G(s, t) = \sum_{x=0}^{\infty} s^x P(x, t) \tag{7.1.27}$$

使得

$$\boxed{\partial_t G(s, t) = k_2 a(s-1)G(s, t) - k_1(s-1)\partial_s G(s, t)} \tag{7.1.28}$$

若代入

$$\varphi(s, t) = G(s, t)\exp(-k_2 a s/k_1) \tag{7.1.29}$$

则(7.1.28)变成

$$\partial_t \varphi(s, t) = -k_1(s-1)\partial_s \varphi(s, t) \tag{7.1.30}$$

进一步作代换 $s-1=e^z$，则由 $\varphi(s, t) = \psi(z, t)$ 给出

$$\partial_t \psi(z, t) + k_1 \partial_z \psi(z, t) = 0 \tag{7.1.31}$$

其解是 $k_1 t - z$ 的任意函数。为方便起见把它写为

$$\psi(z, t) = F[\exp(-k_1 t + z)]e^{-k_2 a/k_1} = F[(s-1)e^{-k_1 t}]e^{-k_2 a/k_1}$$

而使得有

$$\boxed{G(s, t) = F[(s-1)e^{-k_1 t}]\exp[(s-1)k_2 a/k_1]} \tag{7.1.32}$$

归一化条件要求 $G(1, t) = 1$，因而

$$F(0) = 1 \tag{7.1.33}$$

b) 条件概率

初始条件确定了 F，这就是(令 $t'=0$ 后)

$$P(x, 0|N, 0) = \delta_{x,N} \tag{7.1.34}$$

它意味着

$$G(s,0)=s^N=F(s-1)\exp[(s-1)k_2a/k_1] \tag{7.1.35}$$

从而

$$G(s,t)=\exp\left[\frac{k_2a}{k_1}(s-1)(1-e^{-k_1t})\right][1+(s-1)e^{-k_1t}]^N \tag{7.1.36}$$

它可展开成 s 的幂级数而且给出

$$\begin{aligned}P(x,t|N,0)=&\exp\left[-\frac{k_2a}{k_1}(1-e^{-k_1t})\right]\\&\times\sum_{r=0}^{x}\frac{N!}{(N-r)!r!(x-r)!}\\&\times\left(\frac{k_2a}{k_1}\right)^{x-r}(1-e^{-k_1t})^{N+x-2r}e^{-k_1tr}\end{aligned} \tag{7.1.37}$$

这个非常复杂的答案是本问题的完整解答，不过实际用处很小。更好的做法是直接用母函数(7.1.36)或者由均值的方程来做。

从母函数出发，我们可以计算出

$$\begin{aligned}\langle x(t)\rangle&=\partial_sG(s=1,t)\\&=\frac{k_2a}{k_1}(1-e^{-k_1t})+Ne^{-k_1t}\end{aligned} \tag{7.1.38}$$

$$\begin{aligned}\langle x(t)[x(t)-1]\rangle&=\partial_s^2G(s=1,t)\\&=\langle x(t)\rangle^2-Ne^{-2k_1t}\end{aligned} \tag{7.1.39}$$

$$\mathrm{var}\{x(t)\}=\left(Ne^{-k_1t}+\frac{k_2a}{k_1}\right)(1-e^{-k_1t}) \tag{7.1.40}$$

c) 矩方程

由微分方程(7.1.28)我们有

$$\begin{aligned}\partial_t[\partial_s^nG(s,t)]=\{&n[k_2a\partial_s^{n-1}-k_1\partial_s^n]\\&+(s-1)[k_2a\partial_s^n-k_1\partial_s^{n+1}]\}G(s,t)\end{aligned} \tag{7.1.41}$$

令 $s=1$ 并利用

$$\partial_s^nG(s,t)|_{s=1}=\langle x(t)^n\rangle_t \tag{7.1.42}$$

则得到

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\langle x(t)^n\rangle_f = n[k_2 a\langle x(t)^{n-1}\rangle_f - k_1\langle x(t)^n\rangle_f]} \tag{7.1.43}$$

并且这些方程形成封闭的有层次性系统。自然，其均值和方差解即对应着 (7.1.38, 40)。

d) 自相关函数与平稳分布

当 $t\to\infty$ 时，从(7.1.32, 33) 可看出对任何 F

$$G(s,\ t\to\infty) = \exp[(s-1)k_2 a/k_1] \tag{7.1.44}$$

而且它对应着泊松分布解：

$$P_s(x) = \exp(-k_2 a/k_1)(k_2 a/k_1)^x/x! \tag{7.1.45}$$

由于 $\langle x(t)\rangle$ 的时间发展方程是线性的，所以可应用 3.7.4 节的方法，也就是回归定理，它说明了平稳自相关函数具有像均值之间那样的时间相依性，而且它在 $t=0$ 处之值就是平稳方差。所以

$$\boxed{\begin{aligned}\langle x(t)\rangle_s &= k_2 a/k_1 \qquad (7.1.46)\\ \mathrm{var}\{x(t)\}_s &= k_2 a/k_1 \qquad (7.1.47)\\ \langle x(t),\ x(0)\rangle_s &= \mathrm{e}^{-k_1 t}k_2 a/k_1 \qquad (7.1.48)\end{aligned}}$$

此泊松平稳分布解也可以由(7.1.13)用直接代入得到。

e) 时间相依的泊松分布解

这个方程的十分有趣的一个性质是存在时间相依的泊松分布解。这是因为若选取

$$P(x,\ 0) = \mathrm{e}^{-\alpha_0}\alpha_0^x/x! \tag{7.1.49}$$

则

$$G(s,\ 0) = \exp[(s-1)\alpha_0] \tag{7.1.50}$$

而且由(7.1.32)

$$G(s,\ t) = \exp[(s-1)(\alpha_0\mathrm{e}^{-k_1 t} + k_2 a/k_1)] \tag{7.1.51}$$

它对应于

$$P(x,\ t) = \mathrm{e}^{-\alpha(t)}\alpha(t)^x/x! \tag{7.1.52}$$

其中

$$\alpha(t) = \alpha_0\mathrm{e}^{-k_1 t} + k_2 a/k_1 \tag{7.1.53}$$

这里的 $\alpha(t)$ 被看作为确定性方程

$$\frac{dx}{dt} = k_2\alpha - k_1 x \tag{7.1.54}$$

在具有初始条件 $x(0) = \alpha_0$ 时的解。 (7.1.55)

这个结果可以推广到多变量场合，并且构成 7.7 节将进一步展开的泊松表示理论的基础。泊松传播解的存在是系统的线性之后果。

7.1.3 一个化学双稳系统

我们考察被许多作者[7.1]研究过的系统

$$A + 2X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3X \tag{7.1.56}$$

$$A \underset{k_4}{\overset{k_3}{\rightleftharpoons}} X \tag{7.1.57}$$

A 的浓度保持固定，所以

$$\begin{aligned} t^+(x) &= k_1 A x(x-1) + k_3 A \\ t^-(x) &= k_2 x(x-1)(x-2) + k_4 x \end{aligned} \tag{7.1.58}$$

自然，其对应的确定性方程是

$$\frac{dx}{dt} = t^+(x) - t^-(x) \simeq -k_2 x^3 + k_1 A x^2 - k_4 x + k_3 A \tag{7.1.59}$$

这里假定了 $x \gg 1$，从而我们置 $x(x-1)(x-2) \simeq x^3$ 以及类似近似等等，此方程具有初始条件 $x(0) = x_0$ 的解则由

$$\left(\frac{x-x_1}{x_0-x_1}\right)^{x_2-x_3}\left(\frac{x-x_2}{x_0-x_2}\right)^{x_1-x_3}\left(\frac{x-x_3}{x_0-x_3}\right)^{x_3-x_1}$$

$$= \exp[-k_2(x_1-x_2)(x_2-x_3)(x_3-x_1)t] \tag{7.1.60}$$

给出，这里的 x_1, x_2, x_3 是

$$k_2 x^3 - k_1 A x^2 + k_4 x - x_3 A = 0 \tag{7.1.61}$$

的根，同时 $x_3 \geqslant x_2 \geqslant x_1$。

显然，这些根是(7.1.59)的解 $x(t)$ 的平稳值。由(7.1.59)我们看到

$$
\begin{aligned}
&x<x_1 && \Longrightarrow \frac{dx}{dt}>0\\
&x_2>x>x_1 && \Longrightarrow \frac{dx}{dt}<0\\
&x_3>x>x_2 && \Longrightarrow \frac{dx}{dt}>0\\
&x>x_3 && \Longrightarrow \frac{dx}{dt}<0
\end{aligned}
\tag{7.1.62}
$$

于是，在区域 $x<x_2$ 中 $x(t)$ 将被吸引到 x_1，在区域 $x>x_2$ 中 $x(t)$ 将被吸引到 x_3。解 $x(t)=x_2$ 对于小摄动来说都是不稳定的。这产生了一个具有两个确定性稳定平稳态的系统。

a）随机平稳解

由(7.1.13)

$$
P_s(x)=P_s(0)\prod_{s=1}^{x}\left\{\frac{B[(z-1)(z-2)+P]}{z(z-1)(z-2)+Rz}\right\} \tag{7.1.63}
$$

其中

$$
\begin{aligned}
&B=k_1A/k_2\\
&R=k_4/k_2\\
&P=k_3/k_1
\end{aligned}
\tag{7.1.64}
$$

注意若 $P=R$，则解 (7.1.63) 是具有均值 B 的泊松分布。在此情况下我们有一个平稳状态，在此状态下反应 (7.1.56, 57) 同时处于平衡之中。这是化学平衡，以后我们将要指出，这时总存在泊松分布解(见第7.5.1节及7.7b节)。根据 (7.1.21)，当

$$
B=x[(x-1)(x-2)+R]/[P+x(x-1)] \tag{7.1.65}
$$

时(7.1.63)出现极大值。代入(7.1.65)而得到的函数 $x=x(B)$ 则按照已给的 P 及 R 所定的 B 值给出此极大(或极小)值。

有两种渐近形式：

$$
\begin{aligned}
&x(B)\sim B && \text{对大值的 } B\\
&x(B)\sim PB/R && \text{对小值的 } B
\end{aligned}
\tag{7.1.66}
$$

如果 $R>9P$，则可证明对于某个 $x>0$ 的区域 $x(B)$ 的斜率会成负的，因而对一给定的 B 会得到三个解，正如图 7.1 所表明的那

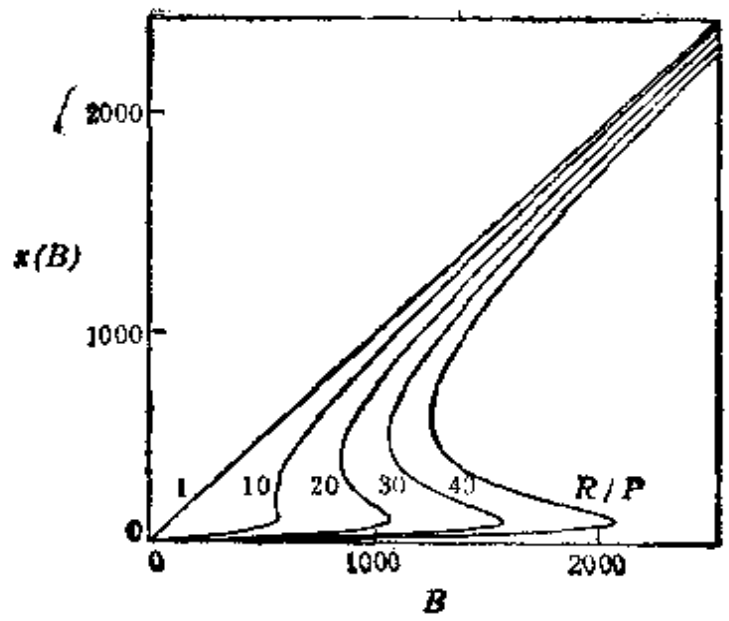


图7.1 $x(B)$对B的曲线

B是作为(7.1.65)的解对各种R/P的值来给出的，并且$P=10,000$

样。从一直线到另一直线的转换则给出了所看到的扭折。

还请注意，对于已标明的这些参数值来说，在B的一个非常小的范围上出现的双峰值形状是有重要含义的。这个值域比$P(x)$出现双峰值的范围远为狭窄，因为后者峰值间的大小之比值可以非常大。

可以给出一个更精确的结果。假定系统的体积V变得很大而且X的浓度y由

$$y=x/V$$

给定为常数。显然转移概率应该按V那样的尺度来度量，因为X的产生速率是像$x=yV$这样来度量的。

因此

$$\begin{aligned} &k_1A\sim 1/V \\ &k_3A\sim V \\ &k_2\sim 1/V^2 \\ &k_4\sim 1 \end{aligned} \tag{7.1.67}$$

而这意味着

$$B\sim V$$

$$R\sim V^1 \qquad (7.1.68)$$
$$P\sim V^2$$

于是我们记

$$B=\tilde{B}V$$
$$R=\tilde{R}V^2$$
$$P=\tilde{P}V^2$$

以使得(7.1.65)变成

$$\tilde{B}\simeq y(y^2+\tilde{R})/(y^2+\tilde{P})$$

此外，若 y_1 及 y_2 为 y 的两个值，则

$$\log[P_s(y_2)/P_s(y_1)]\simeq\sum_{s=y_1V}^{y_2V}\{\log\tilde{B}V+\log(z^2+\tilde{P}V^2)-\log[z(z^2+\tilde{R}V^2)]\} \qquad (7.1.69)$$

再用一个积分作为近似，

$$\sim V\int_{y_1}^{y_2}dy\left[\log\left(\frac{\tilde{B}(y^2+\tilde{P})}{y(y^2+\tilde{R})}\right)\right]$$

就有

$$\frac{P_s(y_2)}{P_s(y_1)}\simeq\exp\left[V\int_{y_1}^{y_2}\log\left(\frac{\tilde{B}(y^2+\tilde{P})}{y(y^2+\tilde{R})}\right)\right] \qquad (7.1.70)$$

并且当 $V\to\infty$ 时，按积分所取值的正负比值不是趋向零就是趋为无穷大。这样一来，在体积增大的极限过程下，这两个峰值除非大小完全相等，否则就会越来越不相等而只剩下一个。

分布的方差可用一简单技巧来得到。注意由 (7.1.63) $P_s(x)$ 可写为

$$P_s(x)=B^xG(x) \qquad (7.1.71)$$

其中 $G(x)$ 是(7.1.63)所定义的函数。于是

$$\left.\begin{aligned}\langle x^h\rangle&=\left[\sum_{x=0}^{\infty}x^hB^xG(x)\right]\left[\sum_{x=0}^{\infty}B^xG(x)\right]^{-1}\\ B\frac{\partial}{\partial B}\langle x^h\rangle&=\langle x^{h+1}\rangle-\langle x\rangle\langle x^h\rangle\end{aligned}\right\} \qquad (7.1.72)$$

从而

$$\mathrm{var}\{x\} = B\frac{\partial}{\partial B}\langle x\rangle \tag{7.1.73}$$

由此我们看到当 $V \to \infty$ 时

$$\mathrm{var}\{x\} \sim \frac{1}{V} \to 0 \tag{7.1.74}$$

这就获得了一个确定性极限。不仅如此，若$\langle x\rangle$与B是成比例的，则方差等于均值。事实上，我们发现(7.1.66)的两支为

$$\langle x(B)\rangle = \mathrm{var}\{x(B)\} = B \qquad B \text{ 值大时}$$

$$\langle x(B)\rangle = \mathrm{var}\{x(B)\} = PB/R \qquad B \text{ 值小时}$$

这意味着在这些极限化了的分布情况下，它们大体上是泊松分布。

事实上随机均值并非由这些峰值所精确给出，而是由它们很好地近似。当然，对于任何B都存在一个完全确定的$\langle x(B)\rangle$，而不是三个值。数值计算表明均值很接近下面的一支，并且然后在B_0处突然作一变移而接近上面的一支。这个值就是使得两个极大有相等大小的地方，原则上可由(7.1.70)来确定它。

b) 与时间相依的行为

不可能对它作精确推演。几乎所有的近似方法都依赖于取大体积极限，它们的性质在平稳情况下已提到过了，下一节，将要对它系统地加以研究。

7.2 用福克-普朗克方程逼近主方程

使一些明确的标度定律得以成立的参数V之存在，导致了系统大小展开的概念，并且范坎彭 [7.2] 最早把它建立在系统化的基础上。关于这一点历史上曾有过混淆，引起人们再三重复地想寻找主方程的极限形式，从中产生出一个福克-普朗克方程来。但是基本的结论是一个扩散过程总可以由跳过程来近似，而不是相反。

7.2.1 扩散过程的跳过程近似

其原型是在3.8.2节中用随机游动得到的结果，在这里，当取

无穷小跳跃度的极限时，主方程就变成了一个福克-普朗克方程。显然跳变必须变得更可能且更小，这可以汇总为用一个标度假设来概括：存在一个参数 δ，使得一步的平均长度和一步步长的方差与 δ 成比例，而且使得 δ 变小时跳变出现的概率增大。

假设跳变概率可写为

$$W_\delta(x'|x)=\Phi\left(\frac{x'-x-A(x)\delta}{\sqrt{\delta}},x\right)\delta^{-\frac{3}{2}} \tag{7.2.1}$$

其中

$$\int \mathrm{d}y\,\Phi(y,x)=Q \tag{7.2.2}$$

以及

$$\int \mathrm{d}y\,y\,\Phi(y,x)=0 \tag{7.2.3}$$

这意味着

$$\begin{aligned}
&\alpha_0(x)\equiv\int \mathrm{d}x' W_\delta(x'|x)=Q/\delta\\
&\alpha_1(x)\equiv\int \mathrm{d}x'(x'-x)W_\delta(x'|x)=A(x)Q\\
&\alpha_2(x)\equiv\int \mathrm{d}x'(x'-x)^2W_\delta(x'|x)=\int \mathrm{d}y\,y^2\Phi(y,x)
\end{aligned} \tag{7.2.4}$$

我们进一步假设 $y\to\infty$ 时 $\Phi(y,x)$ 充分快地趋于零，从而使

$$\lim_{\delta\to0}W_\delta(x'|x)=\lim_{y\to\infty}\left[\left(\frac{y}{x'-x}\right)^3\Phi(y,x)\right]=0,\quad x'\neq x \tag{7.2.5}$$

条件(7.2.4, 5)与3.4节中的条件，即(3.4.1, 4, 5)非常类似，并且取两次可微函数 $f(z)$ 后，我们就可进行3.4节中所作过的许多相同运算来证明

$$\lim_{\delta\to0}\left\langle\frac{\partial f(z)}{\partial t}\right\rangle=\left\langle\alpha_1(z)\,\frac{\partial f}{\partial z}+\frac{1}{2}\,\alpha_2(z)\,\frac{\partial^2 f}{\partial x^2}\right\rangle \tag{7.2.6}$$

意味着 $\delta\to0$ 时主方程

$$\frac{\partial P(x)}{\partial t}=\int \mathrm{d}x'[W(x|x')P(x')-W(x'|x)P(x)] \tag{7.2.7}$$

变成为福克-普朗克方程

$$\frac{\partial P(x)}{\partial t} = -\frac{\partial}{\partial x}\alpha_1(x)P(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\alpha_2(x)P(x) \qquad (7.2.8)$$

这样一来，已给定(7.2.8)时，总可以构造一个依赖于一参数 δ 的主方程，它能按所希望的密切程度逼近前者。这样的主方程必将具有满足(7.2.4)要求的转移概率。如果它们不满足这个要求，则不可能有这个近似。用某些例子来说明是恰当的。

a) 随机游动(3.8.2节)

设 $x = nl$，则

$$W(x|x') = \mathrm{d}(\delta_{x',x-1} + \delta_{x',x+1}) \qquad (7.2.9)$$

于是

$$\begin{aligned} \alpha_0(x) &= 2d \\ \alpha_1(x) &= 0 \\ \alpha_2(x) &= 2l^2d \end{aligned} \qquad (7.2.10)$$

令

$$\delta = l^2 \qquad (7.2.11)$$

和

$$D = l^2d \qquad (7.2.12)$$

则所有要求都满足了，从而其极限方程正和3.8.2节一样是，

$$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial x^2} \qquad (7.2.13)$$

b) 泊松过程(见3.8.3节)

这里令 $x = nl$ 后，就有

$$W(x|x') = d\delta_{x,x'+l} \qquad (7.2.14)$$

以及

$$\begin{aligned} \alpha_0(x) &= d \\ \alpha_1(x) &= ld \\ \alpha_2(x) &= l^2d \end{aligned} \qquad (7.2.15)$$

不存在某种方法用 δ 来表现 l 与 d，使之成为参数，且从而使得 $l \to 0$ 并且 $\alpha_1(x)$ 及 $\alpha_2(x)$ 是有限值的。此情况下不存在福克-普朗克极限。

c) 用生灭主方程逼近扩散过程的一般方法

假设有主方程满足

$$W_\delta(x'|x)=\left(\frac{A(x)}{2\delta}+\frac{B(x)}{2\delta^2}\right)\delta_{x',x+\delta}+\left(-\frac{A(x)}{2\delta}+\frac{B(x)}{2\delta^2}\right)\delta_{x',x-\delta} \tag{7.2.16}$$

从而对充分小的 $\delta W_\delta(x'|x)$ 为正的，我们还假设在所关心的 x 的值域上它是一致正的。那末此过程是发生在 δ 的整数倍之点组成的 x 的值域上的。这不是具有 (7.2.1) 形式的那种过程，尽管如此，当 $\delta\to 0$ 时它仍然给出一个福克-普朗克方程。这是因为

$$a_0(x)=B(x)/\delta^2 \tag{7.2.17a}$$

$$a_1(x)=A(x) \tag{7.2.17b}$$

$$a_2(x)=B(x) \tag{7.2.17c}$$

以及

$$\lim_{\delta\to 0}W_\delta(x'|x)=0 \quad \text{对 } x'\neq x \tag{7.2.17d}$$

不过在这里，$a_0(x)$ 像 $1/\delta^2$ 那样发散，而不是像(7.2.14)中那样 $1/\delta$ 似地发散，且跳跃也不再是按某个光滑分布律发生。但由于 $a_0(x)$ 的变化无关紧要，而且极限福克-普朗克方程是

$$\frac{\partial P(x)}{\partial t}=-\frac{\partial}{\partial x}A(x)P(x)+\frac{1}{2}\frac{\partial^2}{\partial t^2}B(x)P(x) \tag{7.2.18}$$

所以证明还是没有问题的。我们看到，用这种形式我们已有了一种手段，用一个近似的生灭过程去模拟一个扩散过程。如果在 x 的值域的每一点都有 $B(x)=0$，则此方法不适用，因为它导致 $W_\delta(x'|x)$ 取负值。注意在这种情况下(*)主方程的平稳解是

$$\begin{aligned}P_s(x)&=P_s(0)\prod_{z=\delta}^{x}\left[\frac{\delta A(z-\delta)+B(z-\delta)}{-\delta A(z)+B(z)}\right]\\&=P_s(0)\left[\frac{-\delta A(0)+B(0)}{\delta A(x)+B(x)}\right]\\&\quad\times\prod_{z=0}^{x}\left[\frac{1+\delta A(z)/B(z)}{1-\delta A(z)/B(z)}\right]\end{aligned} \tag{7.2.19}$$

(*) 原文如此，是指在(7.2.18)成立情况下。——译注

从而对足够小的 δ

$$\log P_s(x) \to \text{const} - \log B(x) + \sum_{z=0}^{x} 2\delta A(z)/B(z) \qquad (7.2.20)$$

也就是如要求的那样，

$$P_s(x) \to \frac{\mathcal{N}}{B(x)} \exp\left[2\int_0^x \mathrm{d}z A(z)/B(z)\right] \qquad (7.2.21)$$

倘若在 x 的任何有限区间上 $A(x)/B(x)$ 是有界的，则极限在此区间上显然是一致收敛的。

7.2.2 克雷默斯-莫耶尔展开

克雷默斯[7.3]给出了它的简单但不严格的推导，其后，莫耶尔 (Moyal, J. E.)[7.4]作了相当的改进。爱因斯坦[7.5]曾不明显地用过它，正如 1.2.1 节中说明过的。

若在主方程(7.2.7)的第一项中定义

$$y = x - x'$$

并且在第二项中定义

$$y = x' - x$$

来代换 x'，再定义

$$t(y, x) = W(x + y | x) \qquad (7.2.22)$$

则主方程就成了

$$\frac{\partial P(x)}{\partial t} = \int \mathrm{d}y [t(y, x-y)P(x-y) - t(y, x)P(x)] \qquad (7.2.23)$$

把它展开为幂级数，则

$$= \int \mathrm{d}y \sum_{n=1}^{\infty} \frac{(-y)^n}{n!} \frac{\partial^n}{\partial x^n} [t(y, x)P(x)] \qquad (7.2.24)$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} [\alpha_n(x)P(x)] \qquad (7.2.25)$$

其中

$$\alpha_n(x) = \int \mathrm{d}x' (x' - x)^n W(x' | x) = \int \mathrm{d}y y^n t(y, x) \qquad (7.2.26)$$

取级数(7.2.25)的前二项，就得到福克-普朗克方程(7.2.8)。

在引入系统大小展开时，范坎彭对这个"证明"不满，因为它没有顾及要考虑到有小参数。尽管这样，这个做法还是受到广泛欢迎的——主要是因为方便以及结果简单。但是7.2.1节的证明已表明其有效性是有一些限制的。事实上，若假设 $W(x'|x)$ 具有形式 (7.2.1)，则可得到

$$a_n(x) = \delta^{\frac{n}{2}-1}\int \mathrm{d}y y^n \Phi(y, x) \tag{7.2.27}$$

所以当 $\delta \to 0$ 时展开式(7.2.25)(克雷默斯-莫耶尔展式)中比第二项高阶的项就要消失。在莫耶尔的论文 [7.4] 中的确要求了等价于(7.2.4, 5)的条件。

7.2.3 范坎彭的系统大小展开

生灭主方程为克雷默斯-莫耶尔展开失效的情况提供了一些很好的例子，其中最简单的就是7.2.1节提到过的泊松过程。

在所有这些例子中，跳跃度是 ± 1 或者某个小整数，但是变量的典型值却可以取大值，例如分子的总数或者在长格子点集上作随机游动者的位置都是这样。

在这种情况下，我们可以引入系统大小参数 Ω 使转移概率可按内含变量 x/Ω 等于以表现。例如，在7.1.3节的反应中，Ω 是体积 V 并且 x/Ω 是浓度。我们用范坎彭的记号：

a = 广延变量(如分子数等, $\propto \Omega$)

$x = a/\Omega$ 是内含变量(分子的浓度)

使人关心的限制是给定了 x 时 Ω 是很大的。这对应于接近宏观系统。我们可以把转移概率表示为

$$\begin{aligned} &W(a|a') = W(a', \Delta a) \\ &\Delta a = a - a' \end{aligned} \tag{7.2.28}$$

实质的要点是跳跃度大小用广延量 Δa 表示但对 a' 的相依关系则用内含量 x 来表示更好些。

于是假设可写出

$$W(a';\Delta a)=\Omega\psi\left(\frac{a'}{\Omega};\ \Delta a\right) \tag{7.2.29}$$

若是这种情况，则可作一个展开。我们选取新变量 z 使得

$$a=\Omega\phi(t)+\Omega^{1/2}z \tag{7.2.30}$$

其中 $\phi(t)$ 是个待定函数。现在的情况是 $\alpha_n(a)$ 与 Ω 成比例，我们记

$$\alpha_n(a)=\Omega\alpha_n(x) \tag{7.2.31}$$

取克雷默斯-莫耶尔展开(7.2.25)并代换变量，就得到

$$\frac{\partial P(z,\ t)}{\partial t}-\Omega^{1/2}\phi'(t)\ \frac{\partial P(z,\ t)}{\partial z}$$

$$=\sum_{n=1}^{\infty}\frac{\Omega^{1-n/2}}{n!}\left(-\frac{\partial}{\partial z}\right)^n\tilde{\alpha}_n[\phi(t)+\Omega^{-1/2}z]P(z,\ t) \tag{7.2.32}$$

若 $\phi(t)$ 服从所期待的确定性方程

$$\dot{\phi}(t)=\tilde{\alpha}_1[\phi(t)] \tag{7.2.33}$$

则两边 $\Omega^{1/2}$ 阶的项都消掉了。现在展开 $\tilde{\alpha}_1[\phi(t)+\Omega^{-1/2}z]$ 成 $\Omega^{-1/2}$ 的幂级数，经过整理，就得到

$$\frac{\partial P(z,\ t)}{\partial t}=\sum_{m=2}^{\infty}\frac{\Omega^{-(m-2)/2}}{m!}\sum_{n=1}^{\infty}\frac{m!}{n!\,(m-n)!}$$

$$\times\tilde{\alpha}_n{}^{(m-n)}[\phi(t)]\left(-\frac{\partial}{\partial z}\right)^n z^{m-n}P(z,\ t) \tag{7.2.34}$$

取 Ω 趋于无穷的极限，则只剩下 $m=2$ 的项，并且给出

$$\boxed{\begin{aligned}\frac{\partial P(z,\ t)}{\partial t}=&-\tilde{\alpha}_1[\phi(t)]\ \frac{\partial}{\partial z}\,zP(z,\ t)\\&+\frac{1}{2}\,\tilde{\alpha}_2[\phi(t)]\ \frac{\partial^2}{\partial z^2}\ P(z,\ t)\end{aligned}} \tag{7.2.35}$$

a) 与克雷默斯-莫耶尔结果的比较

由在(7.2.25)中只取头二项而得到的克雷默斯-莫耶尔福克-普朗克方程是

$$\frac{\partial P(a)}{\partial t}=-\frac{\partial}{\partial a}\ [\alpha_1(a)P(a)]$$

$$+\frac{1}{2}\frac{\partial^2}{\partial a^2}[\alpha_2(a)P(a)] \tag{7.2.36}$$

把变量代换到 $x=a/\Omega$，则得到

$$\frac{\partial P(x)}{\partial t}=-\frac{\partial}{\partial x}[\tilde{\alpha}_1(x)P(x)]$$

$$+\frac{1}{2\Omega}\frac{\partial^2}{\partial x^2}[\tilde{\alpha}_2(x)P(x)] \tag{7.2.37}$$

利用

$$\varepsilon^2=\frac{1}{\Omega} \tag{7.2.38}$$

我们就可以使用 6.3 节的小噪声理论，并且看到作了代换

$$z=\Omega^{1/2}[x-\phi(t)] \tag{7.2.39}$$

之后，z 的最低阶福克-普朗克方程与范坎彭方法(7.2.35)中的最低阶项相同。这就是说若仅对最低阶感兴趣，则可以用克雷默斯-莫耶尔福克-普朗克方程，它可能比范坎彭的方法更易处理。除了到 $\Omega^{-1/2}$ 的最低阶为止外，两者的结果是不同的，而且也只有到这最低阶为止时两者都适用。

这样一来，若福克-普朗克方程由某主方程得到，则其有效性取决于导出它时所用极限过程的种类。若它是由 7.2.1 节中用的 $\delta\to 0$ 这类极限所导出，则它可严格地得到，而且 $\alpha_1(a)$ 及 $\alpha_2(a)$ 对 a 的整个非线性相依关系可以展现出来。

另一方面，若像 7.2.3 节那样，福克-普朗克方程是由一个 Ω 展开产生的，则只有小噪声近似才会有效。关于确定性解去考虑比线性化的(7.2.35)更进一步的内容是没有意义的。根据 4.4.7 节的式(4.4.49)或 4.4.9 节的式(4.4.83)，这个方程的解可用相应的随机微分方程

$$dz=\tilde{\alpha}_1'[\phi(t)]z\,dt+\sqrt{\tilde{\alpha}_2[\phi(t)]}\,dW(t) \tag{7.2.40}$$

来给出。

b) 例子：化学反应 $X\rightleftarrows A$

由 7.1.2 节我们有

$$W(x|x')=\delta_{x,x'+1}k_2a+\delta_{x,x'-1}k_1x' \tag{7.2.41}$$

这里的假设是

$$a = a_0 V$$
$$x = x_0 V \tag{7.2.42}$$

其中V是系统的体积。这意味着我们假定了A与X的总量都要与V成比例(这是个合理的假定)，并且产物产生的速率和X的衰减速率分别跟a和x成比例。

于是

$$W(x_0'; \Delta x) = V[k_2 a_0 \delta_{\Delta x,1} + k_1 x_0' \delta_{\Delta x,-1}] \tag{7.2.43}$$

当$\Omega \to V$，$a \to x$，则它就是(7.2.29)的形式。这样，

$$\begin{aligned} \alpha_1(x) &= \Sigma(x'-x)W(x'|x) \\ &= k_2 a - k_1 x = V(k_2 a_0 - k_1 x_0) \\ \alpha_2(x) &= \Sigma(x'-x)^2 W(x'|x) \\ &= k_2 a + k_1 x = V(k_2 a_0 + k_1 x_0) \end{aligned} \tag{7.2.44}$$

确定性方程是

$$\phi'(t) = [k_2 a_0 - k_1 \phi(t)] \tag{7.2.45}$$

它的解为

$$\phi(t) = \phi(0)e^{-k_1 t} + \frac{k_2 a_0}{k_1}(1 - e^{-k_1 t}) \tag{7.2.46}$$

福克-普朗克方程则是

$$\begin{aligned} \frac{\partial P(z)}{\partial t} &= k_1 \frac{\partial}{\partial z} zP(z) \\ &+ \frac{1}{2}\frac{\partial^2}{\partial z^2}[k_2 a_0 + k_1 \phi(t)]P(z) \end{aligned} \tag{7.2.47}$$

由(4.4.84)，我们可算出

$$\langle z(t) \rangle = z(0)e^{-k_1 t} \tag{7.2.48}$$

通常我们总可以假设$z(0) = 0$，因为用$\phi(t)$的初始条件就可以很好地安排好这初始条件。假定了$z(0)$是零后，就有

$$\text{var}[z(t)] = \left[\frac{k_2 a_0}{k_1} + \phi(0)e^{-k_1 t}\right](1 - e^{-k_1 t}) \tag{7.2.49}$$

从而

$$\langle x(t) \rangle = V\phi(t) = V\phi(0)e^{-k_1 t} + \frac{k_2 a}{k_1}(1 - e^{-k_1 t}) \tag{7.2.50}$$

$$\text{var}\{x(t)\} = V\,\text{var}\{z(t)\}$$

$$= \left[\frac{k_2 a}{k_1} \cdots V\phi(0)e^{-k_1 t}\right](1 - e^{-k_1 t}) \qquad (7.2.51)$$

借助于恒等式 $V\phi(0) = N$，这些等式恰好与7.1.2节中的精确解(7.1.38～40)完全相同．(7.2.47)的平稳解是

$$P_s(x) = \mathcal{N} \exp\left(-\frac{k_1 x^2}{2k_2 a_0}\right) \qquad (7.2.52)$$

它是精确泊松分布的高斯分布逼近。

克雷默斯-莫耶尔方程的平稳解是

$$P_s(x) = \frac{\mathcal{N}}{\alpha_2(x)} \exp\left[\int_0^x 2\frac{\alpha_1(x')}{\alpha_2(x')}\,dx'\right]$$

$$= \mathcal{N}(k_2 a + k_1 x)^{-1 + 4k_1 a/k_2} e^{-2x} \qquad (7.2.53)$$

事实上，我们可明显地验算此极限，这只要令

$$x = V(k_2 a_0/k_1) + \delta \qquad (7.2.54)$$

使得

$$(7.2.53) \simeq \mathcal{N}(2Vk_2 a_0 + k_1\delta)^{-1+4Vk_1 a_0/k_2} e^{-2Vk_1 a_0/k_1 - 2\delta} \qquad (7.2.55)$$

于是

$$\log P_s(x) \simeq \text{const} - \frac{k_2}{2k_2 a_0 V}(\delta - \delta^2) \qquad (7.2.56)$$

如果使用精确的泊松分布解，作相同的代换并利用斯特林公式

$$\log x! \sim \left(x + \frac{1}{2}\right)\log x - x + \text{const} \qquad (7.2.57)$$

则会得到相同于(7.2.56)的结果，不过精确结果是不同的，即使是在对数比的意义下也不同．

事实上 δ 的线性项是 V 的更低阶项．利用(7.2.39)我们就会发现

$$\delta = z\sqrt{V}$$

并且

$$\log P_s(x) \sim \text{const} - \frac{k_1}{2k_2 a_0}\left(\frac{z}{\sqrt{V}} - z^2\right) \qquad (7.2.58)$$

所以取 V 增大之极限，就有个均值为零的简单高斯分布．

c) 矩的分层性系统

我们可以按(7.2.34)这个展开来展开矩的方程

$$\langle z^k\rangle = \int \mathrm{d}z P(z,\ t) z^k \tag{7.2.59}$$

用直接代入及分部积分就得到：

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle z^k\rangle = \sum_{m=2}^{\infty} \frac{\Omega^{-1/2(m-2)}}{m!} \times \sum_{n=1}^{m,k} \frac{m!\,k!}{n!\,(m-n)!\,(k-n)!} \alpha_n^{(m-n)}[\phi(t)]\langle z^{m+k-2n}\rangle \tag{7.2.60}$$

把$\langle z^k\rangle$展开成$\Omega^{1/2}$的倒数之幂次，就可开发出一个有层次的系统：

$$\langle z^k\rangle = \sum_{r=0}^{\infty} M_r^k \Omega^{-r/2} \tag{7.2.61}$$

如果使用那些在6.3.1节中对福克-普朗克方程的小噪声展开处理矩的分层系统时所用过的相同技巧，我们就可以从现在这个分层系统计算各阶平稳矩和自相关函数。范坎彭[7.2]已经完成了这个工作。

7.2.4 库尔茨定理

库尔茨(Kurtz, T. G.)[7.6]证明了在某种意义上说，克雷默斯-莫耶尔展开可以得出比范坎彭展开还要稍强些的结果。对于具有多项分布转移概率的特定的一类生灭过程，他说明了下列结果。我们考虑服从生灭主方程

$$\partial_t P(a,\ t) = \sum_{a'} W(a|a')P(a',\ t) - \sum_{a'} W(a'|a)P(a,\ t) \tag{7.2.62}$$

的极限过程，它满足标度条件(7.2.29)。那末满足随机微分方程

$$\mathrm{d}b(t) = \alpha_1(b)\mathrm{d}t + \sqrt{\alpha_2(b)}\,\mathrm{d}W(t) \tag{7.2.63}$$

的过程$b(t)$是存在的，而且对于(7.2.62)的每一路径$a(t)$，存在着(7.2.63)的$b(t)$之一样本路途，使得对所有有限值的t

$$|b(t) - a(t)| \sim \log V \tag{7.2.64}$$

这个结果蕴涵着范坎彭的关于最低阶的结果。这是因为若取(7.2.30)形式的代换

$$a(t) = V\phi(t) + V^{1/2}z(t) \tag{7.2.65}$$

$$b(t) = V\phi(t) + V^{1/2}y(t) \tag{7.2.66}$$

则 $z(t)$ 的特征函数为

$$\begin{aligned}\langle \exp\,[is z(t)]\rangle &= \langle \exp[isV^{1/2}a(t) - isV^{1/2}\phi(t)]\rangle \\ &= \exp[-isV^{1/2}\phi(t)] \\ &\quad \times \langle \exp[isV^{-1/2}b(t)]\rangle + O(V^{-1/2}\log V) \\ &= \langle \exp[isy(t)]\rangle + O(V^{-1/2}\log V)\end{aligned} \tag{7.2.67}$$

现在使用福克-普朗克方程的渐近展开，就知道 $y(t)$ 的分布函数可接近福克-普朗克方程 (7.2.35) 的分布函数到 $O(V^{-1/2})$，因而得到了这个结果。不过由于这里含 $\log V$ 这一项，所以是一种稍弱的收敛。这样，从能够计算和度量出的一些量如均值、方差等等来衡量，库尔茨这个显得较强的结果等价于范坎彭的系统大小展开。

7.2.5 临界涨落

正如7.2.3节中概述过的，系统大小展开的存在依赖于 $\tilde{\alpha}'(\phi)$ 不为零。但

$$\tilde{\alpha}'(\phi_s) = 0$$

可能是要发生的情况，其中 ϕ_s 是确定性方程的平稳解。譬如，当考察7.1.3节中的反应时这种情况就要出现，此时（用该节的记号）

$$\tilde{\alpha}_1(y) = (\tilde{B}y^2 + \tilde{p} - y^3 - y\tilde{R})\tilde{k}_2 \tag{7.2.68}$$

其中 $\tilde{k}_2 = V^2 k_2$

可能会有两种情况出现，分别对应于图7.2中的 A 和 B。A 对应于一个不稳定的平稳态——任何向左的摄动将使它最终要到达 C，而 B 是稳定的。显然，确定性方程具有形式

$$\dot{y} = -\tilde{k}_2(y - \phi_s)^3 \tag{7.2.69}$$

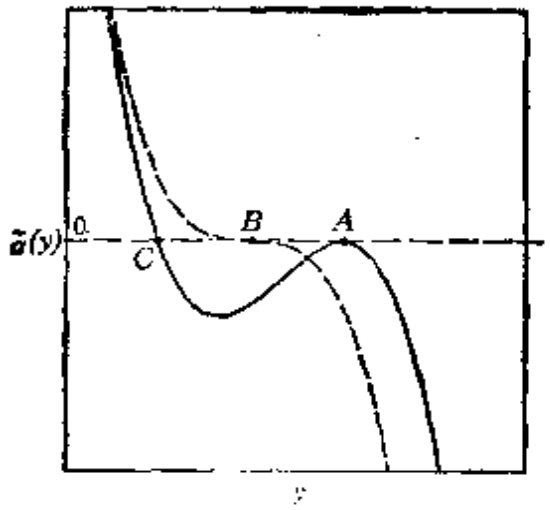


图 7.2　导致 $\alpha_1'(y)=0$ 的不同类型的 $\alpha_1(y)$ 行为

而且我们有个 6.2.4 a 节中立方过程的主方程类同物。

范坎彭[7.13]指出，在这种情况下我们应记

$$a=\Omega\phi(t)+\Omega^{\mu}u \tag{7.2.70}$$

把(7.2.32)变成

$$\frac{\partial P(z,\ t)}{\partial t}-\Omega^{1-\mu}\phi'(t)\frac{\partial P(z,\ t)}{\partial z}$$

$$=\sum_{n=1}^{\infty}\frac{\Omega^{1-\mu n}}{n!}\left(-\frac{\partial}{\partial z}\right)^{n}\tilde{a}_n[\phi(t)+\Omega^{\mu-1}z]P(z,\ t) \tag{7.2.71}$$

现在假设 $\tilde{\alpha}_1(\phi_s)$ 的头 $q-1$ 阶导数为零。那末只要选取 ϕ_s 为 $\phi(t)$，则 (7.2.71) 到最低阶成了

$$\frac{\partial P(z,\ t)}{\partial t}=-\frac{1}{q!}\tilde{\alpha}_1^{(q)}(\phi_s)\Omega^{(1-q)(1-\mu)}\frac{\partial}{\partial z}z^qP(z,\ t)$$

$$+\frac{1}{2}\tilde{\alpha}_2(\phi_s)\Omega^{1-2\mu}\frac{\partial^2P}{\partial z^2}(z,\ t)+\text{高阶项}^{(*)} \tag{7.2.72}$$

为确保 z 保留一阶项，令

$$(1-q)(1-\mu)=(1-2\mu)，\text{也就是}\ \mu=q/(q+1) \tag{7.2.73}$$

结果就成了

$$\frac{\partial P}{\partial t}=\Omega^{-(q-1)/(q+1)}\left(-\frac{1}{q!}\tilde{\alpha}_1^{(q)}\frac{\partial}{\partial z}z^qP+\frac{1}{2}\tilde{\alpha}_2\frac{\partial^2P}{\partial z^2}\right) \tag{7.2.74}$$

(*) 原文此式有多处错漏，——译注

(其中 $\tilde{\alpha}_1^{(q)}$ 及 $\tilde{\alpha}_2$ 是计算 ϕ_s 处的值。)涨落是在一个更慢的时间标度 τ 上变动的，它由

$$\tau = t\Omega^{-(q-1)/(q+1)} \tag{7.2.75}$$

给定，而且均值的方程为

$$\frac{\mathrm{d}\langle x\rangle}{\mathrm{d}t} = \frac{1}{q!}\tilde{\alpha}_1^{(q)}\langle x^q\rangle \tag{7.2.76}$$

已经不再与线性化确定性方程相关了。当然，稳定性取决于 $\tilde{\alpha}_1^{(q)}$ 的正负以及 q 是奇还是偶的。最简单的稳定情况出现在 $q=3$，它出现在图 7.2 的临界点 B，并且这时我们有 6.3.4 节的立方过程。长时间的标度是

$$\tau = \Omega^{-1/2}t \tag{7.2.77}$$

我们看到，系统的时间相依性对大值 Ω 来说是给定为 $\tau = \Omega^{-1/2}t$ 的一个函数的，只有 $t > \Omega^{1/2}$，τ 的确成了一个显著(不为零)的值，并且因此也只有对那些非常长的时间 t，系统才有明显的时间发展。这样一来，在大 Ω 情况下系统的运动变得非常慢。

条件(7.2.68)通常可由某些外界参数(例如温度)来控制，参数空间中使(7.2.68)得到满足的点称为临界点。系统在临界点变化非常慢的性质称为临界慢化。

7.3 生灭过程的边界条件

对生灭过程，我们有一个相当简单的方法来提供边界条件。对限制在区间 $[a, b]$ 中的过程，显然可分别用禁止从区间跑出或返回区间来得到反射边界和吸收边界。即

	反射边界	吸收边界	
在边界 a	$t^-(a)=0$	$t^+(a-1)=0$	(7.3.1)
在边界 b	$t^+(a)=0$	$t^-(b+1)=0$	

不过有时下述做法很有用，即把边界加进过程中去，不是去令某些转移概率为零而是硬性赋予类似于 (5.2.1 节) 对福克-普朗克方程用过的那些边界条件，使得区间 $[a, b]$ 中的解是具有适当零转移概率的主方程的解。这可能正是我们为保持转移概率的特定解

析形式所希望的，它可能会具有某种方便之处。

a) 前向主方程

我们可把前向主方程写成

$$\partial_t P(x, t|x', t') = t^+(x-1)P(x-1, t|x', t') + t^-(x+1)P(x+1, t|x', t') - [t^+(x) + t^-(x)]P(x, t|x', t') \tag{7.3.2}$$

假如我们要在 $x=a$ 有个反射壁，则可通过要求

$$t^-(a) = 0$$

及

$$P(a-1, t|x', t') = 0 \tag{7.3.3}$$

来得到。受此要求影响的方程只限于 $\partial_t P(a, t|x', t')$ 的方程，相对它来说，为得到相同的方程可以不置 $t^-(a)=0$ 而代以引入虚构的 $P(a-1, t|x', t')$，要它满足

$$t^+(a-1)P(a-1, t|x', t') = t^-(a)P(a, t|x', t') \tag{7.3.4}$$

这可以看作为对福克-普朗克方程的反射壁提出的零流要求的类似物。

如果要在 $x=a$ 有个吸收壁，则可令

$$t^+(a-1) = 0 \tag{7.3.5}$$

在一旦到达 $a-1$ 后，过程就不再返回去并且以后的行为是令人不感兴趣的。受它影响的仅是 $\partial_t P(a, t|x', t')$ 的方程，而且为得到相同的方程可再次引进个虚构的 $P(a-1, t|x', t')$ 使得满足

$$P(a-1, t|x', t') = 0 \tag{7.3.6}$$

总之，我们有另外一种表达方式来硬性加上边界条件，在 $[a, b]$ 中产生与(7.3.1)相同的效果：

区间 $[a, b]$ 上的前向主方程	反射	吸收
在边界 a	$t^-(a)P(a) = t^+(a-1)P(a-1)$	$P(a-1)=0$
在边界 b	$t^+(b)P(b) = t^-(b+1)P(b+1)$	$P(b+1)=0$

(7.3.7)

b) 后向主方程

后向主方程(3.6 节) 是

$$\begin{aligned}\partial_t P(x,\ t|x',\ t') = t^+(x')\,[P(x,\ t|x'+1,\ t') \\ -P(x,\ t|x',\ t)] \\ +t^-(x')\,[P(x,\ t|x'-1,\ t) \\ -P(x,\ t|x',\ t')]\end{aligned} \tag{7.3.8}$$

当 $x=a$ 是反射壁时，显然 $t^-(a)=0$ 等价于构造一个虚构的 $P(x, t|a-1, t')$ 使得

$$P(x,\ t|a-1,\ t') = P(x,\ t|a,\ t') \tag{7.3.9}$$

在吸收壁的情况下，对 $x,\ x' \in [a,\ b]$，没有一个 $P(x,\ t|x',\ t')$ 的方程包含着 $t^+(a-1)$。但是由于 $t^+(a-1)=0$，所以在 $x'<a-1$ 时方程显然满足条件

$$P(x,\ t|x',t')=0,\ \ x\in[a,\ b],\ x'\leqslant a-1 \tag{7.3.10}$$

并且这条件对于方程在 $x'=a$ 时的影响是赋予

$$P(x,\ t|a-1,\ t')=0 \tag{7.3.11}$$

因此这就是所要求的边界条件。总起来，就是

区间 $[a,\ b]$ 上的后向主方程	反射	
在边界 a	$P(\cdot\,|a-1)=P(\ \ |a)$	$P(\cdot|a-1)=0$
在边界 b	$P(\cdot\,|b+1)=P(\cdot\,|b)$	$P(\cdot|b+1)=0$

(7.3.12)

7.4 平均首次通过时间

在简单的一步情况下这些量的计算近乎平行于福克-普朗克方程的相应量的计算(见 5.2.7 节)。我们假设系统限制在值域

$$a\leqslant x\leqslant b \tag{7.4.1}$$

之内，而且作为可能的情况，端点不是吸收就是反射。为了确定起见我们取定

在 $x=a$ 为反射壁，在 $x=b+1$ 为吸收壁

的系统加以考察。论证在本质上与5.2.7节的相同，而且我们看到初始位置为 x 的粒子最终被吸收所化平均时间 $T(x)$ 满足与后向主方程(7.3.8)有关的方程：

$$t^+(x)[T(x+1)-T(x)]+t^-(x)[T(x-1)-T(x)]=-1 \tag{7.4.2}$$

其边界条件与(5.2.159)相对应并且来自(7.3.12)，

$$T(a-1)=T(a) \tag{7.4.3a}$$

$$T(b+1)=0 \tag{7.4.3b}$$

现在定义

$$U(x)=T(x+1)-T(x) \tag{7.4.4}$$

使得(7.4.2)变成

$$t^+(x)U(x)-t^-(x)U(x-1)=1 \tag{7.4.5}$$

再定义

$$\phi(x)=\prod_{z=a+1}^{x}\frac{t^-(z)}{t^+(z)} \tag{7.4.6}$$

及

$$S(x)=U(x)/\phi(x) \tag{7.4.7}$$

则(7.4.5)等价于

$$t^+(x)\phi(x)[S(x)-S(x-1)]=-1 \tag{7.4.8}$$

它具有解

$$S(x)=-\sum_{z=a}^{x}1/[t^+(z)\phi(z)] \tag{7.4.9}$$

这个解满足边界条件(7.4.3a)并由此推出

$$U(a-1)-S(a-1)=0 \tag{7.4.10}$$

因此

$$T(x+1)-T(x)=-\phi(x)\sum_{z=a}^{x}1/[t^+(z)\phi(z)] \tag{7.4.11}$$

并且

$$\boxed{T(x)=\sum_{y=x}^{b}\phi(y)\sum_{z=a}^{y}1/[t^+(z)\phi(z)]} \tag{7.4.12}$$

式中 a 为反射壁，b 为吸收壁，$b>a$。此式还满足边界条件(7.4.3b)，$T(b+1)=0$。

类似地，若 a 为吸收壁而 b 为反射壁，则

$$T(x)=\sum_{y=a}^{x}\phi(y)\sum_{z=y}^{b}1/[t^{+}(z)\phi(z)] \tag{7.4.13}$$

式中 a 为吸收壁，b 为反射壁，$b>a$。此外，当 a 及 b 两者均为吸收壁时也可类似地导出对应于(5.2.158)的公式。

7.4.1 吸收概率

a 及 b 均有限时，到被吸收所化的平均时间总是有限的。但是若 b 在无穷远处并且是反射的，则此平均时间可能是发散的。这本身并不意味着不被吸收的有限概率(*)。精确的结果如下［文献 7.7，第 4.7 节］。

若过程发生于区间 (a,∞) 而且 a 是吸收的，则从状态 x 出发被吸收到状态 $a-1$ 中去的吸收概率给出如下。定义函数 $M(x)$ 为

$$M(x)=\sum_{y=a}^{\infty}\left[\prod_{z=a}^{y}\frac{t^{+}(z)}{t^{-}(z)}\right] \tag{7.4.14}$$

那末当 $M(x)<\infty$ 时，从状态 x 出发被吸收到 $a-1$ 的概率是

$$\frac{M(x)}{1+M(x)} \tag{7.4.15}$$

当 $M(x)=\infty$ 时，这个概率是 1；如果这个概率是 1，则到被吸收的平均时间是(7.4.13)。

7.4.2 与福克-普朗克方程的比较

公式 (7.4.12, 13) 确定非常类似于扩散过程的相应公式 (5.2.160, 161) (**)。事实上，用一下 7.2.1c 的模型则不难说明在取 $\delta\to 0$ 的极限下两者变为相同。

如果我们希望处理有关从一势垒上逃逸这类问题 (5.2.7c 节)，只要它可转换为势垒也出现的同类主方程问题，例如，像

(*) 作者的意图是指正概率。——译注

(**) 原文误写为(7.4.1, 2)。——译注

7.1.3节所讨论过的双稳态反应中那样，则可以做出十分类似的近似。在这个例子中我们来考察从一个平稳态x_1到另一个平稳态x_3的平均首次通过时间。

$x=0$处是个反射壁，所以我们所考察的区间是$(0, x_3)$，初始出发点是x_1。注意到有

$$\phi(x)=\prod_{z=1}^{x}\frac{t^-(z)}{t^+(z)}=\frac{P_s(0)}{P_s(x)} \tag{7.4.16}$$

所以

$$T(x_1\to x_3)=\sum_{y=x_1}^{x_3}[P_s(y)t^+(y)]^{-1}\sum_{z=0}^{y}P_s(y) \tag{7.4.17}$$

若假设$P_s(y)^{-1}$在非稳定点x_2有一很陡的极大值，就可以在(7.4.17)的所有其他因子中令$y=x_2$而得到

$$T(x_1\to x_3)\sim\frac{n_1}{t^+(x_2)}\sum_{y=x_1}^{x_3}[P_s(y)]^{-1} \tag{7.4.18}$$

其中

$$n_1=\sum_{z=0}^{x_1}P_s(z) \tag{7.4.19}$$

而且它是处于平稳分布的较低峰值的全概率。这个结果是(5.2.7c)节结果的离散类似物。

7.5 多变量生灭系统

有一类十分广泛的系统，其时间的发展可看作为某个群体中成员之间的个体冲突结果。它们包括，譬如说，

——基于碰撞的分子变化所产生的化学反应；

——有死亡、生育、交配和彼此吞噬活动的群体系统；

——通过接触使疫病从一个个体传染到另一个个体的流行病系统。

所有这些通常都可以用我称之为组合动力学来建立模型。在这里，冲突所引起的某种变化之转移概率大小与那类可能冲突的数目成比例。

例如，在化学反应$X\rightleftarrows 2Y$中，反应$X\longrightarrow 2Y$是因自发衰

变而发生的，这是仅含一个个体的一种退化型的冲突变化。这种冲突的数目就是X的数目；因此我们说

$$t(x\to x-1,\ y\to y+2)=k_1x \tag{7.5.1}$$

对于逆反应，我们可以用$y(y-1)/2$种不同方式配制Y分子对。因此

$$t(x\to x+1,\ y\to y-2)=k_2y(y-1) \tag{7.5.2}$$

一般地则可在分子之间、种类之间等等考虑多种冲突。用化学反应的语言，就可有如下一般的表述。

考虑一个含S个不同反应的n分量反应系统：

$$\sum_\alpha N_\alpha^A X_\alpha \underset{k_A^-}{\overset{k_A^+}{\rightleftharpoons}} \sum_\alpha M_\alpha^A X_\alpha \quad (A=1,\ 2,\ \cdots,\ s) \tag{7.5.3}$$

X_α的系数N_α^A是X_α在左边具有的分子数，而M_α^A是X_α在右边具有的分子数。我们引进向量记号使得

$$\begin{aligned}&\boldsymbol{x}=(x_1,\ x_2,\ \cdots,\ x_n)\\&\boldsymbol{N}^A=(N_1^A,\ N_2^A,\ \cdots,\ N_n^A)\\&\boldsymbol{M}^A=(M_1^A,\ M_2^A,\ \cdots,\ M_n^A)\end{aligned} \tag{7.5.4}$$

还定义

$$\boldsymbol{r}^A=\boldsymbol{M}^A-\boldsymbol{N}^A \tag{7.5.5}$$

显然，当反应A沿向前方向进行一步时，

$$\boldsymbol{x}\longrightarrow\boldsymbol{x}+\boldsymbol{r}^A \tag{7.5.6}$$

并且沿向后方向进行一步时，

$$\boldsymbol{x}\longrightarrow\boldsymbol{x}-\boldsymbol{r}^A \tag{7.5.7}$$

速率常数则用下式来定义：

$$\begin{aligned}t_A^+(\boldsymbol{x})&=k_A^+\prod_\alpha\frac{x_\alpha!}{(x_\alpha-N_\alpha^A)!}\\t_A^-(\boldsymbol{x})&=k_A^-\prod_\alpha\frac{x_\alpha!}{(x_\alpha-M_\alpha^A)!}\end{aligned} \tag{7.5.8}$$

它们分别与从$\boldsymbol{x}$个分子中选取组合$\boldsymbol{N}^A$或$\boldsymbol{M}^A$的方式数成正比。于是，主方程是

$$\partial_tP(\boldsymbol{x},\ t)=\sum_A\{t_A^-(\boldsymbol{x}+\boldsymbol{r}^A)P(\boldsymbol{x}+\boldsymbol{r}^A,\ t)$$

$$-t_A^+(\boldsymbol{x})P(\boldsymbol{x},\ t)]+[t_A^+(\boldsymbol{x}-\boldsymbol{r}^A)P(\boldsymbol{x}-\boldsymbol{r}^A,\ t)$$
$$-t_A^-(\boldsymbol{x})P(\boldsymbol{x},\ t)]\} \tag{7.5.9}$$

对于可出现 $\boldsymbol{r}^A$ 步反应的任一整值变量 $\boldsymbol{x}$ 来说，这个形式自然是表示其齐次主方程的十分一般的方式。只要对给出每单位时间的转移概率的 (7.5.8) 作了特定选择，就产生了一般的组合主方程。另一种名称是化学主方程，因为这种方程特别适用于化学反应。

7.5.1 细致平衡成立时的平稳解

一般并没有明确的方法把平稳解写成为实用的形式。但是，如果细致平衡得以满足，则平稳解就容易导出。仅由数组成的向量变量 $\boldsymbol{x}$ 只能是偶变量，因此对于所有的 A，细致平衡的形式只能取为(由 5.3.5 节)

$$t_A^-(\boldsymbol{x}+\boldsymbol{r}^A)P_s(\boldsymbol{x}+\boldsymbol{r}^A)=t_A^+(\boldsymbol{x})P_s(\boldsymbol{x}) \tag{7.5.10}$$

上式对所有的 A 成立，这是加在 t_A^+ 上的十分严厉的要求，它来自于这个事实，即(7.5.10)提供了对任何 n 及任何 $\boldsymbol{x}_0$ 计算 $P_s(\boldsymbol{x}_0+n\boldsymbol{r}^A)$ 的方法。对所有能有的 A 用这个方法，就可以在能表示成

$$\boldsymbol{x}=\boldsymbol{x}_0+\Sigma n_A\boldsymbol{r}^A\quad(n_A\text{ 为整数}) \tag{7.5.11}$$

的所有 $\boldsymbol{x}$ 组成的空间上产生 $P_s(\boldsymbol{x})$；但是所得到的解可能是不明确的，譬如说，因为由(7.5.10)可写出

$$\left.\begin{aligned}P_s[(\boldsymbol{x}+\boldsymbol{r}^A)+\boldsymbol{r}^B]&=\frac{P_s(\boldsymbol{x})t_A^+(\boldsymbol{x})t_B^+(\boldsymbol{x}+\boldsymbol{r}^A)}{t_A^-(\boldsymbol{x}+\boldsymbol{r}^A)t_B^-(\boldsymbol{x}+\boldsymbol{r}^A+\boldsymbol{r}^B)}\\ \text{但}\qquad\\ P_s[(\boldsymbol{x}+\boldsymbol{r}^B)+\boldsymbol{r}^A]&=\frac{P_s(\boldsymbol{x})t_B^+(\boldsymbol{x})t_A^+(\boldsymbol{x}+\boldsymbol{r}^B)}{t_B^-(\boldsymbol{x}+\boldsymbol{r}^B)t_A^-(\boldsymbol{x}+\boldsymbol{r}^A+\boldsymbol{r}^B)}\end{aligned}\right\} \tag{7.5.12}$$

用了组合公式(7.5.8)并代入(7.5.12)之后，就看出这个条件是自动满足的。

当相同的两个点可以用本质不同的两种方式彼此联接时，也就是譬如说，若

$$\left.\begin{aligned}&\boldsymbol{N}^A\neq\boldsymbol{N}^B\\ &\boldsymbol{M}^A\neq\boldsymbol{M}^B\end{aligned}\right. \tag{7.5.13}$$

但

$$\boldsymbol{r}^A = \boldsymbol{r}^B \equiv \boldsymbol{r}$$

时，此条件就并非毫无意义的了。

在这种情况下，(7.5.10) 中 $P_s(\boldsymbol{x}+\boldsymbol{r}^A)$ 的唯一性要求有

$$\frac{t_A^+(\boldsymbol{x})}{t_A^-(\boldsymbol{x}+\boldsymbol{r})} = \frac{t_B^+(\boldsymbol{x})}{t_B^-(\boldsymbol{x}+\boldsymbol{r})} \tag{7.5.14}$$

并且这意味着

$$\frac{k_A^+}{k_A^-} = \frac{k_B^+}{k_B^-} \tag{7.5.15}$$

如果有两个反应链 $A, B, C, \cdots$ 及 $A', B', C', \cdots$，使得

$$\boldsymbol{r}^A + \boldsymbol{r}^B + \boldsymbol{r}^C + \cdots = \boldsymbol{r}^{A'} + \boldsymbol{r}^{B'} + \boldsymbol{r}^{C'} + \cdots \tag{7.5.16}$$

则直接代入表明要

$$P_s(\boldsymbol{x}+\boldsymbol{r}^A+\boldsymbol{r}^B+\cdots) = P_s(\boldsymbol{x}+\boldsymbol{r}^{A'}+\boldsymbol{r}^{B'}+\cdots) \tag{7.5.17}$$

成立就只要

$$\begin{vmatrix} k_A^+ & k_B^+ & k_C^+ & \cdots \\ k_A^- & k_B^- & k_C^- & \end{vmatrix} = \begin{vmatrix} k_{A'}^+ & k_{B'}^+ & k_{C'}^+ & \cdots \\ k_{A'}^- & k_{B'}^- & k_{C'}^- & \end{vmatrix} \tag{7.5.18}$$

因而这也就是服从组合动力学的主方程中使细致平衡成立的条件。

此时，$P_s(\boldsymbol{x})$ 的一个解是多元泊松分布

$$P_s(\boldsymbol{x}) = \prod_a \alpha_a^{x_a} \mathrm{e}^{-\alpha_a} / x_a! \tag{7.5.19}$$

把它代入 (7.5.10) 作验算，就得到

$$\prod_a \frac{\alpha_a^{(x_a+r_a^A)}\mathrm{e}^{-\alpha_a}}{(x_a+r_a^A)!} \frac{k_A^-(x_a+r_a^A)!}{(x_a+r_a^A-M_a^A)!} = \prod_a \frac{\alpha_a^{x_a}\mathrm{e}^{-\alpha_a}}{x_a!} \frac{k_A^+ x_a!}{(x_a-N_a^A)!} \tag{7.5.20}$$

利用

$$r_a^A = M_a^A - N_a^A$$

的事实，就会发现

$$k_A^+ \prod_a \alpha_a^{N_a^A} = k_A^- \prod_a \alpha_a^{M_a^A} \tag{7.5.21}$$

但是，只有是遵从各种类型的守恒律最一般的解才具有这种形式。

例如在反应

$$X \rightleftarrows 2Y \tag{7.5.22}$$

中，量 $2x+y$ 是守恒的，于是平稳分布是

$$\frac{e^{-\alpha_1}\alpha_1^x}{x!}\frac{e^{-\alpha_2}\alpha_2^y}{y!}\Phi(2x+y) \tag{7.5.23}$$

其中 ϕ 是个任意函数。选取 $\phi(2x+y)=1$ 就给出了泊松分布解。另一种选择是

$$\phi(2x+y)=\delta(2x+y,\ N) \tag{7.5.24}$$

它对应着把 $2x+y$ 之和数固定在 N。

作为它的一种退化形式，人们有时考虑一种记为

$$A \rightleftarrows 2Y \tag{7.5.25}$$

的反应，不过这里的 A 被考虑为固定的确定性数值，并且可能有的反应是

$$\begin{aligned} &A \to 2Y;\quad t^+(y)=k_1 a \\ &2Y \to A;\quad t^-(y)=k_2 y(y-1) \end{aligned} \tag{7.5.26}$$

在此情况下守恒律就只是 y 总为偶的或总为奇的。平稳解则具有

$$P_s(y)=\frac{\alpha^y}{y!}\Psi(y,\ \alpha) \tag{7.5.27}$$

的形式，其中 $\Psi(y,\ \alpha)$ 是个这样的函数，它对 y 的相依性仅通过 y 的偶性或奇性来体现。

7.5.2 没有细致平衡时的平稳解

有一种在原则上一般地确定平稳解的方法，虽然它并未发现有很多实际用处。有兴趣的读者可在哈肯 [文献 7.8，第 4.8 节] 和施纳肯伯格 (Schnakenberg, J.) 的著作 [7.4] 中找到详细的处理方法。但是一般说来，近似方法却更应注意。

7.5.3 系统大小展开与有关的展开

总的来说我们发现了在化学主方程中确实存在系统大小展开，产生与吸收速率被期待是与系统的大小 Ω 成比例的。这意味着当 $\Omega \to \infty$ 时我们预料

$$x \sim \Omega\rho \tag{7.5.28}$$

其中 ρ 是化学浓度的配置。这样一来，我们必须使 $t_A^{\pm}(x)$ 在 $\Omega \to \infty$ 时是与 Ω 成比例的，从而提出了要求

$$\boxed{\begin{aligned} k_A^+ &\sim k_A^+ \Omega^{-\sum_\alpha N_\alpha^A + 1} \\ k_A^- &\sim k_A^- \Omega^{-\sum_\alpha M_\alpha^A + 1} \end{aligned}} \tag{7.5.29}$$

在这些场合下，就可以开发出多变量形式的范坎彭系统大小展开。它是如此之复杂，以致于不可能在这里明显导出它，但是正如单变量情况那样，我们有克雷默斯-莫耶尔展开，其头两项给出了一个扩散过程，而这个过程的渐近形式与系统大小展开所得到的相同。

如 7.2.2 节那样，这个克雷默斯-莫耶尔展开可由(7.5.9)完全相同地导出来。事实上是更容易，因为(7.5.9)已经具有适用的形式。这样，我们有

$$\begin{aligned} \partial_t P(x,\ t) = \sum_{A,n} \Big\{ &\frac{(r^A \cdot \nabla)^n}{n!} \ [t_A^-(x) P(x,\ t)] \\ &+ \frac{(-r^A \cdot \nabla)^n}{n!} \ [t_A^+(x) P(x,\ t)] \Big\} \end{aligned} \tag{7.5.30}$$

对它在二阶截尾后就得到

$$\boxed{\begin{aligned} \partial_t P(x,\ t) &= -\sum_\alpha \partial_\alpha [A_\alpha(x) P(x,\ t)] \\ &\quad + \frac{1}{2} \sum_{\alpha,\beta} \partial_\alpha \partial_\beta [B_{\alpha\beta}(x) P(x,\ t)] \end{aligned}} \tag{7.5.31}$$

$$\boxed{\begin{aligned} A_\alpha(x) &= \sum_A r_\alpha^A [t_A^+(x) - t_A^-(x)] \\ B_{\alpha\beta}(x) &= \sum_A r_\alpha^A r_\beta^A [t_A^+(x) + t_A^-(x)] \end{aligned}} \tag{7.5.32}$$

于是，以这种形式我们获得了对应于主方程的化学福克-普朗克方程。但是要注意，这实际上仅仅是作为一个近似才适用，其大体积渐近展开做到 $1/\Omega$ 阶为止，则与相应的主方程对应的项相同。

如果我们已满足于这种程度的近似，则用福克-普朗克方程常常比用主方程更为简单。

7.6 某些例子

7.6.1 $X+A \rightleftarrows 2X$

在这里

$$\begin{aligned} t^+(x) &= k_1 ax \\ t^-(x) &= k_2 x(x-1) \end{aligned} \tag{7.6.1}$$

因而，到 $1/\Omega$ 阶，

$$\begin{aligned} A(x) &= k_1 ax - k_2 x(x-1) \sim k_1 ax - k_2 x^2 \\ B(x) &= k_1 ax + k_2 x(x-1) \sim k_1 ax + k_2 x^2 \end{aligned} \tag{7.6.2}$$

7.6.2 $X \underset{k}{\overset{\gamma}{\rightleftarrows}} Y \underset{\gamma}{\overset{k}{\rightleftarrows}} A$

这里有

$$\left.\begin{aligned} t_1^+(x) &= \gamma y \\ t_1^-(x) &= kx \end{aligned}\right\} \quad r^1 = (1,\ -1) \tag{7.6.3}$$

$$\left.\begin{aligned} t_2^+(x) &= ka \\ t_2^-(x) &= \gamma y \end{aligned}\right\} \quad r^2 = (0,\ 1) \tag{7.6.4}$$

因此

$$A(x) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} (\gamma y - kx) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} (ka - \gamma y) \tag{7.6.5}$$

$$= \begin{bmatrix} \gamma y - kx \\ kx + ka - 2\gamma y \end{bmatrix} \tag{7.6.6}$$

$$\begin{aligned} \underset{\sim}{B}(x) = &\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} (1,\ -1) \right\} (\gamma y + kx) \\ &+ \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} (0,\ 1) \right\} (ka + \gamma y) \end{aligned} \tag{7.6.7}$$

$$= \begin{bmatrix} \gamma y + kx & -\gamma y - kx \\ -\gamma y - kx & 2\gamma y + kx + ka \end{bmatrix} \tag{7.6.8}$$

若利用平稳态附近的线性化形式，则

$$\gamma y = kx = ka \tag{7.6.9}$$

$$B_\varepsilon = \begin{bmatrix} 2ka & -2ka \\ -2ka & 4ka \end{bmatrix} \tag{7.6.10}$$

7.6.3 被捕食-捕食系统

1.3 节中的被捕食-捕食系统为我们感兴趣的一类系统提供了一个很好的例子。我们可以把它作为一个化学反应写成为

$$\begin{aligned} &\text{i)} \quad X + A \to 2X \qquad r^1 = (1,\ 0) \\ &\text{ii)} \quad X + Y \to 2Y \qquad r^2 = (-1,\ 1) \\ &\text{iii)} \quad Y \to B \qquad\quad r^3 = (0,\ -1) \end{aligned} \tag{7.6.11}$$

这些反应全是不可逆的(虽然可以引进可逆性)，所以

$$t_A^-(x) = 0 \quad (A = 1,\ 2,\ 3)$$

而

$$\begin{aligned} t_1^+(x) &= k_1 a \frac{x!}{(x-1)!} \frac{y!}{y!} = k_1 ax \\ t_2^+(x) &= k_2 \frac{x!}{(x-1)!} \frac{y!}{(y-1)!} = k_2 xy \\ t_3^+(x) &= k_3 \frac{x!}{x!} \frac{y!}{(y-1)!} = k_3 y \end{aligned} \tag{7.6.12}$$

现在主方程可以利用(7.5.9)明确地写出来，我们得到

$$\begin{aligned} \partial_t P(x,\ y) = {} & k_1 a(x-1)P(x-1,\ y) \\ & + k_2(x+1)(y-1)P(x+1,\ y-1) \\ & + k_3(y+1)P(x,\ y+1) \\ & - (k_1 ax + k_2 xy + k_3 y)P(x,\ y) \end{aligned} \tag{7.6.13}$$

这个方程没有精确解，所以必须用近似方法。

克雷默斯-莫耶尔方法。由式(7.5.32)

$$A(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} k_1 ax + \begin{bmatrix} -1 \\ 1 \end{bmatrix} k_2 xy + \begin{bmatrix} 0 \\ -1 \end{bmatrix} k_3 y \tag{7.6.14}$$

$$= \begin{bmatrix} k_1 ax - k_2 xy \\ k_2 xy - k_3 y \end{bmatrix} \tag{7.6.15}$$

$$\underset{\sim}{\boldsymbol{B}}(x)=\begin{bmatrix}1\\0\end{bmatrix}(1,\ 0)k_1ax+\begin{bmatrix}-1\\1\end{bmatrix}(-1,\ 1)k_2xy$$

$$+\begin{bmatrix}0\\-1\end{bmatrix}(0,\ -1)k_3y \tag{7.6.16}$$

$$=\begin{bmatrix}k_1ax+k_2xy & -k_2xy\\ -k_2xy & k_2xy+k_3y\end{bmatrix} \tag{7.6.17}$$

确定性方程是

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}k_1ax-k_2xy\\k_2xy-k_3y\end{bmatrix} \tag{7.6.18}$$

平稳态在

$$\begin{bmatrix}x_s\\y_s\end{bmatrix}=\begin{bmatrix}k_3/k_2\\k_1a/k_2\end{bmatrix} \tag{7.6.19}$$

为确定此状态的稳定性，我们来检验线性化了的确定性方程

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\delta_x\\\delta_y\end{bmatrix}=\frac{\partial A(x_s)}{\partial x_s}\delta_x+\frac{\partial A(x_s)}{\partial y_s}\delta_y$$

$$=\begin{bmatrix}k_1a-k_2y_s\\k_2y_s\end{bmatrix}\delta_x+\begin{bmatrix}-k_2x_s\\k_2x_s-k_3\end{bmatrix}\delta y \tag{7.6.20}$$

$$=\begin{bmatrix}0 & -k_3\\k_1a & 0\end{bmatrix}\begin{bmatrix}\delta_x\\\delta_y\end{bmatrix} \tag{7.6.21}$$

此矩阵的本征值是

$$\lambda=\pm i(k_1k_3a)^{\frac{1}{2}} \tag{7.6.22}$$

它表明了对平稳态的某种小偏离的周期性运动。由于扰动既不增长也不衰减，故我们有着随遇稳定性

这显与守恒量

$$V=k_2(x+y)-k_3\log x-k_1a\log y \tag{7.6.23}$$

的存在性有关的，V 的守恒性容易由验证它满足 $\mathrm{d}V/\mathrm{d}t=0$ 而得知。于是系统使 V 守恒，而且这意味着存在着不同的使 V 等于常值的循环轨迹。

再一次记

$$x = x_s + \delta x$$
$$y = y_s + \delta y \tag{7.6.24}$$

并展开到二阶为止，就看到

$$V = \frac{k_3}{2}\left(\frac{\delta x^2}{k_3^2} + \frac{\delta y^2}{(k_1 a)^2}\right) \tag{7.6.25}$$

从而这些轨线最初是椭圆(这也可以由线性化分析导出)。

随着轨线越来越大，它们就变得椭圆性越来越少，最终不是 x 就是 y 会变成零。

如果 x 首先变成零(所有的被捕食物已被吃掉)，则人们会看到 y 不可避免地也同样会变成零。若 y 变成零（所有捕食者都饿死了)，则被捕食物以指数增长速度不受抑制地增长起来。

随机行为 因为量 V 的守恒，轨道具有随意稳定性，这意味着当含有一个涨落时，系统趋向于随着时间的变化改变轨道的大小。我们可以直接从等价的随机微分方程

$$\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix} = \begin{bmatrix} k_1 ax - k_2 xy \\ k_2 xy - k_3 y \end{bmatrix} \mathrm{d}t = \underset{\sim}{C}(x, y) \begin{bmatrix} \mathrm{d}W_1(t) \\ \mathrm{d}W_2(t) \end{bmatrix} \tag{7.6.26}$$

中看到这点，其中

$$\underset{\sim}{C}(x, y)\underset{\sim}{C}(x, y)^T = \underset{\sim}{B}(x) \tag{7.6.27}$$

用一下伊藤公式

$$\mathrm{d}V(x, y) = \frac{\partial V}{\partial x}\mathrm{d}x + \frac{\partial V}{\partial y}\mathrm{d}y + \frac{1}{2}\left(\frac{\partial^2 V}{\partial x^2}\mathrm{d}x^2 + 2\frac{\partial^2 V}{\partial x \partial y}\mathrm{d}x\mathrm{d}y + \frac{\partial^2 V}{\partial y^2}\mathrm{d}y^2\right) \tag{7.6.28}$$

故

$$\begin{aligned} \langle \mathrm{d}V(x, y)\rangle = &\left\langle \frac{\partial V}{\partial x}(k_1 ax - k_2 xy) \right. \\ &\left. + \frac{\partial V}{\partial y}(k_2 xy - k_3 y)\right\rangle \mathrm{d}t \\ &+ \left\langle B_{11}\frac{k_3}{2x^2} + B_{22}\frac{k_1}{2y^2}\right\rangle \mathrm{d}t \end{aligned} \tag{7.6.29}$$

由于 V 是确定性地守恒，所以第一个平均为零，我们看到

$$\langle \mathrm{d}V(x,y)\rangle = \frac{1}{2}\left\langle \frac{k_2k_1a}{x} + \frac{k_3k_2y}{x} + \frac{k_1k_2ax}{y} + \frac{k_1k_3a}{y} \right\rangle \mathrm{d}t \tag{7.6.30}$$

所有 Ω^{-1} 阶的项在 x 与 y 是正值时都为正的，于是，在平均意义上说 $V(x,y)$ 是稳定增加的。当然最终不是这根就是那根坐标轴要与之相交，而且在确定性情况时也会出现类似的结果。我们看到当 x 或 y 为零时 $V=\infty$。

在这种情况下，直接给出系统大小展开则是十分繁琐的，而矩方程则表明十分有用，它们可以直接由主方程或福克-普朗克方程导出。到体积倒数的阶数前这两种结果稍有不同。为了简单起见我们使用福克-普朗克方程，使得有

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\langle x\rangle\\ \langle y\rangle\end{bmatrix} = \begin{bmatrix}k_1a\langle x\rangle - k_2\langle xy\rangle\\ k_2\langle xy\rangle - k_3\langle y\rangle\end{bmatrix} \tag{7.6.31}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\langle x^2\rangle\\ \langle xy\rangle\\ \langle y^2\rangle\end{bmatrix} = \frac{1}{\mathrm{d}t}\begin{bmatrix}\langle 2x\mathrm{d}x + \mathrm{d}x^2\rangle\\ \langle x\mathrm{d}y + y\mathrm{d}x + \mathrm{d}x\mathrm{d}y\rangle\\ \langle 2y\mathrm{d}y + \mathrm{d}y^2\rangle\end{bmatrix} \tag{7.6.32}$$

$$= \begin{bmatrix}2k_1a\langle x^2\rangle - 2k_2\langle x^2y\rangle + k_1a\langle x\rangle + k_2\langle xy\rangle\\ k_2\langle x^2y - y^2x\rangle + (k_1a - k_3 - k_2)\langle xy\rangle\\ 2k_2\langle xy^2\rangle - 2k_3\langle y^2\rangle + k_2\langle xy\rangle + k_3\langle y\rangle\end{bmatrix} \tag{7.6.33}$$

知道了系统大小展开有效，就意味着我们知道所有的相关与方差跟均值相差是 $1/\Omega$ 阶大小的量。

$$x = \langle x\rangle + \delta x$$
$$y = \langle y\rangle + \delta y \tag{7.6.34}$$

并且只保留最低阶。注意到 $\langle \mathrm{d}x^2\rangle$，$\langle \mathrm{d}x\mathrm{d}y\rangle$ 及 $\langle \mathrm{d}y^2\rangle$ 产生的项是比其它阶项都小的 Ω 的一阶项，我们有

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\langle x\rangle\\ \langle y\rangle\end{bmatrix} = \begin{bmatrix}k_1a\langle x\rangle - k_2\langle x\rangle\langle y\rangle\\ k_2\langle x\rangle\langle y\rangle - k_3\langle x\rangle\langle y\rangle\end{bmatrix} + \begin{bmatrix}-k_2\langle \delta x\delta y\rangle\\ k_2\langle \delta x\delta y\rangle\end{bmatrix} \tag{7.6.35}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\langle\delta x^2\rangle\\ \langle\delta x\delta y\rangle\\ \langle\delta y^2\rangle\end{bmatrix}=\begin{bmatrix}k_1a\langle x\rangle+k_2\langle x\rangle\langle y\rangle\\ -k_2\langle x\rangle\langle y\rangle\\ k_2\langle x\rangle\langle y\rangle+k_3\langle y\rangle\end{bmatrix}$$

$$+\begin{bmatrix}2k_1a-2k_2\langle y\rangle & -2k_2\langle x\rangle & 0\\ k_2\langle y\rangle & k_1a-k_3+k_2(\langle x\rangle-\langle y\rangle) & -k_2\langle x\rangle\\ 0 & 2k_2\langle y\rangle & 2k_2\langle x\rangle-2k_3\end{bmatrix}$$

$$\times\begin{bmatrix}\langle\delta x^2\rangle\\ \langle\delta x\delta y\rangle\\ \langle\delta y^2\rangle\end{bmatrix} \tag{7.6.36}$$

我们注意到均值精确到最低阶是服从确定性方程的，但是到再高一阶，则会出现含 $\langle\delta x\delta y\rangle$ 的项。所以，让我们选取一个简化了的即

$$k_1a=k_3=1,\quad k_2=\alpha \tag{7.6.37}$$

情况来看(这总是可以用改变变量尺度来做到的)，再用简化记号

$$\langle x\rangle\to x,\ \langle y\rangle\to y,\ \langle\delta x^2\rangle\to f,\ \langle\delta x\delta y\rangle\to g,\ \langle\delta y^2\rangle\to h \tag{7.6.38}$$

则得到

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}x-\alpha xy\\ \alpha xy-y\end{bmatrix}+\begin{bmatrix}-\alpha g\\ \alpha g\end{bmatrix}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}f\\ g\\ h\end{bmatrix}=\begin{bmatrix}x+\alpha xy\\ -\alpha xy\\ \alpha xy+y\end{bmatrix}$$

$$+\begin{bmatrix}2-2\alpha y & -2\alpha x & 0\\ \alpha y & \alpha(x-y) & -\alpha x\\ 0 & 2\alpha y & 2\alpha x-2\end{bmatrix}\begin{bmatrix}f\\ g\\ h\end{bmatrix} \tag{7.6.39}$$

我们可以尝试在平稳态下解出这些方程。注意到 f、g、h 是比 x 和 y 小 Ω^{-1} 的一个因子，这就要求 α 是 Ω^{-1} 阶的[这也可由标度要求(7.5.29)得到]。因此 α 是个小量。到最低阶，我们有

$$x_s=y_s=1/\alpha \tag{7.6.40}$$

但是，此时 f、g、h 在平稳状态下的方程就成了不相容的

$$\begin{bmatrix} 2g \\ h-f \\ -2g \end{bmatrix} = \begin{bmatrix} 2/\alpha \\ -1/\alpha \\ 2/\alpha \end{bmatrix} \tag{7.6.41}$$

于是这个方法不产生平稳态。换一个方法可在平稳态下解出(7.6.39)全部方程。

经一些运算和整理后可得

$$x_s = y_s$$

$$g_s = \alpha^{-1}(x_s - \alpha x_s^2) \tag{7.6.42}$$

所以

$$f_s = x_s(-2\alpha x_s^2 + x_s(2-\alpha) - 1/(2-2\alpha x_s) \tag{7.6.43}$$

$$h_s = x_s(-2\alpha x_s^2 + x_s(2+\alpha) + 1)/(2-2\alpha x_s)$$

此外，g_3 的方程还给出了

$$-\alpha x_s^2 + \alpha x_s(f_s - h_s) = 0 \tag{7.6.44}$$

这些方程给出了 x_s, y_s 等等解为

$$\left.\begin{aligned} x_s = y_s &= \frac{1}{2} \\ f_s &= \frac{1}{2}\,\alpha/(\alpha-2) \\ g_s &= \frac{1}{4}\,(2-\alpha)/\alpha \\ h_s &= -1/(\alpha-2) \end{aligned}\right\} \tag{7.6.45}$$

以后，对于使得此方法有效的小值 α，上式导致 f_s 为负值，而按定义它却是正的。这样，就不存在平稳解。

再用近似

$$x_s = y_s = \frac{1}{\alpha}$$

则 f, g, h 的方程即可容易解出。我们发现，假设了系统初始时具有零方差和相关后，就有

$$f(t) = \frac{1}{2\alpha}(\cos 2t - 1) + \frac{2t}{\alpha}$$

$$g(t) = -\frac{1}{2\alpha}\sin 2t$$

$$h(t) = -\frac{1}{2\alpha}(\cos 2t - 1) + \frac{2t}{\alpha} \tag{7.6.46}$$

要注意 $f(t)$ 及 $h(t)$ 事实上总是正的和稳定增加的。这个解仅仅在一短时间内有效，因为 $g(t)$ 的增加终究要产生一个与时间相依的均值。

7.6.4 母函数方程

在组合动力学情况下，相对简单的微分方程可以由母函数

$$G(s,t) = \sum_x (\prod_a s_a^{x_a}) P(x, t) \tag{7.6.47}$$

导出。我们看到

$$\partial t G(s, t) = \partial_t^+ G(s, t) + \partial_t^- G(s, t) \tag{7.6.48}$$

其中两项对应于主方程的 t^+ 及 t^- 部分。这样，

$$\partial_t^+ G(s, t) = \sum_{A,x} k_A^+ \left\{ \prod_a \left[\frac{(x_a - r_a^A)!}{(x_a - r_a^A - N_a^A)!} s_a^{x_a} \right] P(x - r^A, t) - \prod_a \left[\frac{x_a!}{(x_a - N_a^A)!} s_a^{x_a} \right] P(x, t) \right\} \tag{7.6.49}$$

在第一项中，把变量改变为 $x - r^A$，并把它称为 x，则

$$\partial_t^+ G(s, t) = \sum_{A,x} k_A^+ \left[\prod_a \frac{x_a!}{(x_a - N_a^A)!} s_a^{x_a + r_a^A} - \prod_a \frac{x_a!}{(x_a - N_a^A)!} s_a^{x_a} \right] P(x, t) \tag{7.6.50}$$

注意到

$$\prod_a \frac{s_a^{x_a} x_a!}{(x_a - N_a^A)!} = \prod_a (\partial_a^{N_a^A} s_a^{x_a}) s_a^{N_a^A} \tag{7.6.51}$$

以及

$$\prod_a \frac{s_a^{x_a + r_a^A} x_a!}{(x_a - N_a^A)!} = \prod_a (\partial_a^{N_a^A} s_a^{x_a}) s_a^{M_a^A} \tag{7.6.52}$$

故

$$\partial_t^+ G(s, t) = \sum_A k_A^+ \left(\prod_a s_a^{M_a^A} - \prod_a s_a^{N_a^A} \right) \partial_a^{N_a^A} G(s, t) \tag{7.6.53}$$

类似地还可导出 $\partial_t^- G(s, t)$ 的公式，把这两个公式结合在一起就得到

$$\partial^r G(s,\ t) = \sum_A \left(\prod_a s_a^{M_a^A} - \prod_a s_a^{N_a^A}\right) \times \left(k_A^+ \prod_a \partial_a^{N_a^A} - k_A^- \prod_a \partial_a^{M_a^A}\right) G(s,\ t) \tag{7.6.54}$$

这就是母函数微分方程的一般公式，下面给出两个例子。

a) 一个完全可解模型

反应：(A, B, C 保持不变)

$$A + X \xrightarrow{k_2} 2X + D \quad N^1 = 1,\ M^1 = 2;\ r^1 = 1$$

$$k_1^+ = k_2 A \tag{7.6.55}$$

$$k_1^- = 0$$

$$B + X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} C \qquad N^2 = 1,\ M^2 = 0;\ r^2 = -1$$

$$k_2^+ = k_1 B \tag{7.6.56}$$

$$k_2^- = k_3 C$$

因而由(7.6.54)知母函数方程是

$$\partial_t G = (s^2 - s)(k_2 A \partial_s G) + (1 - s)(k_1 B \partial_s G - k_3 C G) \tag{7.6.57}$$

现用其特征线求解。令

$$k_1 B = \beta,\ k_2 A = \alpha,\ k_3 C = \gamma \tag{7.6.58}$$

特征线是

$$\frac{dt}{1} = \frac{ds}{(1-s)(\beta - \alpha s)} = \frac{dG}{\gamma(1-s)G} \tag{7.6.59}$$

它具有解

$$\left(\frac{1-s}{\beta - \alpha s}\right) e^{(\alpha - \beta)t} = u \tag{7.6.60}$$

$$(\beta - \alpha s)^{\gamma/\alpha} G = v \tag{7.6.61}$$

通解可写为 $v = F(u)$，也就是说

$$G = (\beta - \alpha s)^{-\gamma/\alpha} F\left\{e^{(\alpha - \beta)t}\left(\frac{1-s}{\beta - \alpha s}\right)\right\} \tag{7.6.62}$$

由此可得到各种时间相依解。条件概率 $P(x,\ t|y,\ 0)$ 则由初值条件得到，

$$G_y(s,\ 0) = s^y \tag{7.6.63}$$

$$\Longrightarrow F(z)=(1-\beta z)^{\nu}(1-\alpha z)^{-\gamma/\alpha-\nu}(\beta-\alpha)^{\gamma/\alpha} \tag{7.6.64}$$

$$\boxed{\begin{aligned}\Longrightarrow G_{\nu}(s,\ t)=\lambda^{\gamma/\alpha}&[\beta(1-\mathrm{e}^{-\lambda t})-s(\alpha-\beta\mathrm{e}^{-\lambda t})]^{\nu}\\ &\times[(\beta-\alpha\mathrm{e}^{-\lambda t})-\alpha s(1-\mathrm{e}^{-\lambda t})]^{-\gamma/\alpha-\nu}\end{aligned}} \tag{7.6.65}$$

(其中 $\lambda=\beta-\alpha$)。当 $t\to\infty$ 时，仅当 $\beta>\alpha$ 时存在平稳态，它是

$$G_{\nu}(s,\ \infty)=(\beta-\alpha s)^{-\gamma/\alpha}(\beta-\alpha)^{\gamma/\alpha} \tag{7.6.66}$$

$$\Longrightarrow P_{s}(x)=\frac{\Gamma(x+\gamma/\alpha)}{\Gamma(\gamma/\alpha)x!}(\alpha/\beta)^{x}(\beta-\alpha)^{\gamma/\alpha} \tag{7.6.67}$$

只要注意到

$$\partial_{s}G(s,\ t)\big|_{s=1}=\langle x(t)\rangle \tag{7.6.68}$$

$$\partial_{s}G(s,\ t)\big|_{s=1}=\langle x(t)\,[x(t)-1]\rangle$$

我们还可由母函数方程导出矩方程。沿此做下去，就有

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x(t)\rangle=(k_{2}A-k_{1}B)\langle x(t)\rangle+k_{3}C \tag{7.6.69}$$

以及

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle x(t)\,[x(t)-1]\rangle=&\,2(k_{2}A-k_{1}B)\langle x(t)\,[x(t)-1]\rangle\\&+2k_{2}A\langle x(t)\rangle+2k_{3}C\langle x(t)\rangle\end{aligned} \tag{7.6.70}$$

倘若

$$k_{2}A<k_{1}B,\quad \text{即，}\ \alpha<\beta$$

则这些方程有稳定的平稳解。此时，平稳解均值及方差为

$$\langle x\rangle_{s}=k_{3}C/(k_{1}B-k_{2}A) \tag{7.6.71}$$

$$\mathrm{var}\{x\}_{s}=k_{1}k_{3}BC/(k_{2}A-k_{1}B)^{2} \tag{7.6.72}$$

这个模型是在核反应中所发生的一些过程的简化表示。这里的X是一中子。第一个反应表示着一中子因被A吸收而发生的裂变，它产生了剩余物D及另外二个中子。第二个反应表示着中子的吸收和由裂变之外的其它方法而来的中子产生。

当 $k_{2}A$ 接近 $k_{1}B$ 时，我们就接近于一个临界状态，中子的吸收和产生数几乎相等。对于 $k_{2}A>k_{1}B$，则出现了一个爆炸性的

链锁反应。注意当接近一个临界点时$\langle x_s\rangle$及$\mathrm{var}\{x\}$，两者都变得十分大。事实上

$$\frac{\mathrm{var}\{x_s\}}{\langle x\rangle_s}=\frac{k_1B}{k_1B-k_2A}\longrightarrow\infty \tag{7.6.73}$$

这样，在接近临界点的地方$\langle x_s\rangle$存在着非常大的涨落。

还要注意系统具有关于均值的线性方程，而且系统是马尔可夫过程，所以第3.7.4节的方法(回归定理)指出

$$\langle x(t),\ x(0)\rangle_s=\exp[(k_2A-k_1B)t]\,\mathrm{var}\{x\}。 \tag{7.6.74}$$

所以当临界点已迫近了时，其涨落慢得越来越看不见，也就是说，时间相关函数随着时间变化衰减得十分慢。

b) 化学反应 $X_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} X_2$

一个反应

$$N=\begin{bmatrix}1\\0\end{bmatrix},\quad M=\begin{bmatrix}0\\1\end{bmatrix},\quad r=\begin{bmatrix}-1\\1\end{bmatrix}$$

$$k_1^+=k_1,\qquad k_1^-\quad k_2 \tag{7.6.75}$$

$$\partial_t G(s_1,\ s_2,\ t)=(s_2-s_1)(k_1\partial_{s_1}-k_2\partial_{s_2})G(s_1,\ s_2,\ t)$$

可以由特征方法解出。其母函数是

$$\frac{\mathrm{d}t}{1}=-\frac{\mathrm{d}s}{k_1(s_2-s_1)}=\frac{\mathrm{d}s_2}{k_2(s_2-s_1)} \tag{7.6.76}$$

之解的任意函数。两个积分是

$$k_2\mathrm{d}s_1+k_1\mathrm{d}s_2=0\Longrightarrow k_2s_1+k_1s_2-v \tag{7.6.77}$$

$$(k_1+k_2)\mathrm{d}t=\frac{\mathrm{d}(s_2-s_1)}{s_2-s_1}$$

$$\Longrightarrow (s_2-s_1)\mathrm{e}^{-(k_1+k_2)t}=u \tag{7.6.78}$$

的解。所以

$$G(s_1,\ s_2,\ t)=F[k_2s_1+k_1s_2,\ (s_2-s_1)\mathrm{e}^{-(k_1+k_2)t}] \tag{7.6.79}$$

(泊松分布化的)初始条件

$$G(s_1,\ s_2,\ 0)=\exp[\alpha(s_1-1)+\beta(s_2-1)] \tag{7.6.80}$$

给出了泊松分布解。

$$G(s_1, s_2, t) = \exp\left\{\frac{k_2\beta - k_1\alpha}{k_1 - k_2}(s_2 - s_1)e^{-(k_1+k_2)t}\right.$$

$$\left. + \frac{\alpha - \beta}{k_1 + k_2}[k_1(s_2 - 1) + k_2(s_1 - 1)]\right\} \quad (7.6.81)$$

在这种情况下，因 $x + y$ 是守恒的，平稳解不唯一。由 (7.6.79) 知一般平稳解为

$$G(s_1, s_2, \infty) = F(k_2 s_1 + k_1 s_2, 0) \quad (7.6.82)$$

的形式，于是

$$k_1^n \frac{\partial^n G}{\partial s_1^n} = k_2^n \frac{\partial^n G}{\partial s^n} \quad (7.6.83)$$

而令 $s_1 = s_2 = 1$ 之后，上式即推得

$$k_1^n \langle x_1^n \rangle_t = k_2^n \langle x_2^n \rangle_t \quad (7.6.84)$$

7.7 泊松表示[7.10]

这是一种特别精致的技巧，它产生与(7.5.9)的化学主方程等价的福克-普朗克方程。

我们假设可把 $P(x, t)$ 展开成多变量不相关的泊松分布的叠加：

$$P(x, t) = \int d\alpha \prod_a \frac{e^{-\alpha_a}\alpha_a^{x_a}}{x_a!} f(\alpha, t) \quad (7.7.1)$$

这意味着母函数 $G(s, t)$ 可表示为

$$G(s, t) = \int d\alpha \exp\left[\sum_a (s_a - 1)\alpha_a\right] f(\alpha, t) \quad (7.7.2)$$

把它代入母函数方程(7.6.54)，就得到

$$\partial_t G(s, t) = \sum_A' \int d\alpha \left\{\left[\prod_a \left(\frac{\partial}{\partial \alpha_a} + 1\right)^{M_a^A}\right.\right.$$

$$\left. - \prod_a \left(\frac{\partial}{\partial \alpha_a} + 1\right)^{N_a^A}\right] \times \left(k_A^+ \prod_a \alpha_a^{N_a^A} - k_A^- \prod_a \alpha_a^{M_a^A}\right)$$

$$\left. \times \exp\left[\sum_a (s_a - 1)\alpha_a\right]\right\} f(\alpha, t) \quad (7.7.3)$$

在进行分部积分时，去掉曲面项并最后令指数的系数相等，就得到

$$\frac{\partial f(\alpha,\ t)}{\partial t} = -\sum_A \left[\prod_a \left(1 - \frac{\partial}{\partial \alpha_a}\right)^{M_a^A} - \prod_a \left(1 - \frac{\partial}{\partial \alpha_a}\right)^{N_a^A}\right] \times \left[k_A^+ \prod_a \alpha_a^{N_a^A} - k_A^- \prod_a \alpha_a^{M_a^A}\right] f(\alpha,\ t) \tag{7.7.4}$$

a) 双分子反应系统的福克-普朗克方程

如果像实际化学反应中通常的情况那样

$$\sum_a M_a^A \leqslant 2$$
$$\sum_a N_a^A \leqslant 2 \tag{7.7.5}$$

这表明反应中最多只有分子对参加反应，则方程是福克-普朗克形式的。因此方程可作如下简化。定义流

$$J_A(\alpha) = k_A^+ \prod_a \alpha_a^{N_a^A} - k_A^- \prod_a \alpha_a^{M_a^A} \tag{7.7.6}$$

和漂移

$$A_a[J(\alpha)] = \sum_A r_a^A J_A(\alpha) \tag{7.7.7}$$

并且用

$$B_{ab}[J(\alpha)] = \sum_A J_A(\alpha)\,(M_a^A M_b^A - N_a^A N_b^A - \delta_{a,b} r_a^A) \tag{7.7.8}$$

来定义扩散矩阵元素，则泊松表示的福克-普朗克方程是

$$\frac{\partial f(\alpha,\ t)}{\partial t} = -\sum_a \frac{\partial}{\partial \alpha_a} \{A_a[J(\alpha)]\, f(\alpha,\ t)\} + \frac{1}{2} \sum_{a,b} \frac{\partial^2}{\partial \alpha_a \partial \alpha_b} \{B_{ab}[J(\alpha)]\, f(\alpha,\ t)\} \tag{7.7.9}$$

还要注意，如果我们用在7.5.3节 (7.5.29) 中所给参数对体积的明显相依关系，并且定义

$$\eta_a = \alpha_a / V \tag{7.7.10}$$
$$\varepsilon = V^{-1/2} \tag{7.7.11}$$

和 $F(\eta,\ t)$ 是涉及变量 η 的拟概率，那末变量 η 的福克-普朗克方程取形式

$$\frac{\partial F(\eta, t)}{\partial t} = -\sum_a \frac{\partial}{\partial \eta_a}[\hat{A}_a(\eta)F(\eta, t)] + \frac{\varepsilon^2}{2}\sum_{a,b}\frac{\partial^2}{\partial\eta_a\partial\eta_b}[\hat{B}_{a,b}(\eta)F(\eta, t)] \quad (7.7.12)$$

其中

$$\hat{A}_a(\eta) = \sum_A r_a^A \hat{J}_A(\eta) \quad (7.7.13a)$$

$$\hat{J}_a(\eta) = k_A^+ \prod_a \eta_a^{N_a^A} - k_A^- \prod_a \eta_a^{M_a^A} \quad (7.7.13b)$$

$$\hat{B}_{ab}(\eta) = \sum_A \hat{J}_A(\eta)(M_a^A M_b^A - N_a^A N_b^A - \delta_{a,b} r_a^A) \quad (7.7.13c)$$

在这个形式中，我们看见了按 $V^{-\frac{1}{2}}$ 所作的系统大小展开，恰好完全对应于福克-普朗克方程 (7.7.12) 按 η 所作的小噪声展开。对于这样的生灭主方程来说，这个方法在技巧上远比一个直接的系统大小展开简单得多。

b) 单分子反应

若对于所有的 A

$$\sum_a M_a^A \leqslant 1 \quad \text{并且} \quad \sum_a N_a^A \leqslant 1$$

则易验证(7.7.13)中的扩散系数 $\hat{B}_{ab}(\eta)$ 为零，而且我们有了刘维方程。初始取泊松分布 $P(x, t_0)$，则它对应着 δ 函数 $F(\eta, t_0)$，而且由该刘维方程产生的时间发展将产生出一个 δ 函数解 $\delta(\eta - \eta(t))$，其中 $\eta(t)$ 是

$$d\eta/dt = A(\eta)$$

的解。这意味着 $P(x, t)$ 将保持泊松分布的形式，均值等于 $\eta(t)$。于是，我们导出了一般性结果，即对任何单分子反应系统来说存在着传播的多元泊松分布解，也存在没有泊松分布解的情况——这些解对应着不是 δ 函数初始 $F(\eta, t_0)$。

c) 例子

作为例子，我们考察对反应

i) $A + X \underset{k_4}{\overset{k_2}{\rightleftharpoons}} 2X$

ii) $B+X \underset{k_3}{\overset{k_1}{\rightleftarrows}} C$ (7.7.14)

由于

$$N^1=1,\quad M^1=2,\quad k_1^+=k_2A,\quad k_1^-=k_4$$
$$N^2=1,\quad M^2=0,\quad k_2^+=k_1B,\quad k_2^-=k_3C$$

故(7.7.4)具有形式

$$\frac{\partial f}{\partial t}=\left[\left(1-\frac{\partial}{\partial x}\right)^2-\left(1-\frac{\partial}{\partial x}\right)\right](k_2A\alpha-k_4\alpha^2)f + \left[1-\left(1-\frac{\partial}{\partial\alpha}\right)\right](k_1B\alpha-k_3C)f \tag{7.7.15}$$

$$\frac{\partial f}{\partial t}=\left\{-\frac{\partial}{\partial\alpha}[k_3C+(k_2A-k_1B)\alpha-k_4\alpha^2] + \frac{\partial^2}{\partial\alpha^2}[k_2A\alpha-k_4\alpha^2]\right\}f \tag{7.7.16}$$

只要 $k_2A\alpha-k_4\alpha^2>0$，它便是福克-普朗克形式的。此外，在矩之间存在着简单的关系，它(在单变量情况下)取如下形式：

$$\boxed{\langle x^r\rangle_f=\sum_x\int d\alpha[x(x-1)\cdots(x-r+1)\frac{e^{-\alpha}\alpha^x}{x!}f(\alpha) =\int d\alpha\alpha^r f(\alpha)\equiv\langle\alpha^r\rangle}$$

(7.7.17)

这由泊松分布的阶乘矩导出（参见第 2, 8, 3 节）。但是 $f(\alpha)$ 不是概率分布，至少就它所定义的方式来看不能保证它应该是概率。这是很清楚的，因为一些泊松分布的任何正叠加必然有着至少像泊松分布那样大小的方差。因此任何概率分布，只要其方差是小于泊松分布之方差，也就不能够被一个正的 $f(\alpha)$ 来表示。

通过广义函数的表示总是可能的，至少形式上可以做到，因为只要定义

$$f_\nu(\alpha)=(-1)^\nu\delta^\nu(\alpha)e^\alpha \tag{7.7.18}$$

那末

$$\int d\alpha f_\nu(\alpha)e^{-\alpha}\alpha^x/x!=\int d\alpha\alpha^x\left(-\frac{d}{d\alpha}\right)^\nu\delta(\alpha)/x! \tag{7.7.19}$$

并且由分部积分

$$=\delta_{x,y} \tag{7.7.20}$$

它意味着我们可以写出

$$P(x)=\int \mathrm{d}\alpha(\mathrm{e}^{-\alpha}\alpha^x/x!)\left[\sum_y(-1)^y P(y)\delta^y(\alpha)\mathrm{e}^{\alpha}\right] \tag{7.7.21}$$

可以在形式的意义下对任何 $P(x)$ 总可以找到一个 $f(\alpha)$。

事实上刚才给出的这种较奇异的形式通常是不发生的，这是由于，譬如说，我们可以找到一个象位势解(精确到一个归一化常数)因子那样的福克-普朗克方程(7.7.16)的平稳解

$$f_s(\alpha)=\mathrm{e}^{\alpha}(k_2A-k_4\alpha)^{(k_1B/k_4-k_3C/k_2A-1)}\alpha^{k_3C/k_2A-1} \tag{7.7.22}$$

这是一个相对光滑的函数。但是只有当 $f_s(\alpha)$ 是正值或零值的而且是可归一的时候，才可能作为一个概率分布来理解。

如果我们定义

$$\delta=(k_1B/k_4-k_3C/k_2A) \tag{7.7.23}$$

则只要

$$\delta>0$$

$$k_3>0 \tag{7.7.24}$$

$f_s(\alpha)$ 就是在区间 $(0,\ k_2A/k_4)$ 上可归一化的。显然，由定义知 k_3 必是正的。

必须进一步验证在导出福克-普朗克方程(7.7.4)中用到的分部积分使得在这些条件下曲面项是零。对于一个区间 $(a,\ b)$ 来说，反应(7.7.14)的情况下所产生的曲面项可以写出为

$$\begin{aligned}&[\{(k_2A\alpha-k_4\alpha^2-k_1B\alpha+k_3C)f\\&\quad-\partial_\alpha[(k_2\alpha-k_4\alpha^2)f]\}\langle\mathrm{e}^{(s-1)\alpha}\rangle]_b^a\\&\quad+[(k_2\alpha-k_4\alpha^2)f[(s-1)\mathrm{e}^{(s-1)\alpha}]]\end{aligned} \tag{7.7.25}$$

因为在第二行有个外加因子 $(s-1)$，所以每一行都必然分别为零。容易验证，只要 δ，k 两者都大于零，在区间 $(0,\ k_2A/k_4)$ 上，对于(7.7.22)所选定的 f，在此区间的每个端点每项都是零。

当 k_3 与 δ 两者都为零时，我们就有真正的福克-普朗克方程，它等价于随机微分方程

$$\mathrm{d}\alpha = [k_3C + (k_2A - k_1B)\alpha - k_4\alpha^2]\mathrm{d}t + \sqrt{2}(k_2A\alpha - k_4\alpha^2)\,\mathrm{d}W(t) \tag{7.7.26}$$

其运动发生在值域$(0,\ k_2A/k_4)$上，而且两个边界满足入口边界的判据，这意味着运动不可能离开值域 $(0,\ k_2A/k_4)$（见第 5.2.1 节）。

如果(7.7.24)中的某一个不成立，就会看到漂移向量会使运动跑出到区间 $(0,\ k_2A/k_4)$之外。例如，在靠近 $\alpha=0$ 的地方我们有

$$\mathrm{d}\alpha \sim k_3C\mathrm{d}t \tag{7.7.27}$$

并且若 k_3C 是负的，则 α 将取负值。此时，在(7.7.26)中 $\mathrm{d}W(t)$ 的系数变成了虚数，如果没有进一步的解释就不可能理解了。

当然，看作为复变量

$$\alpha = \alpha_x + i\alpha_y \tag{7.7.28}$$

的随机微分方程的话，此随机微分方程完全是有意义的，而且实际上就是两个变量 α_x 及 α_y 的一对随机微分方程。但是其相应的福克-普朗克方程也就不再是单变量方程 (7.7.16) 而是一个两变量福克-普朗克方程了。我们可以通过下面就要研究的各种泊松表示来导出这种福克-普朗克方程。

7.7.1 泊松表示的种类

让我们来考虑单变量情况，记

$$P(x) = \int_{\mathscr{D}} \mathrm{d}\mu(\alpha)\,(\mathrm{e}^{-\alpha}\alpha^x/x!)\,f(\alpha) \tag{7.7.29}$$

那末 $\mu(\alpha)$ 是一个测度，我们将证明它有三种选择方法，所有这三种方法都导致有用的表示，$\mathscr{D}$是积分区域，$\mathscr{D}$依赖于测度的选择可以取各种形式。

7.7.2 实泊松表示

这里我们选取

$$\mathrm{d}\mu(\alpha) = \mathrm{d}\alpha \tag{7.7.30}$$

并且$\mathscr{D}$是直线段，如上面例子中已提到的，这个表示不一定存在，但是当存在时，则借助于福克-普朗克方程的简单解释是可能的。

7.7.3 复泊松表示

这里

$$\mathrm{d}\mu(\alpha) = \mathrm{d}\alpha \tag{7.7.31}$$

并且$\mathscr{D}$是复平面中的围道。我们可以证明在某种限制性条件下这种情况是存在，因为我们可以不用形式(7.7.18)而代之选用

$$f_\nu(\alpha) = \frac{\nu!}{2\pi i}\,\alpha^{-\nu-1}\mathrm{e}^{\alpha} \tag{7.7.32}$$

并且C是围绕着原点的围道。这意味着

$$P_\nu(x) = \frac{1}{2\pi i}\oint_C \frac{\nu!}{x!}\,\alpha^{x-\nu-1} = \delta_{x,\nu} \tag{7.7.33}$$

用

$$f(\alpha) = \frac{1}{2\pi i}\sum_\nu P(\nu)\mathrm{e}^{-\alpha}\alpha^{-\nu-1}\nu! \tag{7.7.34}$$

给出$f(\alpha)$之后，借适当的求和就可以用$f(\alpha)$来表示已给的$P(x)$。如果$p(\nu)$是使得$\nu!P(\nu)$对所有的ν都有界，则对此级数有着一个有限的收敛半径，在它外面$f(\alpha)$是解析的。选取C在这个收敛圆之外，我们就可以把积分放入求和之中去而看出$P(x)$是由

$$P(x) = \oint_C \mathrm{d}\alpha(\mathrm{e}^{-\alpha}\alpha^x/x!)\,f(\alpha) \tag{7.7.35}$$

给出

a) 例子：反应

(1) $A + X \rightleftharpoons 2X$

(2) $B + X \rightleftharpoons C$

我们使用第7.7节的记号并按δ的大小区别对待三种情况。量δ给出了一个测度，它度量着当存在一个稳态时系统(7.7.14)所进行的反应方向。如果$\delta>0$，我们发现，当x取其稳态值时，反应(1)产生着X而反应(2)则消耗X。当$\delta=0$，则两个反应都分别平衡——于是我们有了化学平衡。当$\delta<0$，反应(1)消耗X而反

应(2)则产生X。

i) $\delta>0$：按照(7.7.24)，这是使 $f_s(\alpha)$ 在实区间$(0,\ k_2A/k_4)$上为有效拟概率的条件。在这个范围内扩散系数$(k_2A\alpha-k_4\alpha^2)$是正的。由

$$\alpha=\frac{k_2A-k_1B+[(k_2A-k_1B)^2+4k_3k_4C]^{\frac{1}{2}}}{2k_4} \tag{7.7.36}$$

所给的α的确定性均值在区间$(0,\ k_2A/k_4)$内。因此我们要处理的是真正福克-普朗克方程的情况，并且 $f_s(\alpha)$ 是一在区间两端均为零而峰值点接近确定性稳态的函数。

ii) $\delta=0$：因为两个反应都分别是平衡的，我们可以有个泊松稳态。我们注意，在此情况下 $f_s(\alpha)$ 在 $\alpha=k_2A/k_4$ 处有一极点，并且我们要选择 α 的值域为复平面上包含这个极点在内的围道。因为这是个闭围道，从偏积分中不产生边界上的项，而且由选这类泊松表示来给出的 $P_s(x)$ 显然满足稳态主方程。用留数分析之后，我们就会看到

$$P_s(x)=e^{-\alpha_0}\alpha_0^x/x! \tag{7.7.37}$$

其中

$$\alpha_0=k_2A/k_4$$

iii) $\delta<0$：当 $\delta<0$ 时我们会看到一些非常有趣的特征：现在稳态解(7.7.22)不再满足条件$\delta>0$。但是，如果α的值域选为复平面上的一围道C（图 7.3），并且使用复泊松表示，则被构造为

$$P_s(x)=\int_C d\alpha f_s(\alpha)\frac{e^{-\alpha}\alpha^x}{x!} \tag{7.7.38}$$

的 $P_s(x)$ 就是主方程的一个解。确定性稳态出现在直线上，在$\alpha=k_2A/k_4$ 的奇异点右边的一点上，并且均值、矩等等的渐近估值也可通过选取C通过鞍点出现的地方而得到，这样做，我们就会看到定义为

$$\mathrm{var}\langle x\rangle=\langle\alpha^2\rangle-\langle\alpha\rangle^2 \tag{7.7.39}$$

是负的，所以

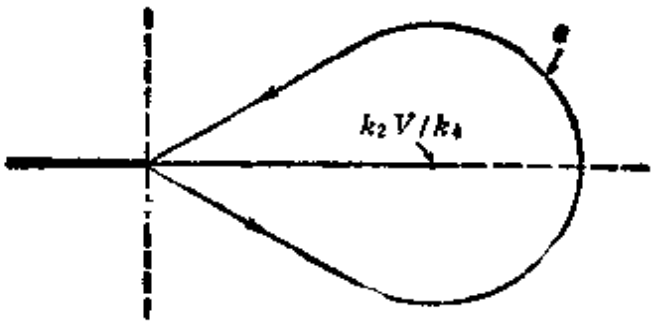


图 7.3　计算(7.7.38)的复平面的围道 C

$$\mathrm{var}\langle x\rangle = \langle x^2\rangle - \langle x\rangle^2 = \langle \alpha^2\rangle - \langle\alpha\rangle^2 + \langle\alpha\rangle < \langle x\rangle \qquad (7.7.40)$$

这意味着稳态是比泊松分布要更狭些。最后应该注意的是所有这三种情况都可由围道 C 得到。在 $\delta=0$ 的情况下，从在

$$\alpha = k_2A/k_4$$

的奇异点到 $-\infty$ 的剖线不见了，C 可能畸变为围绕极点的一个简单围道，而若 $\delta>0$，则在 $\alpha=k_2A/k_4$ 的奇性是可积的，以致围道可能萎缩到剖线上，并且积分是作为在值域 $[0,\ k_2A/k_4]$ 上的一个不连续性积分而计算大小的（当 δ 是正整数时，这个论证要求作修正）。

b) 例子：反应 $B \xrightarrow{k_1} X$，$2X \xrightarrow{k_2} A$

其福克-普朗克方程是

$$\frac{\partial f(\alpha,\ t)}{\partial t} = -\frac{\partial}{\partial\alpha}\left[(K_1V - 2K_2V^{-1}\alpha^2)\,f(\alpha,\ t)\right]$$

$$-\frac{\partial^2}{\partial\alpha^2}\left[(K_2V^{-1}\alpha^2)\,f(\alpha,\ t)\right] \qquad (7.7.41)$$

其中 $K_1V = K_1B$，$K_2V^{-1} = k_2$ 并且 V 是系统的体积。注意上面的福克-普朗克方程中扩散系数在整个实直线上都是负的。

(7.7.41)的(精确到归一化因子的)位势解是

$$f(\alpha) = \alpha^{-2}\exp(2\alpha + aV^2/\alpha) \qquad (7.7.42)$$

其中 $a=2K_2/K_1$ 并且对 α 的积分沿着含原点的一闭围道进行。当然，原则上讲存在着完整地解出此平稳福克-普朗克方程而得到的其他解，但是只有此位势解是单值的，而且允许我们选取一个能进

行偏积分的合用围道。

于是在令 $\alpha = \eta V$ 后就得到

$$\langle x^r \rangle_f = \frac{V^r \oint d\eta e^{V(2\eta - a/\eta)} \eta^{r-2}}{\oint d\eta e^{V(2\eta + a/\eta)} \eta^{-2}} \tag{7.7.43}$$

函数 $(2\eta + a/\eta)$ 在确定性稳态点并不取极大值。事实上它在确定性的稳态

$$\eta = +(a/2)^{\frac{1}{2}}$$

有一极小值。但是在复 η 平面中这个点是一鞍点，并提供了对积分的一控制分布。

这样一来，(7.7.41)中的负扩散系数本身由在确定性稳态下产生一个鞍点而它是反射的。其结果造成了 X 的方差比 $\langle \alpha \rangle$ 要小的。

由(7.7.43)所有稳态矩可精确算出，结果是

$$\langle x^r \rangle_f = \left[V \left(\frac{a}{2} \right)^{\frac{1}{2}} \right]^r \frac{I_{r-1}(2(2a)^{\frac{1}{2}} V)}{I_1(2(2a)^{\frac{1}{2}} V)} \tag{7.7.44}$$

其中 $I_r(2(2a)^{\frac{1}{2}}V)$ 是修正贝塞尔函数。利用对 $I_r(2(2a)^{\frac{1}{2}}V)$ 的大变量展开，我们得到

$$\begin{aligned} \langle x \rangle &= V \left(\frac{a}{2} \right)^{\frac{1}{2}} + \frac{1}{8} + O\left(\frac{1}{V} \right) \\ \mathrm{var}\{x\} &= \frac{3}{4} V \left(\frac{a}{2} \right)^{\frac{1}{2}} - \frac{1}{16} + O\left(\frac{1}{V} \right) \end{aligned} \tag{7.7.45}$$

这些渐近结果也可借直接应用最速下降法于(7.7.43)而得到。一般说，在明确地显示了参数对体积的相依关系后这类展开总是可以做到的。

c) 优点的总结

复泊松表示产生解析形式的平稳解，使渐近或精确的方法都易应用，在与时间相依的情况下它就不是很有用处了。不过，最大的好处出现于量子力学系统中，此时我们可以把类似的技巧用于

复P表示，它可以给出其它方法所无法提取的信息。这将在第10章中说明。

7.7.4 正泊松表示

这里我们选取α是一复变量$\alpha_x + i\alpha_y$

$$d\mu(\alpha) = d^2\alpha = d\alpha_x d\alpha_y \tag{7.7.46}$$

并且$\mathscr{D}$是整个复平面，在第10.6.3节我们证明了对任何$P(\alpha)$，存在$f(\alpha)$使得

$$P(\alpha) = \int d^2\alpha (e^{-\alpha}\alpha^x/x!) f(\alpha) \tag{7.7.47}$$

于是正的泊松表示总是存在的，但不是唯一的。例如，选取

$$f_p(\alpha) = (2\pi\sigma^2)^{-1} \exp(-|\alpha-\alpha_0|^2/2\sigma^2) \tag{7.7.48}$$

并注意到若$g(\alpha)$是α的任何解析函数，则可写出

$$g(\alpha) = g(\alpha_0) + \sum_{n=1}^{\infty} g^{(n)}(\alpha_0)(\alpha-\alpha_0)^n/n! \tag{7.7.49}$$

使得

$$\int (2\pi\sigma^2)^{-1} d^2\alpha \exp(-|\alpha-\alpha_0|^2/2\sigma^2) g(\alpha) = g(\alpha_0) \tag{7.7.50}$$

这是因为(7.7.50)中的积分对$n \geqslant 1$的项都是零。注意到泊松形式$e^{-\alpha}\alpha^x/x!$本身关于α是解析的，对于任何正值的σ^2

$$P(x) = \int d^2\alpha f_p(\alpha) e^{-\alpha}\alpha^x/x! = e^{-\alpha_0}\alpha_0^x/x! \tag{7.7.51}$$

实际中这个不唯一性与其说是个问题还不如说是个方便之处。

a) 福克-普朗克方程

我们利用泊松概率分布的解析性以及其母函数来导出具有正扩散矩阵的福克-普朗克方程。一个形式为(7.7.9)的福克-普朗克方程是从母函数方程

$$\partial_t G(s, t) = \int d^2\alpha f(\alpha, t) \left(\sum_a A_a \frac{\partial}{\partial \alpha_a} + \sum_{a,b} \frac{1}{2} B_{ab} \frac{\partial^2}{\partial\alpha_a \partial\alpha_b} \right) \exp\left[\sum_a (s_a - 1)\alpha_a\right] \tag{7.7.52}$$

产生出来的. 现在我们把 $\boldsymbol{\alpha}$ 明确地考虑为一复变量

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_x+i\boldsymbol{\alpha}_y \tag{7.7.53}$$

同时记

$$\boldsymbol{A}(\boldsymbol{\alpha})=\boldsymbol{A}_x(\boldsymbol{\alpha})+i\boldsymbol{A}_y(\boldsymbol{\alpha}) \tag{7.7.54}$$

此外,我们还记

$$\boldsymbol{B}(\boldsymbol{\alpha})=\boldsymbol{C}(\boldsymbol{\alpha})\boldsymbol{C}^T(\boldsymbol{\alpha}) \tag{7.7.55}$$

以及

$$\boldsymbol{C}(\boldsymbol{\alpha})=\boldsymbol{C}_x(\boldsymbol{\alpha})+i\boldsymbol{C}_y(\boldsymbol{\alpha}) \tag{7.7.56}$$

为简单起见,我们使用记号

$$\begin{aligned}\partial_a&=\frac{\partial}{\partial\alpha_a}\\ \partial_a^x&=\frac{\partial}{\partial\alpha_{x,a}}\\ \partial_a^y&=\frac{\partial}{\partial\alpha_{y,a}}\end{aligned} \tag{7.7.57}$$

因为 $\exp\left[\sum_a(s_a-1)\alpha_a\right]$ 的解析性,在母函数方程(7.7.52)中我们总是可以作彼此是可交换的选择

$$\partial_a\longleftrightarrow\partial_a^x\longleftrightarrow i\partial_a^y \tag{7.7.58}$$

那末我们对 B_{ab} 代入 (7.7.54) 式并且按照 A 或 C 的相应足标是 x 还是 y 来分别,用 ∂_a^x 或 $-i\partial_a^y$ 代替 ∂_a 则导出

$$\begin{aligned}\partial_t G(\boldsymbol{s},t)=\int \mathrm{d}^2\boldsymbol{\alpha}\, f(\boldsymbol{\alpha},t)\{&\left[\sum_a\left(A_{a;x}\partial_a^x+A_{a;y}\partial_a^y\right)\right.\\&+\frac{1}{2}\sum_{a,b,c}\left(C_{a,c;x}C_{c,b;x}\partial_a^x\partial_b^x+C_{a,c;y}C_{c,b;y}\partial_a^y\partial_b^y\right.\\&\left.\left.+2C_{a,c;x}C_{c,b;y}\partial_a^x\partial_b^y\right)\exp\left[\sum_a(s_a-1)\alpha_a\right]\right\}\end{aligned} \tag{7.7.59}$$

再作分部积分并弃去曲面项,就得到了用 $(\boldsymbol{\alpha}_x,\boldsymbol{\alpha}_y)$ 作变量的福克-普朗克方程

$$\begin{aligned}\partial_t f(\boldsymbol{\alpha},t)=&\left[-\sum_a\left(A_{a;x}\partial_a^x+A_{a;y}\partial_a^y\right)\right.\\&+\frac{1}{2}\sum_{a,b,c}\left(C_{a,c;x}C_{c,b;x}\partial_a^x\partial_b^x+C_{a,c;y}C_{c,b;y}\partial_a^y\partial_b^y\right.\end{aligned}$$

$$+2C_{a,c;x}C_{c',k;y}\partial_a^x\partial_b^y)]\times f(\boldsymbol{\alpha},t) \tag{7.7.60}$$

在维数加倍了的空间中这是个具有半正定扩散的福克-普朗克方程。这是因为对变量$(\boldsymbol{\alpha}_x,\boldsymbol{\alpha}_y)$我们有漂移向量

$$\mathscr{A}(\boldsymbol{\alpha})=[\boldsymbol{A}_x(\boldsymbol{\alpha}),\boldsymbol{A}_y(\boldsymbol{\alpha})] \tag{7.7.61}$$

与扩散矩阵

$$\underset{\sim}{\mathscr{B}}(\boldsymbol{\alpha})=\begin{bmatrix}\underset{\sim}{\boldsymbol{C}}_x\underset{\sim}{\boldsymbol{C}}_x^T & \underset{\sim}{\boldsymbol{C}}_x\underset{\sim}{\boldsymbol{C}}_y^T\\ \underset{\sim}{\boldsymbol{C}}_y\underset{\sim}{\boldsymbol{C}}_x^T & \underset{\sim}{\boldsymbol{C}}_y\underset{\sim}{\boldsymbol{C}}_y^T\end{bmatrix}=\underset{\sim}{\mathscr{C}}(\boldsymbol{\alpha})\underset{\sim}{\mathscr{C}}(\boldsymbol{\alpha})^T \tag{7.7.62}$$

其中

$$\underset{\sim}{\mathscr{C}}(\boldsymbol{\alpha})=\begin{bmatrix}\underset{\sim}{\boldsymbol{C}}_x & 0\\ \underset{\sim}{\boldsymbol{C}}_y & 0\end{bmatrix} \tag{7.7.63}$$

从而$\underset{\sim}{\mathscr{B}}(\boldsymbol{\alpha})$显然是半正定的。

b) 随机微分方程式 (SDE)

对应于漂移和扩散(7.7.61, 62) 有随机微分方程

$$\begin{bmatrix}\mathrm{d}\boldsymbol{\alpha}_x\\ \mathrm{d}\boldsymbol{\alpha}_y\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_x(\boldsymbol{\alpha})\\ \boldsymbol{A}_y(\boldsymbol{\alpha})\end{bmatrix}\mathrm{d}t+\begin{bmatrix}\underset{\sim}{\boldsymbol{C}}_x\mathrm{d}\boldsymbol{W}(t)\\ \underset{\sim}{\boldsymbol{C}}_y\mathrm{d}\boldsymbol{W}(t)\end{bmatrix} \tag{7.7.64}$$

其中$\boldsymbol{W}(t)$是维数与$\boldsymbol{\alpha}_x$相同的维纳过程。请注意，由于如(7.7.63)中所写那样在$\underset{\sim}{\mathscr{C}}(\boldsymbol{\alpha})$中有两个元素是零而使二行中出现了相同的维纳过程。

在把实部和虚部组合起来后，我们就看到了对复变量$\boldsymbol{\alpha}$的SDE，

$$\mathrm{d}\boldsymbol{\alpha}=\boldsymbol{A}(\boldsymbol{\alpha})\mathrm{d}t+\underset{\sim}{\boldsymbol{C}}(\boldsymbol{\alpha})\mathrm{d}\boldsymbol{W}(t) \tag{7.7.65}$$

假如我们把转变福克-普朗克方程为随机微分方程的常用法则直接用于泊松表示下的福克-普朗克方程 (7.7.9)，并且置下列事实而不顾：若$\underset{\sim}{\boldsymbol{B}}$不是一个半正定的扩散矩阵，则这样定义的$\underset{\sim}{\boldsymbol{C}}(\boldsymbol{\alpha})$就要有复元素，那末它所产生的SDE自然与上述方程完全相同。

c) 复平面中的随机微分方程的例子：

我们再次考虑反应(第7.7.6节)

$$\boldsymbol{A}+\boldsymbol{X}\underset{k_4}{\overset{k_3}{\rightleftharpoons}}2\boldsymbol{X}$$

$$B + X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} C \tag{7.7.66}$$

对此系统应用正泊松表示，从福克-普朗克方程(7.7.16)中就可以产生出 SDE：

$$\begin{aligned} \mathrm{d}\alpha &= [k_3 C + (k_2 A - k_1 B)\alpha - k_4 \alpha^2]\mathrm{d}t \\ &\quad + [2(k_2 A\alpha - k_4\alpha^2)]^{\frac{1}{2}}\mathrm{d}W(t) \end{aligned} \tag{7.7.67}$$

在 $\delta>0$ 的情况下，我们注意噪声项在 $\alpha=0$ 及 $\alpha=k_2A/k_4$ 为零而在这两点之间则为正，并且漂移项则是使得它只要逼近 $[0, k_2A/k_4]$ 的端点就要返回这个区间中来，于是对 $\delta>0$，(7.7.67)表示了在实区间 $[0, k_2A/k_4]$ 上的一个实 SDE。

在 $\delta<0$ 的情况下，平稳点在区间 $[0, k_2A/k_4]$ 的外面，而且若一个点起初在此区间内，它将受(7.7.67)的支配而沿此区间移动，直到它碰到区间的右端点，在这里噪声为零而漂移项则继续驱使它向右方前进，一旦离开了此区间以后，噪声变成了虚的，并且这个点沿着一条像图 7.4 中所示的那种路径继续下去，直到它终于再达到区间 $[0, k_2A/k_4]$。

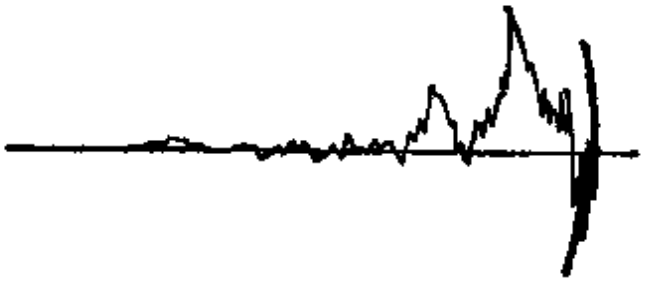


图 7.4　服从随机微分方程(7.7.67)的一点所走的路径

$\delta=0$ 的情况并无太大不同，但是一旦点到达区间 $[0, k_2A/k_4]$ 的右端点，则漂移与扩散都为零而使得它此后就留在那里。

在系统为

$$\begin{aligned} &B \to X \\ &2X \to A \end{aligned} \tag{7.7.68}$$

的情况下



图 7.5　对服从随机微分方程(7.7.69)的一个点之路径所作的模拟

来自于福克-普朗克方程(7.7.41)的 SDE 是

$$\mathrm{d}\eta/\mathrm{d}t = K_1 - 2K_2\eta^2 + i\varepsilon(2K_2)^{\frac{1}{2}}\eta\xi(t) \qquad (7.7.69)$$

其中 $\alpha = \eta V$ 且 $\varepsilon = V^{-\frac{1}{2}}$.

SDE(7.7.69)可进行模拟计算，并在复 η 平面上产生出一个运动的图形. 图 7.5 说明了其行为，可看到这个点停滞在

$$\mathrm{Re}\{\alpha\} = \left(\frac{a}{2}\right)^{\frac{1}{2}}$$

的附近，但是主要沿虚轴方向向两侧发生涨落，这样就产生出 α 的负方差.

7.7.5　时间相关函数

泊松分布的变量 α 的时间相关函数并不与变量 x 的相同. 譬如在反应 $X \rightleftarrows Y$ 的情况下就可看到这点，这个反应给出了没有扩散项的泊松表示下的福克-普朗克方程. 因此，泊松分布变量是没有涨落的. 我们现在来说明它们的关系是怎样的. 为简明起见证明是仅对单变量情况来进行的.

定义

$$\langle\alpha(t)\alpha(s)\rangle = \int \mathrm{d}\mu(\alpha)\mathrm{d}\mu(\alpha')\alpha\alpha' f(\alpha,\ t|\alpha',\ s) f(\alpha',\ s) \qquad (7.7.70)$$

注意到

$$f(\alpha,\ s\,|\alpha',\ s)=\delta_\mu(\alpha-\alpha')$$

它意味着

$$\int \mathrm{d}\mu(\alpha)\mathrm{e}^{-\alpha}(\alpha^x/x!)\,f(\alpha,\ s\,|\alpha',\ s)=\mathrm{e}^{-\alpha'}\alpha'^x/x! \qquad (7.7.71)$$

从而

$$\int \mathrm{d}\mu(\alpha)\alpha f(\alpha,\ t\,|\alpha',\ s)=\sum_{x,x'} xP(x,\ t\,|x',\ s)\mathrm{e}^{-\alpha'}\alpha'^{x'}/x'!$$

因此

$$\begin{aligned}\langle\alpha(t)\alpha(s)\rangle&=\sum_{x,x'} xP(x,\ t\,|x',\ s)\\&\quad\times\int \mathrm{d}\mu(\alpha')\,(\alpha'^{x'+1}\mathrm{e}^{-\alpha'}/x'!)\,f(\alpha',\ s)\\&=\sum_{x,x} xP(x,\ t\,|x',\ s)\int \mathrm{d}\mu(\alpha')\\&\quad\times\left[\left(-\alpha'\frac{\partial}{\partial\alpha'}+x'\right)(\alpha'^{x'}\mathrm{e}^{-\alpha'}/x'!)\right]f(\alpha',\ s)\\&=\sum_{x,x'} xx'P(x,\ t\,|x',\ s)P(x',\ s) \qquad (7.7.72)\\&\quad-\int \mathrm{d}\mu(\alpha')\,f(\alpha',\ s)\alpha'\frac{\partial}{\partial\alpha'}\\&\quad\times\sum_{x,x'} xP(x,\ t\,|x',\ s)(\alpha'^{x'}\mathrm{e}^{-\alpha'}x!) \qquad (7.7.73)\end{aligned}$$

我们定义

$$\langle\alpha(t)\,|[\alpha',\ s]\rangle\equiv\int \mathrm{d}\alpha\alpha f(\alpha,\ t|\alpha',\ s) \qquad (7.7.74)$$

为已给初始在 S 取 α' 值的条件下 $\alpha(t)$ 的均值，则第二项可写成

$$\begin{aligned}&-\int \mathrm{d}\mu(\alpha')\alpha'\frac{\partial}{\partial\alpha'}\langle\alpha(t)\,|[\alpha',\ s]\rangle f(\alpha',\ s)\\&\equiv\left\langle\alpha'\frac{\partial}{\partial\alpha'}\langle\alpha(t)\,|[\alpha,\ s]\rangle\right\rangle \qquad (7.7.75)\end{aligned}$$

而使得

$$\langle x(t)x(s)\rangle=\langle\alpha(t)\alpha(s)\rangle+\left\langle\alpha'\frac{\partial}{\partial\alpha'}\langle\alpha(t)\,|[\alpha,\ s]\rangle\right\rangle \qquad (7.7.76)$$

考虑到多变量的情况并注意到总是成立着

$$\langle x(t)\rangle = \langle \alpha(t)\rangle$$

$$\boxed{\begin{aligned}\langle x_a(t),\ x_b(s)\rangle &= \langle \alpha_a(t),\ \alpha_b(s)\rangle \\ &+\left\langle \alpha'_b \frac{\partial}{\partial \alpha'_b}\langle \alpha_a(t)\ [\alpha',\ s]\rangle\right\rangle\end{aligned}} \tag{7.7.77}$$

这个公式明显地说明了一个事实：泊松表示给出了一个与生灭主方程密切有关但并不同构于它的过程．我们感兴趣的随机量，如时间相关函数，都可以计算出来，但不是直接靠此泊松变量的有关量而给出的。

a) 用统计力学的术语给出的解释

我们暂且假定读者熟悉化学系统的统计力学．如果我们考虑一个由化学反应成分 A, B, C, …组成的系统，这个正则系综总体的分布函数是由

$$P(I) = \exp\{\beta[\Omega + \sum_i \mu_i x_i(I) - E(I)]\} \tag{7.7.78}$$

来给出的，其中 I 是一个刻划系统微观态的指标，$x_i(I)$ 是处于状态 I 中的分子 X_i 的数目，$E(I)$ 是这个状态的能量，μ_i 是成分 X_i 的化学势，Ω 是归一化因子，并且

$$\beta = 1/kT \tag{7.7.79}$$

一些成份要能进行反应是要求它们的化学势之间的某种关系得以满足，因为一个状态 I 只有当

$$\sum_i \nu_i^A x_i(I) = \sum_i \nu_i^A x_i(J) \qquad A = 1,\ 2,\ 3,\ \cdots \tag{7.7.80}$$

才能够转变到一个状态 J 中去，其中 ν_a^A 是某些整数，关系式(7.7.80)是些化学计算的约束。

一个反应系统的正则系综是由要求

$$\sum_i \nu_i^A x_i(I) = \tau^A \tag{7.7.81}$$

对某个 τ^A 成立而定义的．而正则系综的总体则由要求

$$\sum_I P(I) \sum_i \nu_i^A x_i(I) \equiv \sum_i \nu_i^A \langle x_i\rangle = \tau^A \tag{7.7.82}$$

来定义的。

使在约束(7.7.82)下（以及在给定总概率和平均能量的通常约束下)的熵极大，就给出了正则系综总体的形式(7.7.78)，此时，化学势之间也必满足关系式

$$\mu_i = \sum_A k_A \nu_i^A \tag{7.7.83}$$

当我们考虑一个理想化的解或者理想化的气体极限情况时，其反应能(而不是动能或内能)是被忽略不计的，因而在一个反应系统的分布函数和一个理想化的非反应系统的分布函数之间，若撇开化学势能表示为(7.7.83)的形式的要求不管，就没有什么差别了。

正则系综却不是这样简单的，因为约束必然作为

$$\prod_A \delta[\sum_i \nu_i^A x_i(I),\ \tau^A] \tag{7.7.84}$$

形式的一个因子而明显出现，并且每一类反应系统(包括非反应系统作为特况在内)的分布函数在数值上也是不同的。

在一个理想的反应系统的正则系综总体中，反应成份的分子总数 x 的分布是容易计算出来的，它就是

$$P(x) = \exp\ [\beta(\Omega + \sum_i \mu_i x_i)]$$
$$\times \sum_I \prod_i \delta[x_i(I),\ x_i] \exp[-\beta E_k(I)] \tag{7.7.85}$$

在各个状态上的求和是相同于对一个理想化非反应混合体的正则系综的求和，所以

$$P(x) = \exp[\beta(\Omega + \sum_i \mu_i x_i)] \prod_i \frac{1}{x_i!} \{\sum_k \exp[-\beta E_k(i)]\}^{x_i} \tag{7.7.86}$$

其中 $E_k(i)$ 是此物质 A 的分子的能量本征态。其结果是一个多变量泊松分布，其均值用

$$\log\langle x_i\rangle = \beta\mu_i - \log[\sum_k e^{-\beta E_k(i)}] \tag{7.7.87}$$

给出，如所周知，这和(7.7.82)的要求结合起来就给出了质量作用定律。

正则系综是在可推得较弱约束(7.7.82)的较强约束(7.7.81)

下使熵极大而得到。于是正则系综的总分子数的分布函数就由

$$P(x)\propto\left\{\prod_i \frac{1}{x_i!}\left[\sum_k \omega^{(i)}(k)\right]^{x_i}\right\}\sum_A \delta\left[\sum_i \nu_i^A x_i,\ \tau^A\right] \tag{7.7.88}$$

来给定。

我们刚才已经靠泊松分布表示说明了在平衡态下（一个正则系综总体中）拟概率是

$$f(\alpha)_{eq}=\delta[\alpha-\alpha(eq)] \tag{7.7.89}$$

因为 x 空间的分布是泊松分布，对于时间相关函数，有两个有关结果。

i) 变量 $\alpha(t)$ 及 $\alpha(s)$ 是非涨落量，其值为 $\alpha(eq)$。于是

$$\langle\alpha_r(t),\ \alpha_s(s)\rangle_{eq}=0。 \tag{7.7.90}$$

ii) 第二项的平衡均值是平凡的，于是

$$\langle x_a(t),\ x_b(s)\rangle=\left[\alpha_b'\frac{\partial}{\partial\alpha_b'}\langle\alpha_a(t)\,|\,[\alpha',\ s]\rangle\right]_{\alpha'=\alpha(eq)} \tag{7.7.91}$$

这个结果实际上恰好与伯纳德(Bernard, W.)和卡伦(Callen, H.B.)[7.11]的相同，此文把两时刻相关函数跟某个量关于一个热力学共轭变量的平均的导数建立了关系。

我们来考虑一个系统，其中对应于系统的一个构形 I 的各化学物质 X_1、X_2、… 之分子数是 $x_1(I)$、$x_2(I)$、…，并且我们已了解这些化学物质彼此可能进行反应，那么如上面已说明过的，在一个正则系综总体中平衡分布函数是

$$Z^{-1}(\mu)\exp\left[\{\sum_i \mu_i x_i(I)-E(I)\}/kT\right] \tag{7.7.92}$$

同时

$$Z(\mu)=\exp(-\Omega\beta) \tag{7.7.93}$$

其中 $Z(\mu)$ 是总体的正则配分函数。正如上面已指出过的，一个反应系统的化学势 μ_i 不能任意地选取，而必须由容许反应的化学计算约束(7.7.82)形成相关关系。

现在进一步来定义量

$$\langle x_i,\ t|[I,\ s]\rangle \tag{7.7.94}$$

为当系统于S时刻处于构形I的条件下x_i在t时刻的平均值，那么我们感兴趣的是(7.7.94)对初始条件的分布(7.7.92)的平均，即

$$\langle x_i,\ t|[\mu,\ s]\rangle = \sum_J \langle x_i,\ t|[J,\ S]\rangle Z^{-1}(\mu) \times \exp\left\{\frac{1}{kJ}\left[\sum_j \mu_j x_j(J) - E(J)\right]\right\} \tag{7.7.95}$$

当这些化学势满足平衡约束时，这个量将是与时间无关的而且等于平衡下的x_i的均值，但是在其他情况下它将是与时间相依的。现在，稍作一些处理我们就会看到

$$\left[kT\frac{\partial}{\partial\mu_j}\langle x_i,\ t|[\mu,\ s]\rangle\right]_{\mu=\mu(eq)} = \langle x_i(t),\ x_j(s)\rangle_{eq} \tag{7.7.96}$$

左边是均值对化学势在平衡态附近的改变作出反应的响应函数，因而是个散逸的度量，而右边是在平衡态下的两时刻相关函数，是个涨落的度量。

为和泊松表示的结果(7.7.91)联系起来，我们注意在理想溶解理论中的化学势μ_j是由

$$\mu_i(\langle x\rangle) = kT\log\langle x_i\rangle + 常数 \tag{7.7.97}$$

由(7.7.97)，我们看到(7.7.96)成了

$$\langle x_i(t),\ x_j(s)\rangle = \left[\langle x_j\rangle\frac{\partial}{\partial\langle x_j\rangle}\langle x_i,\ t|[\mu(\langle x\rangle),\ s]\rangle\right]_{\langle x\rangle=\langle x\rangle_{eq}} \tag{7.7.98}$$

因为理想溶解理论产生出x_i属泊松分布变量，取极限就得出

$$\langle x_i,\ t|[\mu(\langle x\rangle),\ s]\rangle = \langle\alpha_i,\ t|[\alpha',\ s]\rangle \tag{7.7.99}$$

其中$\alpha' = \langle x\rangle$。于是(7.7.98)成了

$$\langle x_i(t),\ x_j(s)\rangle = \left[\alpha'_j\frac{\partial}{\partial\alpha_j}\langle\alpha_i,\ t|[\alpha',\ s]\rangle\right]_{\alpha'=\alpha(eq)} \tag{7.7.100}$$

这样，(7.7.91)就是一般性结果(7.7.98)的理想溶解极限。

一般公式 (7.7.77) 可以看作是伯纳德-卡伦的结果到非热力学平衡系统上的推广。

但是它相当不同于平衡下的结果，而且对两项立刻可作出解释。第二项是平衡的作用，是一个响应函数，但由于系统不是在一个良定的平衡状态中，我们就对各种会出现的 α 空间状态将平衡结果取平均。第一项是来自于 α 空间涨落本身的效果，并不直接关系到一个响应函数，它表示超过平衡的涨落。

b) 线性化结果

从利用正泊松表示中产生的并对应于福克-普朗克方程 (7.7.12) 的一般微分方程是

$$\mathrm{d}\eta = \hat{A}(\eta)\mathrm{d}t + \varepsilon\underset{\sim}{\hat{C}}(\eta)\mathrm{d}\boldsymbol{W}(t) \tag{7.7.101}$$

其中

$$\underset{\sim}{\hat{C}}\underset{\sim}{\hat{C}}^{T} = \underset{\sim}{\hat{B}} \tag{7.7.102}$$

我们现在可以用第 6.3 节用过的下列程序在平稳态 η 附近作一个一阶小噪声展开。记

$$\eta(t) = \bar{\eta} + \varepsilon\eta(t) \qquad (\varepsilon = V^{-\frac{1}{2}}) \tag{7.7.103}$$

在最低阶上我们看到

$$\hat{A}(\bar{\eta}) = 0 \tag{7.7.104}$$

$$\mathrm{d}\eta_1 = -\boldsymbol{F}\eta_1\mathrm{d}t + \underset{\sim}{\boldsymbol{G}}\mathrm{d}\boldsymbol{W}(\mathrm{t})$$

其中

$$\boldsymbol{F}_{rs} = -\frac{\partial}{\partial\eta_s}\hat{A}_r(\eta) \tag{7.7.105}$$

$$\underset{\sim}{\boldsymbol{G}} = \underset{\sim}{\hat{C}}(\eta)$$

那末，

$$\langle\alpha_r(t),\ \alpha_s(0)\rangle_s = V\sum_{r'}[\exp(-\boldsymbol{F}t)]_{rr'}\langle\eta_{r',1},\ \eta_{s,1}\rangle_s \tag{7.7.106}$$

并且

$$\frac{\partial}{\partial\alpha'_s}\langle\alpha_r(t)\,|\,[\boldsymbol{\alpha}',\ 0]\rangle = -\frac{\partial}{\partial\eta'_{s,1}}\langle\eta_{r,1}(t)\,|\,[\boldsymbol{\eta}'_1,\ 0]\rangle$$
$$= [\exp(-\boldsymbol{F}t)]_{r,s} \tag{7.7.107}$$

因此

$$\langle x_r(t),\ x_s(0)\rangle = V\sum_{r'}[\exp(-Ft)]_{rr'}$$

$$\times[\langle \eta_{r'+1},\ \eta_{s+1}\rangle_s + \delta_{r's}\bar{\eta}_s) \qquad (7.7.108)$$

$$= \sum_{r'}\exp(-Ft)_{r,r'}\langle x_{r'}x_s\rangle_s \qquad (7.7.109)$$

于是，此线性化结果与第 3.7.4 节的回归定理是一致的。在[7.10]中已对各种系统计算了相关函数。

7.7.6 三分子反应

在第 7.1.3 节我们考察了一个包含着

$$A+2X \rightleftarrows 3X \qquad (7.7.110)$$

部分的反应，并对它建立了一个适当的生灭主方程，但是人们都知道，在化学中这种三分子阶段是以近乎零的小概率存在的，而且是通过一个短寿命的中间体来进行这种反应的。因此料想反应(7.7.110)是作为一个双态系统而出现的。

$$\text{i)}\ A+Y \overset{\perp}{\rightleftarrows} X+Y \qquad (7.7.111a)$$

$$\text{ii)}\ Y \overset{\gamma}{\rightleftarrows} 2X \qquad (7.7.111b)$$

这两者都仅仅是双分子的，并且除了假定 γ(Y 的衰减常数)是非常大外，我们还令速率常数等于 1，这样，Y 确实是个短寿命的中间体。确定性的速率方程是

$$\frac{dx}{dt} = ay - xy + 2(\gamma y - x^2)$$

$$\frac{dy}{dt} = x^2 - \gamma y \qquad (7.7.112)$$

并且通常的确定性绝热消去程序置 $y=x^2/\gamma$，并给出

$$\frac{dx}{dt} = (ax^2 - x^3)/\gamma \qquad (7.7.113)$$

尽管这个程序是断然明确的，但并不清楚 7.1.3 中所用的随机主方程是不是一个有效的绝热消去极限。在第 6 章所用的绝热消去技巧并不能轻易地直接用于主方程，但它能够直接适用于泊松表

示下的福克-普朗克方程。

a) 三分子反应的福克-普朗克方程

对于与 (7.7.113) 相对应的前向及后向反应速率等于 $1/\gamma$ 的反应 (7.7.110)，其泊松表示福克-普朗克方程由(7.7.4)变成了

$$\frac{\partial f}{\partial t}=\frac{1}{\gamma}\left\{\left(-\frac{\partial}{\partial\alpha}+2\frac{\partial^2}{\partial\alpha^2}-\frac{\partial^3}{\partial\alpha^3}\right)[\alpha^2(a-\alpha)]f\right\} \tag{7.7.114}$$

而含有三阶导数。不管选用哪种泊松表示，还找不到用某种 α 空间中的实随机过程来表达的合适的概率论解释。在下一节中三阶噪声的概念将得到解释，以说明概率论方法和随机微分方程如何仍然有用。

b) 绝热消去法

使用(7.4.9)中所发展的法则，则系统(7.7.111)的福克-普朗克方程在作对应

$$\begin{bmatrix}x\\y\end{bmatrix}\longleftrightarrow\begin{bmatrix}\alpha\\\beta\end{bmatrix}$$

后是

$$\frac{\partial f}{\partial\alpha}=-\frac{\partial}{\partial\alpha}[(a-\alpha)\beta+2(\gamma\beta-\alpha^2)]+\frac{\partial}{\partial\beta}(\gamma\beta-\alpha^2)$$
$$+\frac{\partial^2}{\partial\alpha^2}(\gamma\beta-\alpha^2)+\frac{\partial^2}{\partial\alpha\partial\beta}[(a-\alpha)\beta] \tag{7.7.115}$$

现在就如同第 6.6.1 节那样地进行绝热消去。我们定义新变量

$$\begin{aligned}x&=\alpha\\y&=\gamma\beta-\alpha^2\end{aligned} \tag{7.7.116}$$

并因而用

$$\begin{aligned}\frac{\partial}{\partial\alpha}&=\frac{\partial}{\partial x}-2x\frac{\partial}{\partial y}\\\frac{\partial}{\partial\beta}&=\gamma\frac{\partial}{\partial y}\end{aligned} \tag{7.7.117}$$

作变量变换后，福克-普朗克方程就成为

$$\frac{\partial f(x,y)}{\partial t}=\left\{-\left(\frac{\partial}{\partial x}-2x\frac{\partial}{\partial y}\right)\left[\frac{(a-x)(y+x^2)}{\gamma}+2y\right]\right.$$

$$+\gamma\frac{\partial}{\partial y}y+\left(\frac{\partial}{\partial x}-2x\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x}-2x\frac{\partial}{\partial y}\right)y$$
$$+\left(\frac{\partial}{\partial x}-2x\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial y}\right)[(y+x^2)(a-x)]\Big\}f \tag{7.7.118}$$

因为 y 是要消去的，所以在给定的 x 控制着 y 运动的算子 L_1 应该有良定的极限，但是此算子是

$$\gamma\frac{\partial}{\partial y}y+\frac{\partial^2}{\partial y^2}[4x^2y-2x(y+x^2)(a-x)] \tag{7.7.119}$$

而且 γ 趋于无穷大的极限把它变成了确定性运动。置

$$y=v\gamma^{-\frac{1}{2}} \tag{7.7.120}$$

它把(7.7.119)转换为

$$\gamma L_1(\gamma)=\gamma\left\{\frac{\partial}{\partial v}v+\frac{\partial^2}{\partial v^2}[2x^3(x-a)+(4x^2-2x)v\gamma^{-\frac{1}{2}}]\right\}$$
$$\xrightarrow[\gamma\to\infty]{}\gamma\left\{\frac{\partial}{\partial v}v+\frac{\partial^2}{\partial v^2}[2x^3(x-a)]\right\}\equiv\gamma L_1 \tag{7.7.121}$$

用了这个代换后，最后我们可建立等式

$$\gamma^{-1}L_3=-\gamma^{-1}\frac{\partial}{\partial x}[x^2(a-x)] \tag{7.7.122}$$

$$L_2(\gamma)=-\frac{\partial}{\partial x}[(a-x)v\gamma^{-\frac{3}{2}}+2v\gamma^{-\frac{1}{2}}]$$
$$-\frac{\partial}{\partial x}2x\frac{\partial}{\partial v}v-2x\frac{\partial}{\partial v}\frac{\partial}{\partial x}v$$
$$+\gamma^{-\frac{1}{2}}\frac{\partial^2}{\partial x^2}v+\gamma^{\frac{1}{2}}\frac{\partial}{\partial x}\frac{\partial}{\partial v}[(a-x)(x^2+v\gamma^{-\frac{1}{2}})] \tag{7.7.123}$$

以及

$$\frac{\partial f}{\partial t}=[\gamma^{-1}L_3+L_2(\gamma)+\gamma L_1(\gamma)]f \tag{7.7.124}$$

投影算子 P 将是映射在 L_1 的零空间上的，并且因为 L_1 依赖于 x，我们有

$$L_3P\neq PL_3 \tag{7.7.125}$$

这意味着我们可借类似于第6.5.4节中使用过的计算方法来找到 $Pf=g$ 的运动方程。结果是

$$s\tilde{g}(s)=\gamma^{-1}PL_3\tilde{g}(s)+P[L_2(\gamma)+\gamma^{-1}L_3][s-\gamma L_1 -(1-P)L_2(\gamma)-\gamma^{-1}(1-P)L_3]^{-1} \times[L_2(\gamma)+\gamma^{-1}(1-P)L_3]\tilde{g}(s)+g(0) \quad (7.7.126)$$

但是注意，由于对 v 的任何函数

$$P\phi(v)=P_x(v)\int dv\phi(v) \quad (7.7.127)$$

其中 $p_x(v)$ 满足

$$L_1p_x(v)=0 \quad (7.7.128)$$

这使得在 $PL_2(\gamma)$ 中凡具 $\frac{\partial}{\partial v}$ 的项都为0。这样一来，直到 γ 的最高阶，都有

$$PL_2(\gamma)=\gamma^{-\frac{1}{2}}\left(-2v\frac{\partial}{\partial x}+v\frac{\partial^2}{\partial x^2}\right) \quad (7.7.129)$$

在(7.7.126)中[ ]$^{-1}$一项是渐近收敛于 $-\gamma^{-1}L_1^{-1}$ 的，并且在另一括号中的像 L_3 的那项，构成了整个 γ^{-1} 阶的表示式，是 $L_2(\gamma)$ 中 $\gamma^{\frac{1}{2}}$ 阶的项，即

$$\gamma^{\frac{1}{2}}\frac{\partial}{\partial x}[(a-x)x^2]\frac{\partial}{\partial v} \quad (7.7.130)$$

于是(7.7.126)的 γ 大值极限是

$$s\tilde{g}(s)=\gamma^{-1}\left\{PL_3\tilde{g}-P\left[-2\frac{\partial}{\partial x}+\frac{\partial^2}{\partial x^2}\right]vL_1^{-1} \times\frac{\partial}{\partial x}\left[(a-x)x^2\frac{\partial}{\partial v}\right]p_x(v)\tilde{p}\right\}+g(0) \quad (7.7.131)$$

其中我们已经用了记号

$$g=p_x(v)p$$
$$\tilde{g}=p_x(v)\tilde{p} \quad (7.7.132)$$

现在我们已被引导到计算

$$\int dv'v'L_1^{-1}\frac{\partial}{\partial x}(a-x)x^2\frac{\partial}{\partial v'}p_x(v') \quad (7.7.133)$$

这个核心问题上来了，它产生于计算(7.7.131)中括号内的第二部

分，我们希望把 $\frac{\partial}{\partial v}$ 放到积分号外面来，但是由于 $\frac{\partial}{\partial x}$ 与 L_1 不是可交换的，这就得小心些，现在

$$\left[L_1^{-1},\ \frac{\partial}{\partial x}\right] = L_1^{-1}\left[\frac{\partial}{\partial x},\ L_1\right]L_1^{-1} \tag{7.7.134}$$

并且由 [7.7.121)，

$$= L_1^{-1}\left\{\frac{\partial^2}{\partial v^2}\ [8x^3 - 6ax^2]\right\}L_1^{-1} \tag{7.7.135}$$

$$\begin{aligned}(7.7.133) = \frac{\partial}{\partial x}\int \mathrm{d}v' v' L_1^{-1}(a-x)x^2\ \frac{\partial}{\partial v'}\ p_s(v') \\ + \int \mathrm{d}v' v' L_1^{-1}\ \frac{\partial^2}{\partial v'^2}\ L_1^{-1} \\ \times [(8x^3 - 6ax^2)(a-x)x^3]\ \frac{\partial}{\partial v'}\ p_s(v')\end{aligned} \tag{7.7.136}$$

第二项是零，虽然其证明较特别。我们知道 L_1 刻划了关于 v 的奥恩斯坦-乌伦贝克过程，$p_s(v)$ 是它的平稳解。第 6.4.2 节用过的本征函数性质表明

$$L_1^{-1}\ \frac{\partial^2}{\partial v^2}\ L_1^{-1}\ \frac{\partial}{\partial v}\ p_s(v) \tag{7.7.137}$$

是与其第三个本征函数成比例的，而它与相应的后向方程之第一特征函数 v 成正交关系。利用 L_1 包含了奥恩斯坦-乌伦贝克过程在内的事实，第一项是容易计算的。再用第 6.6.1 节的相同技巧，我们会发现由 $p_s(v')$ 而产生的对 x 的所有相依性等于零，并且因而

$$(7.7.133) = \frac{\partial}{\partial x}\ (a-x)x^2 \tag{7.7.138}$$

类似地，我们看到

$$\begin{aligned}PL_3\vec{g} &= -p_s(v)\int \mathrm{d}v'\ \frac{\partial}{\partial x}\ [x^2(a-x)]p_s(v')\vec{p} \\ &= -p_s(v)\ \frac{\partial}{\partial x}\ [x^2(a-x)]\vec{p}\end{aligned} \tag{7.7.139}$$

以至最后

$$\frac{\partial p}{\partial t} = \frac{1}{\gamma}\left[\left(-\frac{\partial}{\partial x} + 2\frac{\partial^2}{\partial x^2} - \frac{\partial^3}{\partial x^3}\right)[(a-x)x^2]\,p\right] \quad (7.7.140)$$

它与三分子模型的福克-普朗克方程(7.7.114)完全相同。
这意味着在相同的极限下三分子主方程是有效的。

c) 注解

i) 请注意此系统给出了一个极端结果。它不是以斯特拉托诺维奇形式而是以伊藤形式表现的，具有所有左导数。

ii) (7.7.140) 的导出意味着要求有某些技巧以便理解这种非概率性的福克-普朗克方程。在下一节我们将概述做到这点的一个可能方法。

7.7.7 三阶噪声

为处理产生于三分子反应的三阶福克-普朗克方程，我们引入随机变量 $V(t)$，其条件概率密度 $p(v, t)$ 服从三阶偏微分方程。

$$\frac{\partial p(v, t)}{\partial t} = -\frac{1}{6}\partial^3 p(v, t)/\partial v^3 \quad (7.7.141)$$

由于在第 3.4 节中已经说明过没有哪一个马尔可夫过程能够给出像这样的三阶项，故 $p(v, t)$ 必然是违反了某些基本的要求。结果是 $p(v, t)$ 并不总为正，在拟概率意义下这是允许的。尽管这样，我们将会看到它与正式概率之间的类似是十分有用的。

当赋于边界条件

$$p(v, t_0) = \delta(v - v_0) \quad (7.7.142)$$

时我们知道(7.7.141)的解是由傅里叶变换方法给出为

$$p(v, t|v_0, t_0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dq \exp\left\{i\left[q(v-v_0) + \frac{1}{6}q^3(t-t_0)\right]\right\} \quad (7.7.143)$$

V 的矩可以计算出来，作了偏积分之后得到

$$\langle[V(t) - V_0]^n\rangle = 0 \quad n \text{ 不是 } 3 \text{ 的倍数}$$
$$\langle[V(t) - V_0]^{3m}\rangle = [(t-t_0)/6]^m (3m)!/m! \quad (7.7.144)$$

进一步，我们假设过程(7.7.141)是某种广义马尔可夫过程，其联合概率分布是由

$$p(v_2t_2;v_1t_1)=p(v_2t_2|v_1t_1)p(v_1,t_1) \tag{7.7.145}$$

给出的。由 (7.7.142) 我们看到右边第一个因子仅是 v_2-v_1 与 t_2-t_1 的函数，所以变量 $V(t_2)-V(t_1)$ 是与 $V(t_1)$ 统计独立的，并且这是个具有独立增量的过程。这样，$\mathrm{d}V(t)$ 将是独立于 $V(t)$ 的。严格地定义关于 $V(t)$ 的随机积分是个我们并不想现在就完成的任务。但是，显然它并不是过分不同于伊藤积分的，事实上霍克伯格 (Hockberg, K.) [7.12] 已经严格地定义了偶数次的高阶噪声，并完成了关于它们的随机积分，不过我们可以证明一个形式为

$$\mathrm{d}y(t)=a(y)\mathrm{d}t+b(y)\mathrm{d}W(t)+c(y)\mathrm{d}V(t) \tag{7.7.146}$$

($W(t)$ 及 $V(t)$是独立的过程)的随机微分方程等价于一个三阶福克-普朗克方程。因为 $W(t)$及 $V(t)$是具有独立增量的过程，显然 $y(t)$ 是马尔可夫过程，那末我们可以计算

$$\lim_{t\to 0}\frac{\langle[y(t)-y(t_0)]^n\rangle}{t-t_0}=\lim_{\mathrm{d}t_0\to 0}\frac{\langle[\mathrm{d}y(t_0)]^n\rangle}{\mathrm{d}t_0} \tag{7.7.147}$$

其中 $y(t_0)$是一初始数值而不是一随机变量。由 (7.7.146)，$y(t)$ 仅仅依赖于 $W(t')$ 及 $V(t')$，$t'\leqslant t$。而且因为 $\mathrm{d}W(t)$ 与 $\mathrm{d}V(t)$ 是与$y(t)$ 独立的，我们发现，由于 $y(t_0)$是一初始数值，

$$\begin{aligned}\langle\mathrm{d}y(t_0)\rangle&=\langle a[y(t_0)]\rangle\mathrm{d}t_0+\langle b[y(t_0)]\rangle\langle\mathrm{d}W(t_0)\rangle\\&\quad+\langle c[y(t_0)]\rangle\langle\mathrm{d}V(t_0)\rangle\\&=\langle a[y(t_0)]\rangle\mathrm{d}t_0=a[y(t_0)]\mathrm{d}t_0\end{aligned} \tag{7.7.148}$$

类似地，到 $\mathrm{d}t_0$ 的最低阶，

$$\langle\mathrm{d}y(t_0)^2\rangle=b[y(t_0)]^2\langle\mathrm{d}W(t_0)^2\rangle=b[y(t_0)]^2\mathrm{d}t_0 \tag{7.7.149}$$

$$\langle\mathrm{d}y(t_0)^3\rangle=c[y(t_0)]^3\langle\mathrm{d}V(t_0)^3\rangle=c[y(t_0)]^3\mathrm{d}t_0 \tag{7.7.150}$$

于是我们看到

$$\left.\begin{aligned}&\lim_{t\to t_0}[\langle y(t)-y(t_0)\rangle/(t-t_0)]=a[y(t_0)]\\&\lim_{t\to t_0}[\langle[y(t)-y(t_0)]^2\rangle/(t-t_0)]=b(y(t_0)]^2\\&\lim_{t\to t_0}[\langle[y(t)-y(t_0)]^3\rangle/(t-t_0)]=c[y(t_0)]^3\end{aligned}\right\} \tag{7.7.151}$$

并且所有更高的幂次都给出结果为零，利用类似于第3.4节的分析，只要说明 $y(t)$ 是广义扩散过程就够了，其广义福克-普朗克方程是

$$\frac{\partial p(y,t)}{\partial t} = -\frac{\partial}{\partial y}[a(y)p] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[b(y)^2 p] - \frac{1}{6}\frac{\partial^3}{\partial y^3}[c(y)^3 p] \tag{7.7.152}$$

我们用

$$\mathrm{d}V(t) = \zeta(t)\mathrm{d}t \tag{7.7.153}$$

定义一个噪声源 $\zeta(t)$，其中

$$\langle\zeta(t)\rangle = \langle\zeta(t)\zeta(t')\rangle = 0 \tag{7.7.154}$$

$$\langle\zeta(t)\zeta(t')\zeta(t'')\rangle = \delta(t-t')\delta(t'-t'') \tag{7.7.155}$$

且高阶矩可以由 $\mathrm{d}V(t)$ 的矩很快地计算出来。增量的独立性意味着就像我们在与伊藤积分打交道时情况一样，当上限取为零时，积分具有 δ-函数的奇异性。

使用三阶噪声的例子。考虑化学过程

$$A + 2X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3X \tag{7.7.156}$$

$$A \underset{k_4}{\overset{k_3}{\rightleftharpoons}} X$$

其泊松表示的福克-普朗克方程是

$$\frac{\partial f(\alpha,t)}{\partial t} = -\frac{\partial}{\partial\alpha}[K_1V^{-1}\alpha^2 - K_2V^{-2}\alpha^3 + K_3V - K_4\alpha)f(\alpha,t)] + \frac{1}{2}\frac{\partial^2}{\partial\alpha^2}[4(K_1V^{-1}\alpha^2 - K_2V^{-2}\alpha^3)f(\alpha,t)] - \frac{1}{3!}\frac{\partial^3}{\partial\alpha^3}[6(K_1V^{-1}\alpha^2 - K_2V^{-2}\alpha^3)f(\alpha,t)] \tag{7.7.157}$$

其中 $K_1V^{-1} = k_1A$，$K_2V^{-2} = k_2$，$K_3V = k_3$，$K_4 = k_4$。

在稳态情况下，(7.7.157)归结到一个线性二阶微分方程，它

可以通过超几何函数解出，而且在用了快速下降法后还可得到各阶矩的渐近展开。这个程序虽然在原则上是可能的，但不是十分实际可行的，情况倒是随机微分方程的方法已证明非常有用，因为应用起来容易。

等价于(7.7.157)的随机微分方程是

$$\frac{d\eta(t)}{dt} = K_1\eta(t)^2 - K_2\eta(t)^3 + K_3 - K_4\eta(t) + \mu^3\{4[K_1\eta(t)^2 - K_2\eta(t)^3]\}^{\frac{1}{2}}\xi(t) + \mu^4\{6[K_1\eta(t)^2 - K_2\eta(t)^3]\}^{\frac{1}{8}}\zeta(t) \qquad (7.7.158)$$

其中 $\alpha = \eta V$，$\mu = V^{-\frac{1}{6}}$ 而且以后就称为"三阶噪声"的噪声源，是在(7.7.153-155)定义了的。

方程(7.7.158)可以借展开 $\eta(t)$ 而叠代地解出：

$$\eta(t) = \eta_0(t) + \mu^3\eta_3(t) + \mu^4\eta_4(t) + \mu^6\eta_6(t) + \mu^8\eta_8(t) + \mu^9\eta_9(t) + \cdots \qquad (7.7.159)$$

把它代入 (7.7.158)，则在最低阶产生确定性方程，并在高阶产生线性随机微分方程，它们可以如第 6.2 节那样被解出。

在平稳态时结果是

$$\langle x\rangle = V\eta_0 + \langle\eta_6\rangle + \cdots = V\eta_0 + \frac{2ab}{c^2} + \cdots \qquad (7.7.160a)$$

$$\langle x^2\rangle - \langle x\rangle^2 = V\langle\eta_3^2\rangle + [2\langle\eta_9\eta_3\rangle + 2\langle\eta_8\eta_4\rangle + \langle\eta_6^2\rangle - \langle\eta_6\rangle^2 + \langle\eta_6\rangle] + \cdots$$
$$= V\left[\frac{2a}{c}\right] + \left[\frac{28}{3}\frac{a^2b^2}{c^4} + \frac{8ab^2\eta_0}{c^3} - \frac{36K_2a^2}{c^3} + \frac{8ab}{c^2}\right] + \cdots \qquad (7.7.160b)$$

$$\langle(x - \langle x\rangle)^3\rangle = V\langle\eta_4^3\rangle - 3\langle\eta_3^2\rangle\langle\eta_6\rangle + 3\langle\eta_3^2\rangle + \langle\eta_0\rangle + \cdots$$
$$= V\left[\frac{8a}{c} - \frac{12a^2b}{c^3} + \eta_0\right] + \cdots \qquad (7.7.160c)$$

其中

$$a = K_1\eta_0^2 - K_2\eta_0^3$$
$$b = 2K_1 - 3K_2\eta_0$$

$$c = K_4 - 2K_1\eta_0 + 3K_2\eta_0^2$$

并且 η_0 是稳定态确定性方程

$$K_1\eta_0^3 - K_2\eta_0^2 + K_3 - K_4\eta_0 = 0 \qquad (7.7.161)$$

的解。这里要加几点评注。三阶噪声 $\zeta(t)$ 对均值提供了 $O(V^{-1})$，对方差则提供了 $O(1)$。但对偏斜度系数的作用是提供 $O(V)$。若我们只关心计算均值和方差到 $O(V)$，则此三阶噪声就可以从 (7.7.158) 中除去且展开可以按 $\varepsilon = V^{-\frac{1}{2}}$ 的幂次来完成。还要注意，当 $c \to 0$ 时，方差与高阶相关变成发散的，这当然是由于在该极限之下反应系统显示出一种一级相变型行为所致。

8. 空间分布系统

在本章，我们把反应扩散系统作为自然界中存在的许多空间分布系统的典型例子来处理。用空间相关的朗之万方程逐步而又启发式地引进这个主题，对其不足之处作了解释。然后，介绍了更令人满意的多元主方程的描述，根据系统大小把这种描述展开成级数，并把空间相关的朗之万方程表示为这种描述的一种近似。我们也证明了泊松表示法怎样能够给出非常类似的空间相关的朗之万方程而不需要任何近似。

接着，我们研究这些方程在空间和时间相关结构中的各结果，相关结构特别能够在不稳定点附近产生。然后证明局域描述和整体描述之间的关系，最后用相空间(即速度和位置空间)中的分布所描述的系统的处理方法来结束本章。这种处理是用玻耳兹曼主方程来完成的。

8.1 背景

空间概念是我们感知世界的一个重要概念，这主要是因为很好地被隔离的各个物体相互不会有很大的影响。在宏观决定论的水平上，这就导致可以用诸如局域密度、浓度、温度等等局域量来描述世界。根据决定论，这些量通常被认为服从偏微分方程，例如流体动力学中的内维尔-斯托克斯 (Navier-Stokes) 方程，化学中的反应扩散方程或经典电磁场中的麦克斯韦方程。

反应扩散方程是所讨论的最简单情况。该方程描述化学反应并构成本章的主要论题。为了获得有关这个概念的某些感性知识，我们先讨论化学物质浓度 ρ 随时间演化的朗之万方程的描

述。然后就可以在下面的推导中得出经典的反应扩散方程。扩散流 $j(r.t)$ 以如下形式存在

$$j(r,\ t) = -D\nabla\rho(r,\ t) \tag{8.1.1}$$

而(8.1.1)称为斐克(Fick)定律。若没有化学反应，则该流遵循一个守恒方程。例如，考虑一个任意体积 V，该体积内化学物质的总量只能因通过 V 的边界 S 的迁移而发生变化。因而，如果 N 是 V 内的总量，则

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int_V \mathrm{d}^3r\rho(r,\ t) = -\int_S \mathrm{d}s\cdot j(r,\ t) \tag{8.1.2}$$

$$= -\int_V \mathrm{d}^3r\nabla\cdot j(r,\ t)$$

由于 V 是任意的，所以

$$\partial_t\rho(r,\ t) + \nabla\cdot j(r,\ t) = 0 \tag{8.1.3}$$

将斐克定律(8.1.1)代入守恒方程 (8.1.3)，我们得微分方程

$$\partial_t\rho(r,\ t) = D\nabla^2\rho(r,\ t) \tag{8.1.4}$$

现在我们如何才能把涨落加上去呢？首先应注意到，守恒方程(8.1.3)是精确的；这可以从它的推导看到。所以我们不能把涨落项加到守恒方程上。但是，斐克定律可以通过增加一个斯托克斯源而很好地加以修改。这样，我们把斐克定律重新写成

$$j(r,\ t) = -D\nabla\rho(r,\ t) + f_d(r,\ t)。 \tag{8.1.5}$$

式中 $f_d(r,\ t)$ 是一个矢量朗之万源。关于它的随机性质的最简单假设为

$$\langle f_d(r,\ t)\rangle = 0$$

及

$$\langle f_{d,i}(r,\ t)\, f_{d,j}(r',\ t')\rangle = K_d(r,\ t)\delta_{ij}\delta(r-r')\delta(t-t') \tag{8.1.6}$$

这就是说，不同的分量在相同的时间和空间彼此是独立的，而不同时间和空间的所有涨落都是无关的。这就是局部性假设。因此涨落微分方程为

$$\partial_t\rho(r,\ t) = D\nabla^2\rho(r,\ t) - \nabla\cdot f_d(r,\ t) \tag{8.1.7}$$

应注意，

$$\langle \nabla \cdot f_d(r,\ t) \nabla' \cdot f_d(r',\ t') \rangle$$
$$= \nabla \cdot \nabla' [K_d(r,\ t) \delta(r - r')] \delta(t - t') \tag{8.1.8}$$

现在考虑包括化学反应的扩散方程。斐克定律仍然起作用，但是，我们需要的不是守恒方程，而是如下形式的方程：

$$\frac{dN}{dt} = \frac{d}{dt} \int_v d^3 r \rho(r,\ t) = -\int_s dS \cdot j(r,\ t)$$
$$+ \int_v d^3 r F[\rho(r,\ t)] \tag{8.1.9}$$

式中 $F[\rho(r,\ t)]$ 是浓度的函数并表示局部化学反应中的化学产物。

因此，考虑涨落前我们发现

$$\partial_t \rho(r,\ t) + \nabla \cdot j(r,\ t) = F[\rho(r,\ t)] \tag{8.1.10}$$

化学反应产生的化学产物确实产生涨落，所以我们可以在(8.1.10)中加入 $f_c(r,\ t)$ 项，它满足如下关系：

$$\langle f_c(r,\ t) \rangle = 0$$
$$\langle f_c(r,\ t) f_c(r',\ t') \rangle = K_c(r,\ t) \delta(r - r') \delta(t - t') \tag{8.1.11}$$

上式表示这样的事实，即反应是局部的(也就是说在不同点的涨落是非相关的)和马尔可夫的(即在时间上是 δ 相关的)。现在完整的反应扩散方程成为

$$\boxed{\partial_t \rho(r,\ t) = D \nabla^2 \rho(r,\ t) + F[\rho(r,\ t)] + g(r,\ t)} \tag{8.1.12}$$

式中

$$g(r,\ t) = -\nabla \cdot f_d(r,\ t) + f_c(r,\ t) \tag{8.1.13}$$

而

$$\boxed{\begin{aligned}\langle g(r,\ t) g(r',\ t') \rangle = \{ & K_c(r - r') \delta(r - r') \\ & + \nabla \cdot \nabla' [K_d(r,\ t) \delta(r - r')] \} \delta(t - t')\end{aligned}} \tag{8.1.14}$$

把经典反应扩散方程转变为朗之万方程的最简单过程产生相当复杂的表示式，再说我们对 $K_c(r)$ 或 $K_d(r)$ 一无所知，而且这个过

程是以非常富有启发性的模型为基础的。

尽管如此，但就上述方程与从建立在主方程基础上的更为微观的方法所得结果相一致这点来说，上面所导出的公式本质上是正确的，然而它精确地规定了所有任意常数。

8.1.1 泛函福克-普朗克方程

根据所写出的诸如(8.1.12)这种随机偏微分方程，我们立即提出问题，相应的福克-普朗克方程看来象什么样子？该方程一定是一个无限多个连续变量 $\rho(r)$ 的偏微分方程，这里 r 是区分各变量的连续指标。定义泛函导数的一个笨办法如下。首先，将空间分成边长为 l 的立方格子，位于 r_i 处的立方格子用 i 标示，并引入变量

$$x_i = l^3\rho(r_i) \tag{8.1.15}$$

并考虑变量 $x = \{x_i\}$ 的泛函。

现在考虑所有这些立方格子变量的泛函的计算，偏导数不难用通常的方法来定义。我们形式地引入泛函导数如下：

$$\frac{\delta F(\rho)}{\delta\rho(r_i)} = \lim_{l\to 0} l^{-3}\frac{\partial F(x)}{\partial x_i} \tag{8.1.16}$$

在大多数应用文献中，上式左边这个极限的存在在某种意义上是完全不确定的。可以给出精确的定义，并且象处理泛函问题中常用的那样，收敛性的精确定义是重要的。另外，(8.1.16)这种"显而易见"的定义是无用的。

虽然泛函计算的精确表达不属于本书范围，但是，指明写下这些方程的作者做了些什么工作是合适的。实际上，泛函导数在形式上由(8.1.16)定义，而诸如(8.1.12)这种随机微分方程的相应的离散形式由公式表示。用同样的符号，公式为

$$\mathrm{d}x_i = [\sum_j D_{ij}x_j + \tilde{F}(x_i)]\mathrm{d}t + \sum_j \tilde{g}_{ij}\mathrm{d}w_j(t) \tag{8.1.17}$$

方程中的系数 D_{ij} 产生 $D\nabla^2$ 的离散近似。选取系数 $\tilde{F}$ 和 $\tilde{g}$ 使

$$F[\rho(r_i, t)] = \lim_{l\to 0}\tilde{F}(x_i)l^{-3} \tag{8.1.18}$$

$$\vec{g}(r_i, t) = \lim_{l\to 0} l^{-3} \sum_j \vec{g}_{ij} dw_j(t) \tag{8.1.19}$$

更精确地说，我们设想一个比(8.1.14)更普遍的正确公式，即

$$\langle g(r, t)\, g(r', t)\rangle = G(r, r')\delta(t-t') \tag{8.1.20}$$

并要求

$$G(r_i, r_j) = \lim_{l\to 0} l^{-6} \sum_k \vec{g}_{ik}\vec{g}_{jk} \tag{8.1.21}$$

在这种情况下，x_i 的福克-普朗克方程为

$$\begin{aligned}\partial_t p(x) = &-\sum_{ij} \frac{\partial}{\partial x_i} \{[D_{ij}x_j + \delta_{ij}\tilde{F}(x_i)]p(x)\} \\ &+\sum_{ijk} \frac{1}{2} \frac{\partial^2}{\partial x_i \partial x_j} \vec{g}_{ik}\vec{g}_{jk} p(x)\end{aligned} \tag{8.1.22}$$

现在讨论极限 $l^3\to 0$. 某些计算给出

$$\begin{aligned}\partial_t p(\rho) = &-\int d^3r \frac{\delta}{\delta\rho(r)} \{[D\nabla^2(r) + F[\rho(r)]]P(\rho)\} \\ &+\frac{1}{2}\int\int d^3r d^3r' \left[\frac{\delta^2}{\delta\rho(r)\delta\rho(r')} G(r, r')P(\rho)\right]\end{aligned} \tag{8.1.23}$$

现在 $P(\rho)$ 是一类泛函概率，它的归一化定义要求对 $\rho(r)$ 的概率测度作仔细的说明. 这可以做到，但实际上通常把(8.1.23)理解为离散形式(8.1.22)，而且几乎所有的计算无疑都是离散的.

情况显然并不令人满意. 现在，随机微分方程及其解在数学形式上的存在已经确立了，但是，正如日常的计算工具那样，它没有获得发展. 我们希望读者查阅文献 [8.1]，以便在数学的公式化方面获得更多的信息. 然而，由于大多数工作无疑是离散化的，为了给出比较简单的符号和方便起见，我们简单地利用连续符号，基本上以离散形式用公式直接表示所有问题.

8.2 多变量主方程的描述

8.2.1 扩散

我们设想空间被分解成边长为 l、体积为 ΔV 的立方格子. 各格子用指标 i 表示，而其中的化学物质 X 的分子数称为 x_i. 于是

我们引入一多变量概率

$$P(\boldsymbol{x}, t) = P(x_1, x_2, \cdots, t) = p(x_i, \hat{\boldsymbol{x}}, t) \tag{8.2.1}$$

在上面最后的表示式中，$\hat{\boldsymbol{x}}$ 表示所有没有明确写出的其他 $\boldsymbol{x}$ 的矢量。

我们可以作一个象马尔可夫过程那样的扩散模型，在该过程中，分子在单位时间内以概率 $d_{ij}x_i$ 从格子 i 转移到格子 j，也即转移概率正比于格子内的分子数。对于严格的局部描述来说，我们预期仅当 i 和 j 为近邻时，d_{ij} 才不为零。但这并非必要，在下面的叙述中并不总是作这种假定。

借助 7.5 节中的符号，我们可以写出一个生灭主方程，其参变量由下面的代换给出：

$$\begin{aligned}
&\boldsymbol{N}(i, j) = (0\cdots0, \overset{i}{1}, 0, \cdots0, \overset{j}{0}, 0, \cdots)\\
&\boldsymbol{M}(i, j) = (0\cdots0, 0, 0, \cdots0, 1, 0, \cdots)\\
&\boldsymbol{r}(i, j) = (0\cdots0, -1, 0, \cdots0, 1, 0, \cdots)\\
&K^{+}_{(i,j)} = \mathrm{d}_{ij}\\
&K^{-}_{(i,j)} = 0
\end{aligned} \tag{8.2.2}$$

因此，主方程变为

$$\begin{aligned}
\partial_t p(\boldsymbol{x}, t) = \Sigma \mathrm{d}_{ij}[&(x_i+1)p(\hat{\boldsymbol{x}}, x_i+1, x_j-1, t)\\
&- x_i p(\boldsymbol{x}, t)]
\end{aligned} \tag{8.2.3}$$

这是一个简单的线性主方程，可以用不同的方法求解。

应注意，由于

$$\begin{aligned}
&\boldsymbol{r}^{(i,j)} = -\boldsymbol{r}^{(j,i)}\\
&\boldsymbol{N}^{(i,j)} = \boldsymbol{M}^{(j,i)}
\end{aligned} \tag{8.2.4}$$

因此，我们也可以限制 $i > j$ 并设 $K^{-}_{(i,j)} = d_{ji}$。从 (7.5.15, 18) 看到，用这种书写形式，若

$$d_{ij}\langle x_i\rangle_S = d_{ji}\langle x_j\rangle_S \tag{8.2.5}$$

成立，则细致平衡就满足。

在一个进行扩散的系统中，平稳解是齐次的，也即 $\langle x_i\rangle_S = \langle x_j\rangle_S$。

因此，细致平衡要求

$$d_{ij} = d_{ji} \tag{8.2.6}$$

而(8.2.3)具有多变量泊松平稳解。

均值方程为

$$d_t\langle x_i(t)\rangle = \sum_{j,k} r_i^{(j,k)}[\langle t_{jk}^{+}(x) - t_{jk}^{-}(x)\rangle]$$

$$= \sum_{j,k}(-\delta_{ij} + \delta_{ik}) d_{jk}\langle x_j\rangle$$

因此，

$$d_t\langle x_i(t)\rangle = \sum_j (d_{ji} - \delta_{ji}\sum_k d_{jk})\langle x_j(t)\rangle \tag{8.2.7}$$

$$= \sum_j D_{ji}\langle x_j(t)\rangle \tag{8.2.8}$$

8.2.2 扩散主方程的连续形式

假定格子 i 的中心位于 r_i 并作代换

$$x_i(t) = l^3\rho(r_i, t) \tag{8.2.9}$$

并假定 $d_{ij} = 0$(i, j 非最近邻)

$= d$(i, j 近邻)

因此在 $l \to 0$ 的极限下，(8.2.8)变成

$$\partial_t\langle\rho(r, t)\rangle = D\nabla^2\langle\rho(r, t)\rangle \tag{8.2.10}$$

其中

$$D = l^2 d \tag{8.2.11}$$

这样，扩散方程就恢复到原状，我们将把这个结果作简单的推广。

a) 克雷默斯-莫耶尔方程或按系统大小展开的方程

我们需要一个参量，用它来适当地量度个数和转移概率。可能存在两种极限，它们都对应于分子数的增加：

i) 大格子极限：浓度固定时 $l \to \infty$ 的极限；

ii) 格子大小固定时的高浓度极限。

对纯扩散而言，两者的结果相同。在任一种情况下

$$t_{i,j}^{+}(x) \to \infty \tag{8.2.12}$$

并且系统大小展开是可能的。对于最低级的情况，它等价于克雷

默斯-莫耶尔展开式。根据(7.5.31, 32)，我们发现

$$A_i(x) = -\sum_j D_{ji} x_j \tag{8.2.13}$$

$$B_{lm}(x) = \delta_{lm}\sum_j (D_{lj}x_l + D_{jl}x_j) - D_{lm}x_l - D_{ml}x_m \tag{8.2.14}$$

式中

$$D_{jl} = d_{jl} - \delta_{jl}\sum_k d_{lk} \tag{8.2.15}$$

于是在这种极限下，$p(x, t)$遵循福克-普朗克方程

$$\partial_t P = -\sum_i \partial_i A_i(x) P + \frac{1}{2}\sum_{l,m} \partial_l \partial_m B_{lm}(x) P \tag{8.2.16}$$

b) 克雷默斯-莫耶尔展式的连续形式

通过把点 r 和格子 i 联系起来并写出

$$l^3 \sum_i \to \int \mathrm{d}^3 r \tag{8.2.17}$$

$$D_{ij} \to D(r, r') \equiv \mathscr{D}(r', r - r') \tag{8.2.18}$$

$$l^{-3}\delta_{ij} \to \delta(r - r') \tag{8.2.19}$$

从而引入连续形式。在这步骤中我们对 D_{ij} 来作任何特殊对称性假设，因此各向异性的非均匀扩散也包括在内了。

然而，有着某些由"扩散"概念的含义产生的要求。

i) 仅当存在浓度梯度时，扩散才能被观察到。这意味着稳态与恒定的浓度相对应，而由式(8.2.13, 15)，这意味着

$$\sum_j D_{jl} = 0 \tag{8.2.20}$$

即

$$\sum_j \mathrm{d}_{jl} = \sum_j \mathrm{d}_{lj} \tag{8.2.21}$$

请注意，细致平衡(8.2.6)隐含这些条件。

ii) 扩散不改变系统中物质的总量，即

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_i x_i = 0 \tag{8.2.22}$$

上式对 x_i 的任何体积必然成立。根据均值方程，上式要求

$$\sum_j D_{ij} = 0 \tag{8 2.23}$$

该式由(8.2.15)及(8.2.21)得出。

iii) 在连续符号表示法下，对任何 r 来说，(8.2.20) 意味着

$$\int \mathrm{d}^3\delta \mathscr{D}(r+\delta, -\delta) = 0 \tag{8.2.24}$$

而根据 (8.2.23)，我们也得到

$$\int \mathrm{d}^3\delta \mathscr{D}(r, \delta) = 0 \tag{8.2.25}$$

iv) 如果细致平衡成立，则将(8.2.6)代入 D 的定义，把所得的方程代替(8.2.24)，即

$$D_{ij} = D_{ji} \tag{8.2.26}$$

上式的连续形式为

$$\mathscr{D}(r+\delta, -\delta) = \mathscr{D}(r, \delta) \tag{8.2.27}$$

现在用与克雷默斯-莫耶尔展开相类似的方法，可得到连续形式的导数。

我们定义导数矩

$$M(r) = \int \mathrm{d}^3\delta\, \delta \mathscr{D}(r, \delta) \tag{8.2.28}$$

$$\underset{\sim}{D}(r) = \frac{1}{2}\int \mathrm{d}^3\delta\, \delta\delta \mathscr{D}(r, \delta) \tag{8.2.29}$$

并假定在某种合适的极限下高阶导数矩为零，这与克雷默斯-莫耶尔展开中的极限相似。

细致平衡的要求(8.2.27)给出

$$\begin{aligned} M(r) &= \int \mathrm{d}^3\delta\delta \mathscr{D}(r+\delta, -\delta) \\ &= \int \mathrm{d}^3\delta\delta [\mathscr{D}(r, -\delta) + \delta\cdot\nabla\mathscr{D}(r, -\delta) + \cdots] \\ &= -M(r) + 2\nabla\cdot\underset{\sim}{D}(r) + \cdots \end{aligned} \tag{8.2.30}$$

因此，细致平衡要求

$$M(r) = \nabla\cdot\underset{\sim}{D}(r) \tag{8.2.31}$$

同样地较弱的要求(8.2.24)需要较弱的条件

$$\nabla[M(r) - \nabla\cdot\underset{\sim}{D}(r)] = 0 \tag{8.2.32}$$

现在我们可以构造 $A_i(x)$ 的连续形式：

$$A_i(x) \to -\int d^3\delta \mathscr{D}(r, \delta)\rho(r+\delta)$$

$$= -M(r)\cdot\nabla\rho(r) - \underset{\sim}{D}(r)\cdot\nabla\nabla\rho(r) \qquad (8.2.33)$$

如果细致平衡正确，则由(8.2.31)我们可以把上式改写成

$$\boxed{A_i(x) \to -\nabla\cdot[\underset{\sim}{D}(r)\cdot\nabla\rho(r)]} \qquad (8.2.34)$$

不存在细致平衡时的一般形式可以从下面的定义得到

$$J(r) = M(r) - \nabla\cdot\underset{\sim}{D}(r) \qquad (8.2.35)$$

根据 (8.2.32)，

$$\nabla\cdot J(r) = 0 \qquad (8.2.36)$$

以致我们可以写出

$$J(r) = \nabla\cdot\underset{\sim}{E}(r) \qquad (8.2.37)$$

式中 $E(r)$ 是反对称张量。经代入，由定义

$$\underset{\sim}{H}(r) = \underset{\sim}{D}(r) + \underset{\sim}{E}(r) \qquad (8.2.38)$$

我们发现已经定义了一个非对称的扩散张量 $\underset{\sim}{H}(r)$，并且得到

$$A_i(x) \to -\nabla\cdot[\underset{\sim}{H}(r)\cdot\nabla\rho(r)] \qquad (8.2.39)$$

这意味着在决定论的意义上

$$\boxed{\partial_t\rho(r, t) = \nabla\cdot[\underset{\sim}{H}(r)\cdot\nabla\rho(r, t)]} \qquad (8.2.40)$$

式中 $\underset{\sim}{H}(r)$ 在细致平衡成立时为对称的。

现在我们来讨论由 (8.2.14) 给出的涨落项。为了计算 $B_{im}(x)$，我们首先考虑极限

$$l^3\sum_m B_{im}\phi_m \to \int dr' B(r, r')\phi(r') \qquad (8.2.41)$$

式中 ϕ_m 是一个任意函数。用类似而冗长得多的计算，最后我们得

$$l^3\sum_m B_{im}\phi_m \to -2\nabla\ [\underset{\sim}{D}(r)\rho(r)\cdot\nabla\phi(r)] \qquad (8.2.42)$$

以致

$$\boxed{B(r, r') = 2\nabla'\nabla : [\underset{\sim}{D}(r)\rho(r)\delta(r-r')]} \qquad (8.2.43)$$

至此，8.1 节的唯象理论有了一个合理的基础因为(8.2.43)是根据下面的假设得到的：

$$\boldsymbol{j}(\boldsymbol{r},t) = -\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r})\nabla\rho(\boldsymbol{r},\ t) - \xi(\boldsymbol{r},\ t) \tag{8.2.44}$$

其中

$$\langle\xi(\boldsymbol{r},\ t)\xi(\boldsymbol{r}',\ t')\rangle = 2\delta(t-t')\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r})\delta(\boldsymbol{r}-\boldsymbol{r}')\rho(\boldsymbol{r}) \tag{8.2.45}$$

并因此

$$\partial_t\rho(\boldsymbol{r},\ t) = \nabla\cdot\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r})\ \nabla\rho + \nabla\cdot\xi(\boldsymbol{r},\ t) \tag{8.2.46}$$

这相应于不存在细致平衡时的非均匀各向异性的扩散理论。设

$$\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r}) = D\mathbf{1} \tag{8.2.47}$$

则通常可以简化上式并给出更熟悉的方程。注意，按照(8.2.45)，除非$\underset{\sim}{\boldsymbol{D}}$是对角的，否则流的不同成分中的涨落一般来说是相关的。

与涨落耗散论证的比较。结论(8.2.43)几乎可以从稳态的简单涨落-耗散论证得到，这里我们知道涨落是泊松型的。在这种情况下，

$$\langle x_i,\ x_i\rangle = \langle x_i\rangle\delta_{ij} \tag{8.2.48}$$

相应于

$$g(\boldsymbol{r},\ \boldsymbol{r}') = \langle\rho(\boldsymbol{r}),\ \rho(\boldsymbol{r}')\rangle = \delta(\boldsymbol{r}-\boldsymbol{r}')\langle\rho(\boldsymbol{r})\rangle$$

由于是线性理论，所以我们可应用4.46节中的(4.4.51)式，那儿的矩阵A和A^T变为

$$\begin{aligned} A &\to -\nabla\cdot\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r})\cdot\nabla \\ A^T &\to -\nabla'\cdot\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r}')\cdot\nabla' \end{aligned} \tag{8.2.49}$$

于是

$$\begin{aligned} B(\boldsymbol{r},\ \boldsymbol{r}') &\to BB^T \\ &= A\sigma + \sigma A^T \to [-\nabla\cdot\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r})\cdot\nabla - \nabla'\cdot\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r}')\cdot\nabla']g(\boldsymbol{r},\ \boldsymbol{r}') \end{aligned} \tag{8.2.50}$$

我们应注意，在平稳态$\langle\rho(\boldsymbol{r})\rangle = \langle\rho\rangle$，与$\boldsymbol{r}$无关。于是

$$\begin{aligned} B(\boldsymbol{r},\ \boldsymbol{r}') &= [-\nabla\cdot\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r})\cdot\nabla\langle\rho\rangle\delta(\boldsymbol{r}-\boldsymbol{r}') \\ &\qquad -\nabla'\cdot\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r}')\cdot\nabla'\langle\rho\rangle\delta(\boldsymbol{r}-\boldsymbol{r}')] \\ &= \nabla\nabla' : \{[\underset{\sim}{\boldsymbol{H}}(\boldsymbol{r}) + \boldsymbol{H}^{\dagger}(\boldsymbol{r})]\langle\rho\rangle\delta(\boldsymbol{r}-\boldsymbol{r}')\} \\ &= 2\nu\nabla\nabla' : [\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r})\langle\rho\rangle\delta(\boldsymbol{r}-\boldsymbol{r}')] \end{aligned} \tag{8.2.51}$$

然而这结果并不如(8.2.43)那样普遍，这是因为它只适用于稳态，而且因为它计算得到的是 $\langle\rho\rangle$ 而不是 $\rho(r)$，所以在稳态也不相同。但是，由于福克-普朗克公式仅在大格子尺度极限时才适用，这时涨落很小，所以(8.2.51)的精度与(8.2.43)相符。

8.2.3 反应与扩散的结合

设想一个格子内的分子按照与第7章相似的主方程互相进行反应，我们介绍这种反应并希望讨论几种化学成分，所以引入符号

$$\boldsymbol{X}=(\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \cdots) \tag{8.2.52}$$

式中 $\boldsymbol{x}_i$ 表示一个向量，它的分量 $x_{i,a}$ 是格子 i 内 $\boldsymbol{x}_a$ 处的分子数。于是我们写出

$$\begin{aligned}\partial_t P(\boldsymbol{X},\ t)=\partial_t P(\text{扩散})+\sum_i\{&\sum_A t_A^+(\boldsymbol{x}_i-\boldsymbol{r}^A)P(\boldsymbol{x}_i-\boldsymbol{r}^A,\ \hat{\boldsymbol{X}})\\&+t_A^-(\boldsymbol{x}_i+\boldsymbol{r}^A)P(\boldsymbol{x}_i+\boldsymbol{r}^A,\ \hat{\boldsymbol{X}})\\&-[t_A^+(\boldsymbol{x}_i)+t_A^-(\boldsymbol{x}_i)]P(\boldsymbol{X})\}\end{aligned} \tag{8.2.53}$$

式中扩散部分具有(8.7.3)的形式，而求和对各分量进行。

这就通过克雷默尔-莫耶尔展式得到了具有通常漂移和扩散的福克-普朗克方程，和 (7.5.32) 所给出的福克-普朗克方程相同。

借助于与空间有关的维纳过程 $\boldsymbol{W}(\boldsymbol{r},\ t)$，我们可以写出等价的随机偏微分方程如下：我们考虑一个各向同性的恒定扩散张量

$$\underset{\sim}{\boldsymbol{D}}(\boldsymbol{r})=D\mathbf{1} \tag{8.2.54}$$

$$\mathrm{d}\rho_a=[D_a\nabla^2\rho_a+\sum_A\nu_a^A(\kappa_A^+\prod_a\rho_a^{N_a^A}-\kappa_A^-\prod_a\rho_a^{M_a^A})]\mathrm{d}t+\mathrm{d}W_a(\boldsymbol{r},\ t) \tag{8.2.55}$$

而

$$\begin{aligned}\mathrm{d}W_a(\boldsymbol{r},\ t)\mathrm{d}W_b(\boldsymbol{r}',\ t)=\{&2\delta_{a,b}\nabla'\cdot\nabla[D_c\rho_a(\boldsymbol{r})\delta(\boldsymbol{r}-\boldsymbol{r}')]\\&+\delta(\boldsymbol{r}-\boldsymbol{r}')\sum_A\nu_a^A\nu_b^A(\kappa_A^+\prod_a\rho_a^{N_a^A}+\kappa_A^-\prod_a\rho_a^{M_a^A})\}\mathrm{d}t.\end{aligned} \tag{8.2.56}$$

朗之万方程的有效性取决于该系统大小展式。方程 (8.2.55, 56) 与含有 Ω 的化学反应速率常数的特殊标度有关，Ω 由 7.5.3 节

(7.5.29)给出。在这种情况下，只有把 Ω 解释成

$$\Omega = l^3 = \text{格子的体积}$$

才是有效的。然而应注意，在相同的时刻，由于 $l^2 d$ 必须保持与扩散系数相等，而来自化学反应的那些项象 l^3 那样量度，所以扩散部分应象 l 那样量度。这意味着当格子的体积增加时，我们使扩散的影响越来越小，但是仍比化学反应部分的相应项的影响要大，而化学反应部分的项比前者小 l^3 的整数次幂。

比较扩散和反应的精确方法将在稍后论述。

例子：$X_1 \rightleftarrows 2X_2$。对于这种化学反应，我们发现（利用7.5节的方法）

$$\boldsymbol{N} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{M} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad \boldsymbol{r} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \tag{8.2.57}$$

$$\boldsymbol{K}^+ = \boldsymbol{K}_1 = \boldsymbol{K}_1 \Omega^{-2+1}$$

于是作代换

$$\begin{aligned} \mathrm{d}\rho_1(\boldsymbol{r}) &= (D_1 \nabla^2 \rho_1 - \kappa_1 \rho_1 + \kappa_2 \rho_2^2)\mathrm{d}t + \mathrm{d}W_1(\boldsymbol{r}, t) \\ \mathrm{d}\rho_2(\boldsymbol{r}) &= (D_2 \nabla^2 \rho_2 + 2\kappa_1 \rho_1 - 2\kappa_2 \rho_2^2)\mathrm{d}t + \mathrm{d}W_2(\boldsymbol{r}, t) \end{aligned} \tag{8.2.58}$$

式中

$$\mathrm{d}\boldsymbol{W}(\boldsymbol{r}, t)\mathrm{d}\boldsymbol{W}^T(\boldsymbol{r}, t) \approx \begin{bmatrix} 2\nabla \cdot \nabla' [D_1 \rho_1 \delta(\boldsymbol{r} - \boldsymbol{r}')] + \delta(\boldsymbol{r} - \boldsymbol{r}')[\kappa_1 \rho_1 + \kappa_2 \rho_2^2], & -2\delta(\boldsymbol{r} - \boldsymbol{r}')[\kappa_1 \rho_1 + \kappa_2 \rho_2^2] \\ -2\delta(\boldsymbol{r} - \boldsymbol{r}')[\kappa_1 \rho_1 + \kappa_2 \rho_2^2], & 2\nabla \cdot \nabla' [D_2 \rho_2 \delta(\boldsymbol{r} - \boldsymbol{r}')] + 4\delta(\boldsymbol{r} - \boldsymbol{r}')[\kappa_1 \rho_1 + \kappa_2 \rho_2] \end{bmatrix} \mathrm{d}t \tag{8.2.59}$$

这方程仅当以 Ω^{-1} 作为系统大小展式时才适合，这就是说，只有把格子的大小以及连续极限看作为一个符号时才合适。用了这个符号才能理解我们的真正意思。对于一个格子模型和一个尺度显然很小的格子，我们用足够大的标度来处理它，虽然格子本身大到足以容纳许多分子。

于是，这类方程实际上仅作为确定性态的线性化方程时才成立。这些确定性态具有这样的形式，凯塞[Keizer][8.2]用它把化

学反应表示成公式。在这方面，泊松表示比较好，因为它给出了与主方程完全等价的方程。

8.2.4 泊松表示法

对于一个反应，其步骤不多于双分子反应，由 (7.7.9) 得到前面关于空间扩散的相当简单的福克-普朗克方程 [利用公式 (8.2.2)]。我们发现扩散矩阵消失了。于是，一个空间相关系统的一般表示式可用密度变量

$$\eta_2(\boldsymbol{r}) = 2a(\boldsymbol{r})/l^3 \tag{8.2.60}$$

求得，并且我们发现

$$\mathrm{d}\eta_a(\boldsymbol{r}) = [D_a\nabla^2\eta_a(\boldsymbol{r}) + \sum_A \nu_a^A(\kappa_A^+\prod_a \eta_a^{N_a^A} - \kappa_A^-\prod_a \eta_a^{M_a^A})]\mathrm{d}t + \mathrm{d}W_a(\boldsymbol{r},\ t) \tag{8.2.61}$$

$$\mathrm{d}W_a(\boldsymbol{r},t)\mathrm{d}W_b(\boldsymbol{r}',\ t) = \mathrm{d}t\delta(\boldsymbol{r}-\boldsymbol{r}')\sum_A(\kappa_A^+\prod_a \eta_a^{N_a^A} - \kappa_A^-\prod_a \eta_a^{M_a^A}) \times (M_a^A M_b^A - N_a^A N_b^A - \delta_{ab}\nu_a^A) \tag{8.2.62}$$

这些方程与(8.2.55, 56) 非常相似。当写出它们的表达式时，它们的相似是明显的。例如，再次讨论 $X_1 \rightleftarrows 2X_2$，我们得

$$\mathrm{d}\eta_1(\boldsymbol{r}) = (D_1\nabla^2\eta_1 - \kappa_1\eta_1 + \kappa_2\eta_2^2)\mathrm{d}t + \mathrm{d}W_1(\boldsymbol{r},\ t) \tag{8.2.63}$$

$$\mathrm{d}\eta_2(\boldsymbol{r}) = (D_2\nabla^2\eta_2 + 2\kappa_1\eta_1 - 2\kappa_2\eta_2^2)\mathrm{d}t + \mathrm{d}W_2(\boldsymbol{r},\ t) \tag{8.2.64}$$

$$\mathrm{d}\boldsymbol{W}(\boldsymbol{r},\ t)\mathrm{d}\boldsymbol{W}^T(\boldsymbol{r}',\ t) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}(\kappa_1\eta_1 - \kappa_2\eta_2^2)\delta(\boldsymbol{r}-\boldsymbol{r}')\mathrm{d}t \tag{8.2.65}$$

当把它们与(8.2.57,58)的相应部分比较时，(8.2.63~65)的简单性是非常明显的。它们与主方程完全等价(以一连续的公式)，这一点特别值得重视。

8.3 空间和时间相关结构

这里我们希望讨论线性系统中的空间、时间及空-时相关的几个问题，当然，线性系统是完全精确可解的。重要的相关是阶乘

相关，和用矩来定义普通的相关一样，阶乘相关是借助阶乘矩来定义的。正如我们在下面的几个例子中将看到的，用阶乘矩来写出这些方程要自然得多。

8.3.1 反应 $X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} Y$

对于X和Y，我们假定它们是具有相同扩散常数的均匀而各向同性的扩散，由于反应和扩散两者都是线性的，所以我们得到了关于浓度变量η、μ（分别与X和Y相对应）并具有非随机源的泊松表示的朗之万方程，即

$$\begin{aligned}\partial_t\eta(\boldsymbol{r},\ t) &= D\nabla^2\eta - k_1\eta + k_2\mu \\ \partial_t\mu(\boldsymbol{r},\ t) &= D\nabla^2\mu + k_1\eta - k_2\mu\end{aligned} \tag{8.3.1}$$

a) 空间相关

现在我们注意

$$\begin{aligned}&\langle\eta(\boldsymbol{r},\ t)\rangle = \langle\rho_x(\boldsymbol{r},\ t)\rangle \\ &\langle\mu(\boldsymbol{r},\ t)\rangle = \langle\rho_y(\boldsymbol{r},\ t)\rangle \\ &\langle\eta(\boldsymbol{r},\ t)\eta(\boldsymbol{r}',\ t)\rangle = \langle\rho_x(\boldsymbol{r},\ t)\rho_x(\boldsymbol{r}',\ t)\rangle - \delta(\boldsymbol{r}-\boldsymbol{r}')\langle\rho_x(\boldsymbol{r},\ t)\rangle \\ &\qquad\qquad \equiv g(\boldsymbol{r},\ \boldsymbol{r}',\ t) \\ &\langle\eta(\boldsymbol{r},\ t)\mu(\boldsymbol{r}',\ t)\rangle = \langle\rho_x(\boldsymbol{r},\ t)\rho_y(\boldsymbol{r}',\ t)\rangle \equiv f(\boldsymbol{r},\ \boldsymbol{r}',\ t) \\ &\langle\mu(\boldsymbol{r},\ t)\mu(\boldsymbol{r}',\ t)\rangle = \langle\rho_y(\boldsymbol{r},\ t)\rho_x(\boldsymbol{r}',\ t)\rangle - \delta(\boldsymbol{r}-\boldsymbol{r}')\langle\rho_y(\boldsymbol{r},\ t)\rangle \\ &\qquad\qquad = h(\boldsymbol{r},\ \boldsymbol{r}',\ t)\end{aligned} \tag{8.3.2}$$

这些方程都是下列事实的连续统符号形式，这些事实是：泊松矩等于实数阶乘矩。

平均浓度方程显然与 (8.3.1) 式严格相同。现在设想一种均匀的情况，以致可以假定

$$\left.\begin{aligned}&\langle\rho_x(\boldsymbol{r},\ t)\rangle = \langle\rho_x(t)\rangle \\ &\langle\rho_y(\boldsymbol{r},\ t)\rangle = \langle\rho_y(t)\rangle \\ &g(\boldsymbol{r},\ \boldsymbol{r}',\ t) = g(\boldsymbol{r}-\boldsymbol{r}',\ t) \\ &f(\boldsymbol{r},\ \boldsymbol{r}',\ t) = f(\boldsymbol{r}-\boldsymbol{r}',\ t) \\ &h(\boldsymbol{r},\ \boldsymbol{r}',\ t) = h(\boldsymbol{r}-\boldsymbol{r}',\ t)\end{aligned}\right\} \tag{8.3.3}$$

并计算 $\langle \eta(\boldsymbol{r}, t)\eta(0, t)\rangle$ 等等的运动方程，我们很快发现

$$\frac{\partial g(\boldsymbol{r}, t)}{\partial t} = 2D\nabla^2 g(\boldsymbol{r}, t) - 2k_1 g(\boldsymbol{r}, t) + 2k_2 f(\boldsymbol{r}, t)$$

$$\frac{\partial f(\boldsymbol{r}, t)}{\partial t} = 2D\nabla^2 f(\boldsymbol{r}, t) - (k_1 + k_2) f(\boldsymbol{r}, t) + k_1 h(\boldsymbol{r}, t) + k_2 g(\boldsymbol{r}, t) \tag{8.3.4}$$

$$\frac{\partial h(\boldsymbol{r}, t)}{\partial t} = 2D\nabla^2 h(\boldsymbol{r}, t) - 2k_2 h(\boldsymbol{r}, t) + 2k_1 f(\boldsymbol{r}, t)$$

这些方程的平稳解具有形式

$$g(\boldsymbol{r}) = \xi k_2^2, \quad f(\boldsymbol{r}) = \xi k_1 k_2, \quad h(\boldsymbol{r}) = \xi k_1^2 \tag{8.3.5}$$

式中 ξ 是一任意参变量，均值的相应的平稳解为

$$\langle x(\boldsymbol{r}_i)\rangle = \lambda k_2, \quad \langle Y(\boldsymbol{r}_i)\rangle = \lambda k_1 \tag{8.3.6}$$

式中 λ 是另一个任意参量。如果 $\xi = 0$，则我们重新获得泊松情况

$$\begin{aligned} &\langle \rho_x(\boldsymbol{r}), \rho_x(\boldsymbol{r}')\rangle = \langle \rho_x(\boldsymbol{r})\rangle \delta(\boldsymbol{r} - \boldsymbol{r}') \\ &\langle \rho_y(\boldsymbol{r}), \rho_y(\boldsymbol{r}')\rangle = \langle \rho_y(\boldsymbol{r})\rangle \delta(\boldsymbol{r} - \boldsymbol{r}') \\ &\langle \rho_x(\boldsymbol{r}), \rho_y(\boldsymbol{r}')\rangle = 0 \end{aligned} \tag{8.3.7}$$

（通过选择别的 λ 值，就得到不同的解，这些解相应于对系统中总分子数的不同分布。）

对任何初始条件，很容易求得与时间有关的解。当一开始解是均匀的、非相关的和泊松情况时，(8.3.7)满足初条件，于是 f、g 和 h 开始都为零，而且将保持下去。于是，非相关的泊松形式随时间的推移而保持不变，和 7.7b 节中所导出的形式相同。

向泊松分布豫弛的问题与所设的初相关函数的特殊形式有非常密切的关系。例如，最初是非相关、而且是非泊松的系统，它由

$$g(\boldsymbol{r}, 0) = \alpha\delta(\boldsymbol{r}), \quad f(\boldsymbol{r}, 0) = \beta\delta(\boldsymbol{r}), \quad h(\boldsymbol{r}, 0) = \gamma\delta(\boldsymbol{r}) \tag{8.3.8}$$

表示。与时间有关的解为

$$\begin{bmatrix} g(\boldsymbol{r}, t) \\ f(\boldsymbol{r}, t) \\ h(\boldsymbol{r}, t) \end{bmatrix} = \frac{\exp(-r^2/8Dt)}{(8\pi Dt)^{3/2}}$$

$$\times\begin{bmatrix} k\ \varepsilon_1 - 2k_2\varepsilon_2 e^{-(k_1+k_2)t} + \varepsilon_3 e^{-2(k_1+k_2)t} \\ k_1k_2\varepsilon_1 + (k_2-k_1)\varepsilon_2 e^{-(k_1+2)t} - \varepsilon_3 e^{-2(k_1+k_2)t} \\ k_1\varepsilon_1 + 2k_1\varepsilon_2 e^{-(k_1+k_2)t} + \varepsilon_3 e^{-2(k_1+k_2)t} \end{bmatrix} \tag{8.3.9}$$

式中

$$\varepsilon_1 = (\alpha + 2\beta + \gamma)/(k_1+k_2)^2$$
$$\varepsilon_2 = [k_2(\beta+\gamma) - k_1(\alpha+\beta)]/(k_1+k_2)^2 \tag{8.3.10}$$
$$\varepsilon_2 = [k_1\alpha + k_2\gamma - 2k_1k_2\beta]/(k_1+k_2)^2$$

注解

i) 项 ε_1、ε_2 和 ε_3 分别相应于量$\langle(x_i+y_i),\ (x_j+y_j)\rangle$、$\langle(x_i+y_i),\ (k_1y_j-k_2x_j)\rangle$ 及 $\langle(k_1y_i-k_2x_i),\ (k_1y_j-k_2x_j)\rangle$ 的非相关泊松偏差，这些量实质上是密度涨落、密度涨落和化学不稳定之间的相关以及化学不稳定中的涨落。我们注意到乘有适合于各项的化学时间相依性的一个特征扩散形式。

ii) 与泊松过程相比，(8.3.8)变得微不足道，作为由(8.3.8)给出的泊松非相关形式的偏差的时间，显然决定于初始偏差的量值。然而，假定 α、β、γ、$\langle x_i\rangle$ 及 $\langle y_i\rangle$ 都具有相似的大小，我们可以作如下粗略的估计。我们考虑一个半径为 R 的小球的体积，但比我们的基本格子大得多，这样，在这个小体积 V 中，我们发现

$$\mathbf{var}\{x[V,0]\} = \int_V \mathrm{d}^3r \int_V \mathrm{d}^3r' \langle \rho_x(r,0),\ \rho_x(r',0)\rangle, \text{ 即}$$
$$\mathbf{var}\{x[V,0]\} = \langle x[V,0]\rangle + 2V \tag{8.3.11}$$

以及类似地

$$\mathbf{var}\{y[V,0]\} = \langle y[V,0]\rangle + \gamma V$$
$$\langle x[V,0],\ y[V,0]\rangle = \beta V$$

而在时间 $t \gg R^2/4D$ 以后，这些量近似地满足

$$\left.\begin{aligned} &\mathbf{var}\{x[V,t]\} \simeq \langle x[V,t]\rangle \\ &\quad + \frac{V^2}{(8\pi Dt)^{3/2}}\ (k_1^2\varepsilon_1 - 2k_2\varepsilon_2 e^{-(k_1+k_2)t} + \varepsilon_3 e^{-2(k_1+k_2)t}) \\ &\mathbf{var}\{y[V,t]\} \simeq \langle y[V,t]\rangle \\ &\quad + \frac{V^2}{(8\pi Dt)^{3/2}}\ (k_1^2\varepsilon_1 + 2k_1\varepsilon_2 e^{-(k_1+k_2)t} + \varepsilon_3 e^{-2(k_1+k_2)t}) \end{aligned}\right\}$$

$$\left.\begin{aligned}\langle x[V,\ t],\ y[V,\ t]\rangle \simeq & \frac{V^2}{(8\pi Dt)^{3/2}} \\ & \times [k_1k_2\varepsilon_1 + (k_2 - k_1)\varepsilon_2 e^{-(k_1+k_2)t} - \varepsilon_3 e^{-2(k_1+k_2)t}]\end{aligned}\right\} \tag{8.3.12}$$

于是，扩散把由泊松非相关行为引起的所有偏差减少到原来的 $R^3/(Dt)^{3/2}$ 分之一。但是应注意，在初条件为非泊松，因而也是非相关的，即相应于 $\beta=0$ 的情况下，我们发现 X 和 Y 之间出现了相关。如果化学反应速率常数足够大，则这种相关可以非常显著。

b) 空间-时间相关

由于这里的运动方程是线性的，所以我们可以利用3.74节所阐述的线性理论。定义平稳双时相关矩阵 $\boldsymbol{G}(\boldsymbol{r},\ t)$

$$\boldsymbol{G}(\boldsymbol{r},\ t) = \begin{bmatrix} \langle \rho_x(\boldsymbol{r},\ t),\ \rho_x(0,\ 0)\rangle_s & \langle \rho_y(\boldsymbol{r},\ t),\ \rho_x(0,\ 0)\rangle \\ \langle \rho_x(\boldsymbol{r},\ t),\ \rho_y(0,\ 0)\rangle_s & \langle \rho_y(\boldsymbol{r},\ t),\ \rho_y(0,\ 0)\rangle \end{bmatrix} \tag{8.3.13}$$

因此，相应于(3.7.63)的方程为

$$\partial_t \underset{\sim}{G}(\boldsymbol{r},\ t) = \begin{bmatrix} D\nabla^2 - k_1 & k_2 \\ k_1 & D\nabla^2 - k_2 \end{bmatrix} \underset{\sim}{G}(\boldsymbol{r},\ t) \tag{8.3.14}$$

通过傅里叶变换和用标准方法求解最后的一阶微分矩阵方程，我们可以求得上述方程的解。而该方程具有(8.3.7)所给的 $t=0$ 时的边界条件。结果为

$$\begin{aligned}\underset{\sim}{G}(\boldsymbol{r},\ t) = & \left(\frac{\langle \rho_x\rangle + \langle \rho_y\rangle}{(k_1+k_2)^2}\right) \frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}} \\ & \begin{bmatrix} k_2^2 + k_1k_2e^{-(k_1+k_2)t} & k_1k_2(1 - e^{-(k_1+k_2)t}) \\ k_1k_2(1 - e^{-(k_1+k_2)t}) & k_1^2 + k_1k_2e^{-(k_1+k_2)t} \end{bmatrix}\end{aligned} \tag{8.3.15}$$

如果我们定义变量

$$n(\boldsymbol{r},\ t) = \rho_x(\boldsymbol{r},\ t) + \rho_y(\boldsymbol{r},\ t) \tag{8.3.16}$$

$$c(\boldsymbol{r},\ t) = [k_1\rho_x(\boldsymbol{r},\ t) - k_2\rho_y(\boldsymbol{r},\ t)]/(k_1+k_2) \tag{8.3.17}$$

则解(8.3.15)可以写成

$$\langle n(\boldsymbol{r},\ t),\ n(\boldsymbol{0},\ 0)\rangle_s = \langle n\rangle_s \frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}} \tag{8.3.18}$$

$$\langle n(\boldsymbol{r},\ t),\ c(\boldsymbol{0},\ 0)\rangle_s = \langle c(\boldsymbol{r},\ t),\ n(\boldsymbol{0},\ 0)\rangle_s = 0 \tag{8.3.19}$$

$$\langle c(\boldsymbol{r},\ t),\ c(\boldsymbol{0},\ 0)\rangle_s = \frac{k_1 k_2 \langle n\rangle_s}{(k_1+k_2)^2}\,\frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}}\,e^{-(k_1+k_2)t} \tag{8.3.20}$$

变量 n 和 c 分别对应于总密度和化学不稳定密度，即 $\langle c\rangle_s = 0$。这样，我们就看到，(8.3.18)给出了密度涨落相关，这种相关与由纯扩散所引起的相同。(8.3.20) 给出了化学不稳定时的涨落相关，而(8.3.19)表明这两种涨落是独立的。一个特征扩散项乘上这两种涨落，这种结果的简单性取决于不同种类扩散常数的同一性。

8.3.2 反应 $B+X \underset{k_3}{\overset{k_2}{\rightleftharpoons}} C$，$A+X \xrightarrow[k_4]{} 2X$

这个反应已经在 7.6.4a 节中讨论过，它没有空间相关性，我们发现浓度变量 $\eta(\boldsymbol{r},\ t)$ 的泊松表示方程为[根据(8.2.60～62)，(7.6.55, 56)]

$$\mathrm{d}\eta(\boldsymbol{r},\ t) = [D\nabla^2\eta(\boldsymbol{r},\ t) + (K_2 - K_1)\eta(\boldsymbol{r},\ t) + K_3]\mathrm{d}t + \mathrm{d}W(\boldsymbol{r},\ t) \tag{8.3.21}$$

$$\mathrm{d}W(\boldsymbol{r},\ t)\mathrm{d}W(\boldsymbol{r}',\ t) = 2\mathrm{d}t\,\delta(\boldsymbol{r}-\boldsymbol{r}')K_2\eta(\boldsymbol{r},\ t) \tag{8.3.22}$$

式中

$$\left.\begin{aligned} k_3 C &= K_2 l^3 \\ k_1 B &= K_1 \\ k_2 A &= K_2 \end{aligned}\right\} \tag{8.3.23}$$

因为系统是线性的，所以它可以精确求解，但是由于第二个反应以正比于 X 的速率消灭了 X，所以我们确实得到了泊松表示中的噪声项。

[在克雷默斯-莫耶尔方法中，我们将发现 $\rho(\boldsymbol{r},\ t)$ 与 (8.3.21) 具有相同形式的方程，但 $\mathrm{d}W(\boldsymbol{r},\ t)$ 将满足

$$\mathrm{d}W(\boldsymbol{r},\ t)\mathrm{d}W(\boldsymbol{r}',\ t) = \mathrm{d}t\{2\nabla'\cdot\nabla[D\rho(\boldsymbol{r})\delta(\boldsymbol{r}-\boldsymbol{r}')] + \delta(\boldsymbol{r}-\boldsymbol{r}')[(K_2+K_1)\eta(\boldsymbol{r},\ t)+K_3]\} \tag{8.3.24}]$$

a) 空间相关

现在定义

$$g(\boldsymbol{r}, \boldsymbol{r}', t) = \langle \rho(\boldsymbol{r}, t), \rho(\boldsymbol{r}', t)\rangle - \delta(\boldsymbol{r}-\boldsymbol{r}')\langle \rho(\boldsymbol{r}, t)\rangle$$
$$= \langle \eta(\boldsymbol{r}, t), \eta(\boldsymbol{r}', t)\rangle$$
$$= \langle \eta(\boldsymbol{r}, t)\eta(\boldsymbol{r}', t)\rangle - \langle \eta(\boldsymbol{r}, t)\rangle\langle \eta(\boldsymbol{r}', t)\rangle \tag{8.3.25}$$

我们讨论平稳齐次的情况，这时显然有

$$\langle \eta(\boldsymbol{r}, t)\rangle_s = \langle \rho(\boldsymbol{r}, t)\rangle_s = \langle \rho\rangle_s \tag{8.3.26}$$

因此，

$$\mathrm{d}g(\boldsymbol{r}, \boldsymbol{r}', t) = \mathrm{d}\langle \eta(\boldsymbol{r}, t)\eta(\boldsymbol{r}', t)\rangle$$
$$= \langle \eta(\boldsymbol{r}, t)\mathrm{d}\eta(\boldsymbol{r}', t)\rangle + \langle \mathrm{d}\eta(\boldsymbol{r}, t)\eta(\boldsymbol{r}'t)\rangle$$
$$+ \langle \mathrm{d}\eta(\boldsymbol{r}, t)\mathrm{d}\eta(\boldsymbol{r}', t)\rangle \tag{8.3.27}$$

而利用常用的伊藤法则及(8.3.21, 22)，上式

$$= [D\nabla^2 + D\nabla'^2 + 2(K_2 - K_1)]\langle \eta(\boldsymbol{r}, t)\eta(\boldsymbol{r}', t)\rangle$$
$$+ 2K_3\langle \rho\rangle_s \mathrm{d}t + 2K_2\delta(\boldsymbol{r}-\boldsymbol{r}')\langle \rho\rangle_s \mathrm{d}t \tag{8.3.28}$$

注意

$$\langle \rho\rangle_s = K_3/(K_1 - K_2) \tag{8.3.29}$$

以及在空间齐次的情况下，$g(\boldsymbol{r}, \boldsymbol{r}', t)$ 仅可能是 $\boldsymbol{r}-\boldsymbol{r}'$ 的函数，称为 $g(\boldsymbol{r}, t)$。

利用(8.3.25, 26) 作代换得：

$$\partial_t g(\boldsymbol{r}, t) = 2[D\nabla^2 + (K_2 - K_1)]\, g(\boldsymbol{r}, t) + 2K_2\langle \rho\rangle_s \delta(\boldsymbol{r}) \tag{8.3.30}$$

通过把平稳解 $g_s(\boldsymbol{r})$ 表示成傅里叶积分，就很容易求得：

$$g_s(\boldsymbol{r}) = \int \mathrm{d}^3\boldsymbol{q}\mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{r}}\tilde{g}_s(\boldsymbol{q}) \tag{8.3.31}$$

利用

$$\delta(\boldsymbol{r}) = (2\pi)^{-3}\int \mathrm{d}^3\boldsymbol{q}\mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{r}} \tag{8.3.32}$$

(8.3.31)给出

$$\tilde{g}_s(\boldsymbol{q}) = \frac{K_2\langle \rho\rangle_s}{(2\pi)^3}\,\frac{1}{Dq^2 + K_1 - K_2} \tag{8.3.33}$$

它的反演式为

$$g_s(r) = \frac{K_2\langle\rho\rangle_s}{4\pi D r}\exp\left[-r\left(\frac{K_1-K_2}{D}\right)^{1/2}\right] \tag{8.3.34}$$

因此，

$$\langle\rho(r,\ t),\ \rho(r',\ t)\rangle_s = \delta(r-r')\langle\rho\rangle_s + \frac{K_2\langle\rho\rangle_s}{4\pi D|r-r'|}\exp\left[-|r-r'|\left[\frac{K_1-K_2}{D}\right]^{1/2}\right] \tag{8.3.35}$$

注解

i) 我们应注意两个不同部分：相应于泊松涨落的 δ 相关部分，它在不同的空间点是独立的；而另一部分，具有特征相关长度 l_c 的相关项，l_c 为

$$l_c = \sqrt{D}/(K_1-K_2) \tag{8.3.36}$$

ii) 趋于平衡：当 $K_2\to 0$ 时，我们得到一个简单的可逆反应 $B+X\rightleftarrows C$。我们知道，(8.3.34)中的相关变为零。然而，相关长度本身并不消失。

iii) 局部或整体涨落：尼柯里斯 (Nicolis) 原先提出了如下问题[8.3]。考虑体积 V 内的总分子数

$$x[V,\ t] = \int_V \mathrm{d}^3 r\rho(r,\ t) \tag{8.3.37}$$

接着我们希望知道这个数是怎样变化的。我们不难知道

$$\mathrm{var}\{x(V)\}_s = \int_V \mathrm{d}^3 r\int_V \mathrm{d}^3 r'\langle\rho(r,\ t),\ \rho(r',\ t)\rangle_s \tag{8.3.38}$$

并利用 (8.3.35)，

$$\text{上式} = \langle x(V)\rangle_s + \frac{K_2\langle\rho\rangle_s}{4\pi D}\int_V \mathrm{d}^3 r\int_V \mathrm{d}^3 r'|r-r'|^{-1}\times\exp(-|r-r'|/l_c) \tag{8.3.39}$$

对上式可以计算两种极限。如果 V 比 l^3 大得多，则应注意到，当 $r\to\infty$ 时 $g(r)\to 0$，我们对(8.3.30)的平稳部分进行积分，并略去由拉普拉斯分部积分所得到的曲面项：

$$0 = 2(K_2-K_1)\int_V \mathrm{d}^3 r\int_V \mathrm{d}^3 r' g(r-r') + 2K_2\langle\rho\rangle_s\int_V \mathrm{d}^3 r' \tag{8.3.40}$$

于是，

$$\int_V\int_V \mathrm{d}^3r\mathrm{d}^3r' g(r-r') \sim \frac{K_2\langle x(V)\rangle_s}{K_1-K_2} \tag{8.3.41}$$

所以

$$\mathrm{var}\{x(V)\}_s \sim \frac{K_1\langle x(V)\rangle_s}{K_1-K_2} \quad (V \gg l^3) \tag{8.3.42}$$

这小体积极限是更为巧妙的方法。然而，对于半径 $R \ll l_c$ 的球形体积来说，我们可以忽略(8.3.39)中的指数，而且

$$\begin{aligned}
&\int_V\int_V \mathrm{d}^3r\mathrm{d}^3r' |r-r'|^{-1} \\
&\quad = \int_V \mathrm{d}^3r' \int_0^R 4\pi r^2 \mathrm{d}r\mathrm{d}(\cos\theta)(r^2+r'^2-2rr'\cos\theta)^{-1/2} \\
&\quad = \int_V \mathrm{d}^3r' \int_0^R \frac{4\pi r}{r'}\mathrm{d}r(|r+r'|-|r-r'|) \\
&\quad = 2R^5(4\pi)^2/5
\end{aligned} \tag{8.3.43}$$

所以

$$\mathrm{var}\{x(V)\}_s \sim \langle x(V)\rangle\left(1+\frac{2R^2K_2}{5D}\right) \quad \left(V=\frac{4}{3}\pi R^3 \ll l^3\right) \tag{8.3.44}$$

由此可见，在小体积的情况下，涨落是泊松的；但在大体积时，涨落是非泊松的。而且实际上，方差与(7.6.72)完全一样，它由相同反应的整体主方程给出。

实际上，对于任意球形体积，方差可以直接算出（利用富里叶变换方法），并由下式给出

$$\begin{aligned}
\mathrm{var}\{x(V)\} = \langle x(V)\rangle \Big\{1+\frac{3K_2l_c^5}{2DR^3}\Big[1-\Big(\frac{R}{l_c}\Big)^2 \\
+\frac{2}{3}\Big(\frac{R}{l_c}\Big)^3 - \mathrm{e}^{-2R/l_c}\Big(1+\frac{R}{l_c}\Big)^2\Big]\Big\}
\end{aligned} \tag{8.3.45}$$

而且这个式子给出了更精确的大体积渐近公式

$$\mathrm{var}\{x(V)\} \sim \langle x(V)\rangle\left(\frac{K_1}{K_1-K_2} - \frac{3K_1l_c}{2R(K_1-K_2)}\right) \quad (V \gg l_c^3) \tag{8.3.46}$$

注意，$\langle \rho(r,\ t),\ \rho(r,\ t)\rangle$ 的傅里叶变换是通过加进了 $g_s(r)$ 和 $\delta(r)\langle\rho\rangle_s$ 的傅里叶变换后给出的，由此可以说明上述结果，并且它

显然为

$$\langle\rho\rangle_q(2\pi)^{-3}\frac{Dq^2+k_1}{Dq^2+k_1-k_2} \tag{8.3.47}$$

这意味着对于小的 q，即长波长，上式近似为

$$\langle\rho\rangle_q(2\pi)^{-3}\frac{k_1}{k_1-k_2} \tag{8.3.48}$$

它是

$$\langle\rho\rangle_r\frac{k_1}{k_1-k_2}\delta(r-r') \tag{8.3.49}$$

的傅里叶变换，上式对应于整体主方程的相同的方差。

对于大的 q，即短波长，我们发现

$$\langle\rho\rangle_q(2\pi)^{-3} \tag{8.3.50}$$

与

$$\langle\rho\rangle_r\delta(r-r') \tag{8.3.51}$$

相对应，即为泊松涨落。在物理上，这种差别来源于 8.2.3 节所提到的主方程的扩散和反应部分的不同标度。因此，在小体积情况下，扩散占优势，这是因为扩散引起的涨落是由分子横穿 V 的边界来回跳跃产生的。这是一种表面效应，相对来讲，这时它比化学反应涨落要大。化学反应涨落来自体反应，它与体积成正比。相反，对于大体积 V，我们发现表面效应微不足道，只有体效应是重要的。

b) 空间-时间相关

由于系统是线性的，所以像 8.3.1 节那样，我们可以利用 3.7.4 节的方法。定义

$$G(r,\ t)=\langle\rho(r,\ t),\ \rho(0,\ 0)\rangle, \tag{8.3.52}$$

因而与(3.7.63)相对应的线性方程为

$$\partial_t G(r,\ t)=D\nabla^2 G(r,\ t)-(k_1-k_2)G(r,\ t) \tag{8.3.53}$$

初条件为

$$G(r,\ 0)=\langle\rho(r),\ \rho(0)\rangle. \tag{8.3.54}$$

它由(8.3.35)给出。把 $G(r,\ t)$ 表示成傅里叶积分

$$\tilde{G}(\boldsymbol{r},\ t)=\int \mathrm{d}^3\boldsymbol{q}\mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{r}}\tilde{G}(\boldsymbol{q},\ t) \tag{8.3.55}$$

则(8.3.53)变成

$$\partial_t\tilde{G}(\boldsymbol{q},\ t)=-(Dq^2+K_1^2-K_2)\tilde{G}(\boldsymbol{q},\ t) \tag{8.3.56}$$

代入初条件(8.3.54)的傅里叶表示 (8.3.47)，上式的解为

$$G(\boldsymbol{r},\ t)=(2\pi)^{-3}\frac{Dq^2+K_1}{Dq^2+K_1-K_2}\exp[-(Dq^2+K_1-K_2)t] \tag{8.3.57}$$

如果愿意，可以用很标准的方法求出它的逆变换。虽然如此，但实际上傅里叶变换相关函数常常根据实际情况写成所希望的、而且通常是容易计算的形式，因为这样对我们有利。

于是我们有

$$\begin{aligned}G(\boldsymbol{r},t)=&\frac{\langle\rho\rangle\exp(-r^2/4Dt-Dt/l_c^2)}{(4\pi Dt)^{3/2}}\\&+\frac{K_2\langle\rho\rangle}{8\pi Dr}\left\{[\exp(-r/l_c)]\operatorname{erfc}\left[\frac{(Dt)^{1/2}}{l_c}-\frac{r}{2(Dt)^{1/2}}\right]\right.\\&\left.+[\exp(r/l_c)]\operatorname{erfc}\left[\frac{(Dt)^{1/2}}{l_c}+\frac{r}{2(Dt)^{1/2}}\right]\right\}\end{aligned} \tag{8.3.58}$$

对于小的 t，我们找到

$$\begin{aligned}G(r,\ t)\to&\langle\rho\rangle\exp\left(\frac{-r^2}{4Dt}\right)\left[(4\pi Dt)^{-3/2}-\frac{4K_2(Dt)^{3/2}}{r^3\sqrt{\pi}}\right]\\&+\frac{K_2\langle\rho\rangle\mathrm{e}^{-r/l_c}}{4\pi Dr}\end{aligned} \tag{8.3.59}$$

而对大的 t，

$$\begin{aligned}G(r,\ t)\sim&\left(\frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}}+\frac{K_2l_c}{4\pi r(\pi D^2t)^{1/2}}\right)\\&\times\langle\rho\rangle\exp(-Dt/l_c^2)\end{aligned} \tag{8.3.60}$$

c) 在不稳定性点附近的行为

当 $K_1\to K_2$ 时，反应趋于不稳定，而当 $K_1=K_2$ 时，不再存在任何平稳解。我们同时可看到

i) 相关长度 $l_c\to\infty$ (8.3.36)；

ii) 当体积 $V \gg l_c^3$ 时，涨落方差将变为无穷大 [(8.3.42)]。然而，当 V 有限而 $l_c \to \infty$ 时，我们到达 $l_c^3 \sim V$ 的点，这时(8.3.42)不再适用。最后当 $l_c^3 \gg V$ 时，体积显得很小，其中的涨落取泊松形式；

iii) 相关时间最好取为

$$\tau_c = l_c^2/D \tag{8.3.61}$$

它是时间很长时空时相关函数的指数中 t 的系数（不考虑扩散部分）。这样，在不稳定性点附近，$\tau_c \to \infty$。于是我们就有一个长程相关的慢涨落图景。这种行为是许多类似情况的共同特征。

8.3.3 两级相变的非线性模型

我们考虑反应

$$\begin{aligned} A + X &\underset{k_4}{\overset{k_2}{\rightleftharpoons}} 2X \\ B + X &\underset{k_3}{\overset{k_1}{\rightleftharpoons}} C \end{aligned} \tag{8.3.62}$$

这两个反应在前面 7.7c、7.3.3a、7.7.4c 中已讨论过。这个模型最早是由薛格尔 (Schlögl) 引进的[8.4]，至今这个模型已用许多别的方法处理过。

上述方程的决定性方程为

$$\partial_t \rho(r, t) = D\nabla^2\rho + \kappa_3 + (\kappa_2 - \kappa_1)\rho - \kappa_4\rho^2 \tag{8.3.63}$$

它的平稳解由下式给出

$$\rho(r) = \rho_s = [\kappa_2 - \kappa_1 + \sqrt{(\kappa_2 - \kappa_1)^2 + 4\kappa_4\kappa_3}]/2\kappa_4 \tag{8.3.64}$$

[κ 的定义与(7.5.29)的定义相同，其中 $\Omega = l^3$]。对于 k_3 很小的情况，该平稳解表示一种相变行为，如图 8.1 所示。

在泊松表示中，恰当的随机微分方程为

$$\begin{aligned} d\eta(r, t) = &[(\mathcal{L}\eta)(r, t) + k_3 + (k_2 - k_1)\eta(r, t) \\ &- k_4\eta^2(r, t)]dt + \sqrt{2}[k_2\eta(r, t) \\ &- k_4\eta^2(r, t)]^{1/2}dW(r, t) \end{aligned} \tag{8.3.65}$$

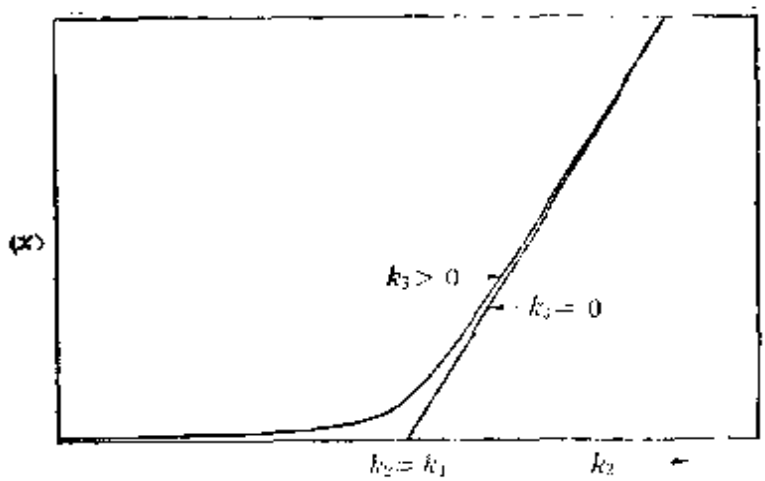


图 8.1　二级相变模型中⟨x⟩随 k_2 的变化曲线

式中

$$\mathrm{d}W(\boldsymbol{r},\ t)\mathrm{d}W(\boldsymbol{r}',\ t)=\delta(\boldsymbol{r}-\boldsymbol{r}')\mathrm{d}t \tag{8.3.66}$$

这里我们没有采用 8.2.1 节中提供的连续极限，而保留了扩散可能具有的非局部性，即

$$(\mathscr{D}\eta)(\boldsymbol{r},\ t)=-\int\mathscr{D}(|\boldsymbol{r}-\boldsymbol{r}'|)\eta(\boldsymbol{r}',\ t)\mathrm{d}^3\boldsymbol{r}' \tag{8.3.67}$$

但是我们只考虑均匀的各向同性扩散。

在这个表示式中，就不存在乘上噪声项 $\mathrm{d}W(\boldsymbol{r},\ t)$ 的小参量以及任何明显的展开式的参量。我们在(8.3.65)中形式地引入参量 λ 如下：

$$\begin{aligned}\mathrm{d}\eta(\boldsymbol{r},\ t)=&[(\mathscr{D}\eta)(\boldsymbol{r},\ t)+k_3+(k_2-k_1)\eta(\boldsymbol{r},\ t)\\&-k_4\eta^2(\boldsymbol{r},\ t)]\mathrm{d}t+\lambda\sqrt{2}\,[k_2\eta(\boldsymbol{r},\ t)\\&-k_4\eta^2(\boldsymbol{r},\ t)]^{1/2}\mathrm{d}W(\boldsymbol{r},\ t)\end{aligned} \tag{8.3.68}$$

用 λ 的幂指数将 $\eta(\boldsymbol{r},\ t)$ 展开到

$$\eta(\boldsymbol{r},\ t)=\eta_0(\boldsymbol{r},\ t)+\lambda\eta_1(\boldsymbol{r},\ t)+\lambda^2\eta_2(\boldsymbol{r},\ t)+\cdots \tag{8.3.69}$$

并在计算结束时令 $\lambda=1$。然而，如果认为所有傅里叶变量的积分截止于 l^{-1}，则实际上该式仍是一个 $(l^3)^{-1}$ 的展式。

将(8.3.69)代入 (8.3.65)，我们得

$$\begin{aligned}\frac{\mathrm{d}\eta_0(\boldsymbol{r},\ t)}{\mathrm{d}t}=&(\mathscr{D}\eta_0)(\boldsymbol{r},\ t)+k_3+(k_2-k_1)\eta_0(\boldsymbol{r},\ t)\\&-k_4\eta_0(\boldsymbol{r},\ t)\end{aligned} \tag{8.3.70}$$

$$d\eta_1(\boldsymbol{r},\ t)=\{(\mathscr{D}\eta_1)(\boldsymbol{r},\ t)+[\kappa_2-\kappa_1-2\kappa_4\eta_0(\boldsymbol{r},\ t)]\eta_1(\boldsymbol{r},\ t)\}dt$$
$$+\sqrt{2}\,[\kappa_2\eta_0(\boldsymbol{r},\ t)-\kappa_4\eta_0(\boldsymbol{r},\ t)]^{1/2}dW(\boldsymbol{r},\ t) \tag{8.3.71}$$

$$d\eta_2(\boldsymbol{r},\ t)=\{(\mathscr{D}\eta_2)(\boldsymbol{r},\ t)+[\kappa_2-\kappa_1-2\kappa_4\eta_0(\boldsymbol{r},\ t)]\eta_2(\boldsymbol{r},\ t)$$
$$-\kappa_4\eta_1(\boldsymbol{r},\ t)\}dt$$
$$+\frac{[\kappa_2-2\kappa_4\eta_0(\boldsymbol{r},\ t)]\eta_1(\boldsymbol{r},\ t)}{\sqrt{2}\,[\kappa_2\eta_0(\boldsymbol{r},\ t)-\kappa_4\eta_0(\boldsymbol{r},\ t)]^{1/2}}\,dW(\boldsymbol{r},\ t) \tag{8.3.72}$$

方程(8.3.70)具有齐次的平稳解 (8.3.64)，以致

$$\kappa_2-\kappa_1-2\kappa_4\eta_0\equiv\kappa=[(\kappa_1-\kappa_2)^2+4\kappa_4\kappa_3]^{1/2} \tag{8.3.73}$$

将(8.3.73)代入 (8.3.70～72)，并作傅里叶变换，得

$$\tilde{\eta}_1(\boldsymbol{q},\ t)=(2\kappa_1\eta_0)^{1/2}\int_0^t[\exp\{-[\mathscr{D}(q^2)+\kappa](t-t')\}]$$
$$\cdot d\tilde{W}(\boldsymbol{q},\ t') \tag{8.3.74}$$

$$\tilde{\eta}_2(\boldsymbol{q},\ t)=-\kappa_4\int d^3q_1\int_0^t dt'[\exp\{-[\mathscr{D}(q^2)+\kappa_1](t-t')\}]$$
$$\cdot\tilde{\eta}_1(\boldsymbol{q}-\boldsymbol{q}_1,\ t')\,\tilde{\eta}_1(\boldsymbol{q}_1,\ t') \tag{8.3.75}$$
$$+\frac{\kappa_1+\kappa}{[2(\kappa_1\eta_0-\kappa_3)]^{1/2}}\int d^3q_1\int_0^t[\exp\{-[\mathscr{D}(q^2)+\kappa]$$
$$(t-t')\}]\tilde{\eta}_1(\boldsymbol{q}-\boldsymbol{q}_1,\ t')d\tilde{W}(\boldsymbol{q}_1,\ t)$$

式中

$$d\tilde{W}(\boldsymbol{q},\ t)=\left(\frac{1}{2\pi}\right)^{3/2}\int\exp(-i\boldsymbol{q}\cdot\boldsymbol{r})dW(\boldsymbol{r},\ t)d^3r \tag{8.3.76}$$

及

$$\tilde{\eta}_1(\boldsymbol{q},\ t)=\left(\frac{1}{2\pi}\right)^{3/2}\int\exp(-i\boldsymbol{q}\cdot\boldsymbol{r})\eta_1(\boldsymbol{r},\ t)d^3r \tag{8.3.77}$$

等，而 $\mathscr{D}(q^2)$ 是 $\mathscr{D}(|\boldsymbol{r}-\boldsymbol{r}'|)$ 的傅里叶变换。我们已忽略了(8.3.74, 75) 中次要的初值项。

对于最低阶的情况，(在平稳态的)平均浓度及相关函数由下式给出，

$$\langle\rho(\boldsymbol{r},\ t)\rangle=\eta_0 \tag{8.3.78}$$

$$\langle\rho(\boldsymbol{r},\ t)\rho(\boldsymbol{r}',\ t)\rangle-\langle\rho(\boldsymbol{r}',\ t)\rangle\langle\rho(\boldsymbol{r}',\ t)\rangle$$
$$=\eta_0\delta(\boldsymbol{r}-\boldsymbol{r}')+\langle\eta_1(\boldsymbol{r},\ t)\eta_1(\boldsymbol{r}',\ t)\rangle \tag{8.3.79}$$

由(8.3.74)得

$$\langle\tilde{\eta}_1(\boldsymbol{q},\ t)\tilde{\eta}_1(\boldsymbol{q}',\ t)\rangle=\frac{k_1\eta_0\delta(\boldsymbol{q}+\boldsymbol{q}')}{[\mathscr{D}(q^2)+k]}$$
$$\cdot[1-\exp\{-2[\mathscr{D}(q^2)+k]t\}] \tag{8.3.80}$$

因此,对稳态相关函数的最低阶贡献为

$$\langle\rho(\boldsymbol{r})\rho(\boldsymbol{r}')\rangle-\langle\rho(\boldsymbol{r})\rangle\langle\rho(\boldsymbol{r}')\rangle=\eta_0\delta(\boldsymbol{r}+\boldsymbol{r}')$$
$$+\frac{\kappa_1\eta_0}{(2\pi)^3}\int d^3q\ \frac{\exp[i\boldsymbol{q}\cdot(\boldsymbol{r}-\boldsymbol{r}')]}{\mathscr{D}(q^2)+k} \tag{8.3.81}$$

若我们假定

$$(\mathscr{D}\eta)(\boldsymbol{r})=DV^2\eta(\boldsymbol{r}) \tag{8.3.82}$$

则

$$\mathscr{D}(q^2)=Dq^2 \tag{8.3.83}$$

因此方程(8.3.81)与前面(8.3.35)中导出的方程具有完全相同的形式,即

$$\langle\rho(\boldsymbol{r}),\ \rho(\boldsymbol{r}')\rangle,=\langle\rho\rangle_0\delta(\boldsymbol{r}-\boldsymbol{r}')$$
$$+\frac{k_1\langle\rho_0\rangle}{4\pi D|\boldsymbol{r}-\boldsymbol{r}'|}\exp(-|\boldsymbol{r}-\boldsymbol{r}'|/l_c) \tag{8.3.84}$$

式中

$$l_c=(D/k)^{1/2} \tag{8.3.85}$$

所有相同的结论都来自有关局部或整体涨落,泊松或非泊松涨落,这是由于这些结论都是式(8.3.84)的结果,与所选参量的特殊值无关。

注意,若 $\kappa_1=\kappa_2$ 时,则当 $\kappa_3\to0$ 时,$\kappa\to0$,因而 $l_c\to\infty$,并且和在那个例子中的情况一样,出现相同的长程相关现象。

高阶相关:发散问题 上面所列的仅是一些最低阶的结果,利用关于相关函数方程的系统大小展式、克雷默斯-莫耶尔方法或因式分解假设,同样可以得到这些结果。

对平均浓度的高一阶的修正是 $\langle\eta_2(\boldsymbol{r},\ t)\rangle$。现在从 (8.3.75) 得

$$\langle\tilde{\eta}_2(\boldsymbol{q},\ t)\rangle = -\int d^3q_1 \int_0^t dt' \exp\{-[\mathscr{D}(q^2)+\kappa](t-t')\}$$
$$\times\langle\eta_1(\boldsymbol{q}-\boldsymbol{q}_1,\ t')\eta_1(\boldsymbol{q}_1,\ t')\rangle \qquad (8.3.86)$$

在稳态，(8.3.86)给出

$$\langle\tilde{\eta}_2(\boldsymbol{q})\rangle = -\kappa_4\kappa_1\eta_0\delta(\boldsymbol{q})\int\frac{d^3q_1}{\mathscr{D}(q^2)+\kappa} \qquad (8.3.87)$$

选取(8.3.83)给出的 $\mathscr{D}$，则(8.3.87)给出

$$\langle\eta_2\rangle = -\frac{\kappa_4\kappa_1\eta_0}{(2\pi)^{3/2}}\int\frac{d^3q_1}{Dq^2+\kappa} \qquad (8.3.88)$$

而这个积分是发散的。问题出在所选取的连续形式中。如果在所有步骤中都取分立的格子，则我们有两个修正：

i) $\int d^3q$ 为对所有允许的 q 求和，q 的最大值为 $|q|\sim 1/l$

ii) $\mathscr{D}(q^2)$ 为 q 和 l 的某种三角函数。如果

$d_{ij}=0$（i 和 j 不相邻）

$=d$（其他情况）

则 $\mathscr{D}(q^2)$ 的形式为

$$\frac{4D}{l^2}\left[\sin\left(\frac{lq_x}{2}\right)+\sin\left(\frac{lq_y}{2}\right)+\sin\left(\frac{lq_z}{2}\right)\right]$$

在这个式中，由于求和有限以及关于

$l|q|\ll 1$

$\mathscr{D}(q^2)\to Dq^2$

的情况，所以不出现发散。

非局部反应 但是，保留连续形式也可以避免发散，只要修改这个方程中的 $\eta^2(r)$ 项。这一项代表了化学反应的碰撞结果，因种种理由它是非局部的。

i) 分子具有有限的大小，所以严格的局部形式不考虑这一点；

ii) 时间的尺度是这样的：在这尺度内系统是马尔可夫的，在这时间内，分子将飞行某个距离；所以反应产物将是前面各个位置碰撞的结果，而且还会在其它地方产生出来。通过代换

$$\eta(r)^2 \to \int d^3r' d^3r'' m(r-r',\ r-r'')\eta(r')\eta(r'') \qquad (8.3.89)$$

我们发现与(8.3.88)不同，有

$$\langle \eta_2 \rangle = -\frac{k_4 k_1 \eta_0}{(2\pi)^{3/2}} \int \frac{d^3q\, \tilde{m}(q, -q)}{D q^2 + k} \qquad (8.3.90)$$

式中 $\tilde{m}(q, q')$ 是 $m(r, r')$ 的傅里叶变换。如果 m 完全是非局部的，则当 q 很大时，$m(q, -q')$ 将衰减，而 $\langle \eta_2 \rangle$ 将为有限。

8.4 局部和整体描述之间的关系

我们已经在 8.3.2 节中看到，在体积元 V 中的涨落方差，对小体积来说是泊松型的，而对充分大的 V 来说，与相应反应中无扩散主方程的涨落方差相接近。这是因为反应引起了涨落，涨落增加了扩散，而由于从一个地方到另一个地方输运的分子较多，所以扩散又影响体积中的涨落，这种影响仅在体积的表面才有。

有一个精确的定理表示了这个事实，即如果扩散足够快，则体积 V 中的分子将到处快速飞行并相互频繁相遇。因此，任何一个分子以同等的可能与任何别的分子相遇。该结果已由阿诺特(Arnold)[8.5]所总结，并已得到非常严格的证实。

实质上，该证明方法有赖第五章提供的绝热消去技巧，并且可以用泊松表示法很容易地加以证明。

8.4.1 非均匀模式的明显绝热消去法

设想我们正在研究单种化学物质的系统，该物质系统按格子模型扩散和反应，并且因而具有泊松表示的福克-普朗克方程，

$$\frac{\partial P}{\partial t} = \left\{ \sum_{i,j} \left[-\frac{\partial}{\partial x_i} D_{ij} x_j \right] + \sum_i \left[\frac{\partial}{\partial x_i} a(x_i) + \frac{\partial^2}{\partial x_i^2} b(x_i) \right] \right\} P \qquad (8.4.1)$$

我们引进 D_{ij} 的本征向量 $f_i(q)$，它满足

$$\sum_j D_{ij} f_j(q) = -D\lambda(q) f_i(q) \qquad (8.4.2)$$

本征函数的这类特性并不很重要，并将依赖于 P_{ij}。对于最简单

的形式，输运仅发生在两个相邻的格子之间，而且反射边界位于总系统的壁上。设系统是一维系统，其长度 $L=vl$（l 是格子的长度）。我们有

$$f_j(q)\propto\cos(qlj) \tag{8.4.3}$$

而反射边界条件要求 q 具有下列形式：

$$q=\frac{n\pi}{Nl}\quad(n=0,\ 1,\ \cdots,\ N) \tag{8.4.4}$$

以及

$$\lambda(q)=4\sin^2\left(\frac{ql}{2}\right)\Big/l^2 \tag{8.4.5}$$

进一步的适当修正必须考虑更多的维数，但是它们的基本结构相同，即

$$\lambda(o)=0$$

$$\lambda(q)>0\quad(q\neq0)\ \text{及} \tag{8.4.6}$$

$$f_i(o)=\text{常数}=N^{-1/2} \tag{8.4.7}$$

这最后的结果表示扩散的时齐平稳态，其归一化常数 $N^{1/2}$ 由态的完备性及正交关系

$$\sum_j f_j(q)f_j(q')=\delta_{q,q'}$$

$$\sum_q f_i(q)f_j(q)=\delta_{i,j} \tag{8.4.8}$$

确定。现在我们引进变量

$$x(q)=\sum_j f_j(q)x_j \tag{8.4.9}$$

我们感兴趣的是正比于 $x(0)$ 的变量，由于

$$x(0)=N^{-1/2}\sum_j x_j \tag{8.4.10}$$

正比于系统内物质的总数，而其他变量都被绝热地消去了。预先考虑适当选取算子 L_1、L_2 及 L_3，我们定义

$$y(q)=x(q)\sqrt{D}\quad(q\neq0) \tag{8.4.11}$$

$$x=x(0)/\sqrt{N}=N^{-1}\sum_j x_j \tag{8.4.12}$$

现在福克-普朗克方程中各项可以写成

$$\sum_{ij}\frac{\partial}{\partial x_i}D_{ij}x_j=-D\sum_q\lambda(q)\frac{\partial}{\partial y(q)}y(q) \tag{8.4.13}$$

定义

$$v_i=\sum_{q\neq 0}f_i(q)y(q) \tag{8.4.14}$$

因而

$$\sum_i\frac{\partial}{\partial x_i}a(x_i)=\frac{1}{N}\sum_i\frac{\partial}{\partial x}a\left(x+\frac{1}{\sqrt{D}}v_i\right)+\sqrt{D}\sum_{q\neq 0}\frac{\partial}{\partial y(q)}\sum_i f_i(q)a\left(x+\frac{1}{\sqrt{D}}v_i\right) \tag{8.4.15}$$

$$\sum_i\frac{\partial^2}{\partial x_i^2}b(x_i)=\frac{1}{N^2}\sum_i\frac{\partial^2}{\partial x^2}b\left(x+\frac{1}{\sqrt{D}}v_i\right)+\frac{1}{N}\sqrt{D}\frac{\partial}{\partial x}\sum_{\substack{i\\ q\neq 0}}\frac{\partial}{\partial y(q)}f_i(q)b\left(x+\frac{1}{\sqrt{D}}v_i\right)+D\sum_{\substack{i\\ q',q\neq 0}}\frac{\partial^2}{\partial y(q)\partial y(q')}f_i(q)f_i(q')b\times\left(x+\frac{1}{\sqrt{D}}v_i\right) \tag{8.4.16}$$

现在我们用$\sqrt{D}$的递减幂来写出上式，为此，定义

$$DL_1=D\sum_{q\neq 0}\left[-\lambda(q)\frac{\partial}{\partial y(q)}y(q)+\frac{\partial^2}{\partial y(q)^2}b(x)\right] \tag{8.4.17}$$

L是上式各项展开式中D的系数(最低阶项被吸收在L_2中)，我们也令

$$L_3=\left\langle\sum_i\left[\frac{1}{N}\frac{\partial}{\partial x}a\left(x+\frac{1}{\sqrt{D}}v_i\right)+\frac{1}{N^2}\frac{\partial^2}{\partial x^2}b\left(x+\frac{1}{\sqrt{D}}v_i\right)\right]\right\rangle \tag{8.4.18}$$

式中平均是对L_1的平稳分布的平均。当D变得很大时，

$$L_3=\left[\frac{\partial}{\partial x}a(x)+\frac{1}{N}\frac{\partial^2}{\partial x^2}b(x)\right]+0\left(\frac{1}{\sqrt{D}}\right) \tag{8.4.19}$$

我们确定L_2到$\sqrt{D}$级，它可以计算到仅包含位于左面的$\frac{\partial}{\partial y(q)}$的一些项。这样，当$D\to\infty$进行变量$y(q)$的绝热消去时，它趋于

最低的量级，将不存在 $PL_2L_1^{-1}L_2$ 的贡献。最后我们得到福克-普朗克方程

$$\frac{\partial P}{\partial t}=\left[\frac{\partial}{\partial x}a(x)+\frac{1}{N}\frac{\partial^2}{\partial x^2}b(x)\right]P \tag{8.4.20}$$

该方程与整体主方程相对应，因为根据(8.4.12)，x 是浓度。

注意，适用条件是

$$D\lambda(1)\gg K \tag{8.4.21}$$

式中 K 代表典型的化学反应速率，或注意 $\lambda(1)\simeq(\pi/Nl^2)$，

$$D\gg\frac{KN^2l^2}{\pi^2} \tag{8.4.22}$$

或

$$\sqrt{\frac{D}{K}}\gg\frac{L}{\pi} \tag{8.4.23}$$

根据扩散的定义，左边部分代表化学反应的时间尺度内的飞行距离的均方根，不等式表示的这个均方根必定比系统的长度大很多。于是，扩散必定能使系统均匀化。

8.5 相空间的主方程

本章的余下部分讨论分子运动论的随机描述，以及引用范坎彭和范德勃洛克(Van den Broek)等人的工作[8.7]。

我们讨论气体系统，其中分子状态用它们的位置 r 和速度 v 来描述。分子到处运动并相互碰撞。当两个分子发生碰撞时，它们的速度因碰撞而改变，但它们的位置不变。然而，在两次碰撞之间，粒子以恒定速度自由运动。

这样，就存在两个过程——碰撞和流动。每一种过程单独存在时很容易处理，但当两种过程结合在一起时就不简单了。

8.5.1 流动的处理方法

我们设想一个由大量粒子组成的系统，粒子的质量为 m，粒子的坐标为 r_n，速度为 v_n，它们不发生碰撞而仅在外界力场 $mA(r)$ 的影响下运动。运动方程为

$$\dot{\boldsymbol{r}}_n = \boldsymbol{v}_n$$
$$\dot{\boldsymbol{v}}_n = \boldsymbol{A}(\boldsymbol{r}_n) \tag{8.5.1}$$

因而相空间密度可以定义为

$$f(\boldsymbol{r},\ \boldsymbol{v},\ t) = \sum_n \delta(\boldsymbol{r} - \boldsymbol{r}_n(t))\delta(\boldsymbol{v} - \boldsymbol{v}_n(t)) \tag{8.5.2}$$

以致

$$\partial_t f(\boldsymbol{r},\ \boldsymbol{v},\ t) = \sum_n [\dot{\boldsymbol{r}}_n \cdot \nabla \delta(\boldsymbol{r} - \boldsymbol{r}_n(t))\delta(\boldsymbol{v} - \boldsymbol{v}_n(t)) + \delta(\boldsymbol{r} - \boldsymbol{r}_n(t))\boldsymbol{v}_n \cdot \nabla_v \delta(\boldsymbol{v} - \boldsymbol{v}_n(t))]$$

利用 δ 函数的性质及运动方程 (8.5.1)，我们得

$$\boxed{[\partial_t + \boldsymbol{v}\cdot\nabla + \boldsymbol{A}(\boldsymbol{r})\ \nabla_v] f(\boldsymbol{r},\ \boldsymbol{v},\ t) = 0} \tag{8.5.3}$$

这是一个相空间密度的确定性的流方程．如果粒子按某种以位置和速度的初概率分布而分布，则 (8.5.3) 仍旧不变，但 $f(\boldsymbol{r}, \boldsymbol{v}, t)$ 是被解释为象(8.5.2)那样的 $f(\boldsymbol{r}, \boldsymbol{v}, t)$ 对初位置和初速度的平均．

方程(8.5.3)是精确的．变量 $f(\boldsymbol{r}, \boldsymbol{v}, t)$ 可看作为随机变量，它的时间演化由该方程给出，而该方程也可看成是零噪声的随机微分方程．相格是体积为 Δ_i 的六维体积元，其中心的相空间坐标为 $(\boldsymbol{r}_i, \boldsymbol{v}_i)$，相格中的粒子数显然为

$$\boldsymbol{X}(\boldsymbol{r}_i,\ \boldsymbol{v}_i) = \int_{\Delta_i} \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v} f(\boldsymbol{r},\ \boldsymbol{v}) \tag{8.5.4}$$

8.5.2 作为生灭过程的流

为了同生灭主方程表示的碰撞相容，若能表示作为格子 Δ_i 中的生灭过程流是很有用的．这对任意大小的格子来说是不可能的．但是在格子大小趋近零的极限情况下，任何流过程的生灭表示是存在的．

为了说明这点，我们考虑一维系统中的密度 $\rho(r, t)$，它遵循流方程

$$\partial_t \rho(r,\ t) = -\kappa \partial_r \rho(r,\ t) \tag{8.5.5}$$

这个确定性方程是关于 x_i 的离散方程的极限，x_i 由下式确定．

$$x_i(t) = \int_{\Delta_i} \mathrm{d}r\rho(r, t) \simeq \lambda\rho(r_i, t) \tag{8.5.6}$$

式中 Δ_i 是 r_i 处的格点周围长为 λ 的区间。因而流方程是下列方程当 $\varkappa \to 0$ 时的极限。

$$\partial_t x_i(t) = \frac{\kappa}{\lambda}[x_{i-1}(t) - x_i(t)] \tag{8.5.7}$$

即

$$\lambda\partial_t\rho(r_i, t) = \frac{\kappa}{\lambda}[\lambda\rho(r_i - \lambda, t) - \lambda\rho(r_i, t)] \tag{8.5.8}$$

当 $\lambda \to 0$ 时，其极限为流方程 (8.5.5)，只要考虑到粒子在单位时间内的转移概率

$$\left.\begin{aligned} t^+(x) &= \kappa x_i/\lambda \\ t^-(x) &= 0 \end{aligned}\right\} \tag{8.5.9}$$

从格子 i 跳到 $i+1$，则可求得上面方程的随机形式。这是 7.5 节中所考察过的形式，那儿的符号(7.5.4, 5) 在这里取为

$$\left.\begin{aligned} &A \to i \\ &N_j^i \to \delta_{i,j} \\ &M_j^i \to \delta_{i+1,j} \\ &r_j^i \to \delta_{j,i+1} - \delta_{j,i} \end{aligned}\right\} \tag{8.5.10}$$

我们考虑克雷默斯-莫耶尔展式，并证明在 $\lambda \to 0$ 的极限下，除了第一级外，所有导数矩均为零。漂移系数由(7.5.32)给出，即

$$\begin{aligned} A_a(x) &= \sum_i (\delta_{a,i+1} - \delta_{a,i})\kappa x_i/\lambda \\ &= \frac{\kappa}{\lambda}(x_{a-1} - x_a) \end{aligned} \tag{8.5.11}$$

而扩散矩阵为

$$B_{a,b}(x) = \frac{\kappa}{\lambda}[(\delta_{a,b} - \delta_{a-1,b})x_{a-1} + (\delta_{a,b} - \delta_{a,b-1})x_a)] \tag{8.5.12}$$

现在我们令 $x_i = \rho(r_i)$，和(8.5.6)一样，并取小 λ 极限；我们得到

$$A_a(x) \to \frac{\kappa}{\lambda}[\lambda\rho(r_a - \lambda, t) - \lambda\rho(r_a, t)] \tag{8.5.13}$$

$$\to -\lambda\kappa\partial_r\rho(r, t) \tag{8.5.14}$$

而 $B_{a,b}(x)$ 的极限值可以用类似的方法得到，但也要利用

$$\delta_{a,b} \to \lambda\delta(r_a - r_b) \tag{8.5.15}$$

$$B_{a,b}(x) \to k\lambda^3 \partial_r \partial_{r'} [\delta(r - r')\rho(r)] \tag{8.5.16}$$

于是，在这极限下，$\rho(r, t)$ 遵循随机微分方程

$$\mathrm{d}\rho(r, t) = -\kappa\partial_r \rho(r, t)\mathrm{d}t + \lambda^{1/2}\mathrm{d}W(r, t) \tag{8.5.17}$$

式中

$$\mathrm{d}W(r, t)\mathrm{d}W(r', t) = \kappa \mathrm{d}t \partial_r \partial_{r'} [\delta(r - r')\rho(r, t)] \tag{8.5.18}$$

我们看到，在 $\lambda \to 0$ 的极限下，(8.5.17) 中的噪声消失，导致预期的确定性的结果。

有趣的问题是这个确定性的极限为什么会发生？它不是该名词在通常意义上的系统大小展式，因为在 $\lambda \to 0$ 的极限下，既不是粒子数 x_i、也不是转移概率变得很大。然而，在 $\lambda \to 0$ 的极限情况下，在 r_i 处单个粒子跳到下一个格子的转移概率确实变成无限大，所以在这个意义上，运动就变得确定性的了。

读者也能经核对发现，这个结果是与空间维数无关的。于是，我们能求得相空间中流的表示，在小格子极限下，它的确变得与流方程(8.5.3)等价。

现在来考虑一个完整的相空间表示公式，它包含流方程的两项。相空间中的格子取成六维盒子，沿位置坐标的边长为 λ，沿速度坐标的边长为 ξ。用一个相格中的总分子数 X 来定义相空间密度，其为

$$f(rq, vq) = X(rq, vq)/\xi^3\lambda^3 \tag{8.5.19}$$

考虑两种跃迁。为简单起见，仅考虑在 x 方向的跃迁，并定义

$$\lambda_x = (\lambda, 0, 0) \tag{8.5.20}$$

类型1：位置空间中的流

$$X(r^i, v^i) \to X(r^i, v^i) - 1$$

或者

$$X(r^i + \lambda_x, v^i) \to X(r^i + \lambda_x, v^i) + 1 \quad (v_x > 0) \tag{8.5.21}$$

或

$$X(r^i - \lambda_x, v^i) \to X(r^i - \lambda_x, v^i) + 1 \quad (v_x < 0)$$

因此我们做好标记 $A\to(i,\ x,\ 1)$：

$$\left.\begin{aligned}
&N_a^{(i,x,1)}=\delta(r^i,\ r^a)\delta(v^i,\ v^a)\\
&M_a^{(i,x,1)}=\delta(r^i-\lambda_x,\ r^a)\delta(v^i,\ v^a)\\
&r_a^{(i,x,1)}=M_a^{(i,x,1)}-N_a^{(i,x,1)}
\end{aligned}\right\}\tag{8.5.22}$$

$$t^+_{(i,x,1)}\ (X)=\lambda^2|v_x|X(r^i,\ v^i)/\lambda^3$$

$$t^-_{(i,x,1)}\ (X)=0\tag{8.5.23}$$

式(8.5.23)这种写法指明，转移概率是格子端面积 (λ_2)、单位体积内的粒子数和速度的 x 分量的乘积，速度的 x 分量近似等于通过这个格子面的流速。

因此，设 $v_x>0$，

$$A_a^{(x,1)}=v_x[X(r^a+\lambda_x,v^a)-X(r^a,\ v^a)]/\lambda\tag{8.5.24}$$

$$=\xi^3\lambda^3v_x[f(r^a+\lambda_x,v^a)-f(r^a,\ v^a)]/\lambda\tag{8.5.25}$$

$$\to\xi^3\lambda^3v_x\frac{\partial}{\partial x}f(r,\ v)\tag{8.5.26}$$

对前面的例子，用类似的方法可以证明，扩散矩阵与 λ 成正比，因此当 $\lambda\to0$ 时，$B_{a,b}$ 为零。

类型 2：速度空间中的流：

定义

$$\xi_x=(\xi,\ 0,\ 0)\tag{8.5.27}$$

因此我们得

$$X(r^i,\ v^i+\xi_x)\to X(r^i,\ v^i+\xi_x)+1\quad(A_x>0)\tag{8.5.28}$$

或

$$X(r^i,\ v^i-\xi_x)\to X(r^i,\ v^i-\xi_x)+1\quad(A_x<0)$$

标记为

$$\left.\begin{aligned}
&A\to(i,\ x,\ 2)\\
&N_a^{(i,x,2)}=\delta(v^i,\ v^a)\delta(r^i,\ r^a)\\
&M_a^{(i,x,2)}=\delta(v^i-\xi_x,v^a)\delta(r^i,\ r^a)\\
&r_a^{(i,x,2)}=[\delta(v^i-\xi_x,v^a)-\delta(v^i,\ v^a)]\delta(r^i,\ r^a)\\
&t^+_{(i,x,2)}\ (X)=\xi^2|A_x(r^i)|X(r^i,\ v^i)/\xi^3\\
&t^-_{(i,x,2)}\ (X)=0
\end{aligned}\right\}\tag{8.5.29}$$

因此，设 $A_x(r^a)>0$，漂移系数为

$$A_a=[X(r^a,\ v^a+\xi_x)A_x(r^a)-X(r^a,\ v^a)A_x(r^a)]/\xi \tag{8.5.30}$$

$$=\xi^3\lambda^3A_x(r^a)[f(r^a,\ v^a+\xi_x)-f(r^a,\ v^a)]/\xi \tag{8.5.31}$$

$$\to\xi^3\lambda^3A_x(r)\frac{\partial}{\partial v_x}f(r,\ v) \tag{8.5.32}$$

用类似的理由再次证明，当 $\xi\to0$ 时扩散系数为零。

把(8.5.26, 32)一起加以考虑，在 λ、$\xi\to0$ 的极限下，我们得到适当的流方程(8.5.3)。

8.5.3 包含碰撞在内的情况——玻耳兹曼主方程

开始我们只考虑速度空间中的粒子，并象以前一样，把速度空间分成各个格子，设 $X(v)$ 为具有速度 v 的分子数目(这里速度 v 当然是离散化了的)。

因此碰撞可以用"化学反应"

$$X(v_i)+X(v_j)\to X(v_k)+X(v_l) \tag{8.5.33}$$

来表示。碰撞用指标 $(i,\ j,\ k,\ l)$ 来标记，我们有(用 7.5 节的符号)

$$\begin{aligned}N_a^{ij,kl}&=\delta_{a,i}+\delta_{a,j}\\ M_a^{ij,kl}&=\delta_{a,k}+\delta_{a,l}\\ r_a^{ij,kl}&=-\delta_{a,i}-\delta_{a,j}+\delta_{a,k}+\delta_{a,l}\end{aligned} \tag{8.5.34}$$

单位时间的跃迁概率取成如下形式

$$\begin{aligned}t^+_{ij,kl}(\boldsymbol{X})&=R(ij,\ kl)X(v_i)X(v_j)\\ t^-_{ij,kl}(\boldsymbol{X})&=0\end{aligned} \tag{8.5.35}$$

存在 5 个碰撞不变量，也就是说，这是一些碰撞过程中的守恒量，它们来自动力学。这些量是：

i) 分子数——式(8.5.33)每边都存在二个分子数；

ii) 动量的三个分量：由于这里的所有分子都是相同的，所以在碰撞过程中意味着

$$v_i+v_j=v_k+v_l \tag{8.5.36}$$

iii) 总能量，这意味着

$$v_i^{\prime} + v_j^{\prime} = v_k^{\prime} + v_l^{\prime} \tag{8.5.37}$$

量 v 和 v^2 是已知的加性不变量，可以证明，与此类似的加性守恒的任何函数都是它们的线性函数(可能带有常数项)[8.8]。

在所有分子碰撞的过程中，我们有时间反演对称性，在这种情况下，对称性意味着

$$R(ij, kl) = R(kl, ij) \tag{8.5.38}$$

最后，因所有粒子的全同性，我们有

$$R(ij, kl) = R(ji, kl) = R(ij, lk) \text{等等。} \tag{8.5.39}$$

现在我们有了各种可能的方法。这些方法是：

i) 试图直接用主方程；

ii) 系统大小展开式：假设相空间中格子的体积 ξ^3 很大，而且可以用克雷默斯-莫耶尔展开写出福克-普朗克方程。由(7.5.32)，我们可以写出具有漂移项的福克-普朗克方程，

$$A_a(X) = \sum_{i,j,k,l} (-\delta_{a,i} - \delta_{a,j} + \delta_{a,k} + \delta_{a,l}) \times X(v_i) X(v_j) R(ij, kl) \tag{8.5.40}$$

并利用所有可供利用的对称性，

$$\text{上式} = 2 \sum_{j,k,l} R(aj, kl)[X(v_k)X(v_l) - X(v_a)X(v_l)] \tag{8.5.41}$$

扩散系数也可以导出：

$$B_{ab}(X) = \sum_{i,j,k,l} r_a^{i,kl} r_b^{ij,kl} R(ij, kl) X(v_i) X(v_j)$$

再利用全部可供利用的对称性，

$$\begin{aligned} B_{ab}(X) = {} & 2\delta_{a,b} \sum_{j,k,l} R(aj, kl)[X(v_a)X(v_j) + X(v_k)X(v_l)] \\ & + 2\sum_{i,j} R(ij, ab)[X(v_i)X(v_j) + X(v_a)X(v_b)] \\ & - 4\sum_{j,l} R(aj, bl)[X(v_a)X(v_j) \\ & + X(v_b)X(v_l)] \end{aligned} \tag{8.5.42}$$

这意味着随机微分方程为

$$\mathrm{d}X(\boldsymbol{v}_a) = \{2\sum_{j,k,l} R(aj,\ kl)[X(\boldsymbol{v}_k)X(\boldsymbol{v}_l) - X(\boldsymbol{v}_a)X(\boldsymbol{v}_j)]\}\mathrm{d}t + \mathrm{d}W(\boldsymbol{v}_a,\ t) \tag{8.5.43}$$

式中

$$\mathrm{d}W(\boldsymbol{v}_a,\ t)\mathrm{d}W(\boldsymbol{v}_a,\ t) = \mathrm{d}tB_{ab}(\boldsymbol{X}) \tag{8.5.44}$$

忽略随机项，我们重新得到离散形式的关于 $X(\boldsymbol{v}_a)$ 的玻耳兹曼方程。如经常所做的那样，这个克雷默斯-莫耶尔方程仅在小噪声极限下才适用，它与系统大小展式相等价，这时动量空间格子的体积大小为 ξ^3。

iii) 泊松表示：玻耳兹曼主方程是泊松表示处理方法的理想候选者，与通常一样，利用变量 $\alpha(\boldsymbol{v}_a)$，我们可以从结果(7.7.6～9)获得具有漂移项

$$A_a(\alpha) = 2\sum_{j,k,l} R(aj,\ kl)[\alpha(\boldsymbol{v}_k)\alpha(\boldsymbol{v}_l) - \alpha(\boldsymbol{v}_a)\alpha(\boldsymbol{v}_j)] \tag{8.5.45}$$

和扩散矩阵

$$\begin{aligned} B_{ab}(\alpha) = {} & 2\delta_{ab}\sum_{i,j,l} R(ij,\ kl)[\alpha(\boldsymbol{v}_i)\alpha(\boldsymbol{v}_j) - \alpha(\boldsymbol{v}_a)\alpha(\boldsymbol{v}_l)] \\ & + 2\sum_{i,j} R(ij,\ ab)[\alpha(\boldsymbol{v}_i)\alpha(\boldsymbol{v}_j) - \alpha(\boldsymbol{v}_a)\alpha(\boldsymbol{v}_b)] \\ & - \delta_{a,b}\sum_{j,l} R(aj,\ bl)[\alpha(\boldsymbol{v}_a)\alpha(\boldsymbol{v}_j) - \alpha(\boldsymbol{v}_b)\alpha(\boldsymbol{v}_l)] \end{aligned} \tag{8.5.46}$$

的福克-普朗克方程以及相应的随机微分方程

$$\begin{aligned} \mathrm{d}\alpha(\boldsymbol{v}_a) = {} & 2\sum_{j,k,l} R(aj,\ kl)[\alpha(\boldsymbol{v}_k)\alpha(\boldsymbol{v}_l) - \alpha(\boldsymbol{v}_a)\alpha(\boldsymbol{v}_j)] \\ & + \mathrm{d}\tilde{W}(\boldsymbol{v}_a,\ t) \end{aligned} \tag{8.5.47}$$

式中

$$\mathrm{d}\tilde{W}(\boldsymbol{v}_a,\ t)\mathrm{d}\tilde{W}(\boldsymbol{v}_b,\ t) = \mathrm{d}tB_{ab}(\alpha) \tag{8.5.48}$$

正如前面已强调过的那样，这个泊松表示的随机微分方程，精确等价于假设的玻耳兹曼主方程。和克雷默斯-莫耶尔方法或系统大小展式不同，它对于速度空间格子 ξ 的所有大小都是适用的。

iv) 玻耳兹曼主方程的平稳解：前面我们已经写出了 $t_{ij,kl}(\boldsymbol{x})$ 为零的玻耳兹曼主方程，但另一方面，我们可以写出

$$t^-_{ij,kl}(X) = t^+_{ij,kl}(X) \tag{8.5.49}$$

以及将所有含R的项适当地除以2，因为其中每项都计算了两次。

细致平衡条件(7.5.18)明显满足。虽然我们曾令$t^-(ij, kl)=0$，但是其逆变换实际上由$t^+(kl, ij)$给出。因此，

$$k^+_{ij,kl} = k^-_{ij,kl} = R(ij, kl) \tag{8.5.50}$$

它证明时间反演对称(8.5.38)被满足。

在这些条件下，平稳解具有均值，它满足

$$a(v_i)a(v_j) = a(v_k)a(v_l) \tag{8.5.51}$$

这意味着$\log[a(v_i)]$是加性守恒的，它必然是不变量(8.5.36, 37)的函数。因此

$$a(v) = \exp[-(v-U)^2/mKT] \tag{8.5.52}$$

式中m是分子的质量，U和kT是参量，当然U是用分子的平均速度来表示的，而后者为绝对温度与玻耳兹曼常数之积。

因而平稳分布是一个由(8.5.52)给出均值的多变量泊松分布，分子涨落与不同的速度无关。

8.5.4 碰撞和流同时存在

在把流的处理方法和碰撞的处理方法结合起来时，存在着基本的困难。困难的原因在于流的随机处理方法要求具有无限小的格子，而玻耳兹曼主方程认为用有限大小的格子比较好。这意味着，严格地写出系统的精确的随机方程几乎是不可能的，除非用后面我们将简短地作出说明的泊松表示才行。

但是，当我们假定相空间格子为有限大小$\lambda^3\xi^3$时，用公式写出多变量的相空间主方程是直接了当的。我们简单地包含所有适用的输运，也即，这些输运在位置空间、速度空间产生流及碰撞。于是，所得到的主方程包含了(8.5.22、23、29)及修改式(8.5.34、35)中确定的可能有的输运。然而，这里我们所具有的每个格子内的碰撞是由单位时间内的转移概率定义的

$$t^{+}_{i,j,kl}(\boldsymbol{X}) = \delta(\boldsymbol{r}_i, \boldsymbol{r}_j)\delta(\boldsymbol{r}_k, \boldsymbol{r}_l)\delta(\boldsymbol{r}_i, \boldsymbol{r}_k)R(ij, kl)X_iX_j \tag{8.5.53}$$

对有限大小的 $\lambda^3\xi^3$，将存在来自有限格子大小的随机影响，正如8.5.2节所指出的那样，当来自流的输运是纯粹确定性的时候，在 λ 和 ξ 取极限的情况下，上述影响则消失。

写出所得到的主方程是相当麻烦的，所以我们将不明显写出它。对它所做的大多数工作是涉及系统大小展式或与其等价的克雷默斯-莫耶尔近似。有效的精密极限取决于 $R(ij, kl)$ 对系统大小的依赖关系。在这种情况下，系统大小是六维相空间体积 $\lambda^3\xi^3$，为了作出密度的确定性方程，我们定义密度为

$$f(\boldsymbol{r}_i, \boldsymbol{v}_i) = X(\boldsymbol{r}_i, \boldsymbol{v}_i)/\lambda^3\xi^3 \tag{8.5.54}$$

它与格子的大小无关，和(8.5.53)的定义相同，$R(ij, kl)$ 必然与 $(\lambda^3\xi^3)^4$ 成正比，即

$$R(ij, kl) = \bar{R}(ij, kl)(\lambda^3\xi^3)^4 \tag{8.5.55}$$

该式的意义可解释为：$\bar{R}(ij, kl)$ 是以每个参量构成的单位相空间体积元中的平均碰撞率。

考虑到守恒定律(8.5.36、37)后，我们可以写出

$$R(ij, kl) = 8\sigma[(\boldsymbol{v}_i-\boldsymbol{v}_j)^2, (\boldsymbol{v}_i-\boldsymbol{v}_j)\cdot(\boldsymbol{v}_k-\boldsymbol{v}_i)] \times \delta(\boldsymbol{v}_i^2+\boldsymbol{v}_j^2-\boldsymbol{v}_k^2-\boldsymbol{v}_l^2)\delta(\boldsymbol{v}_i+\boldsymbol{v}_j-\boldsymbol{v}_k-\boldsymbol{v}_l) \tag{8.5.56}$$

[而我们已假定 σ 是唯一的标量函数，实际上 σ 仅仅是 $(\boldsymbol{v}_i-\boldsymbol{v}_j)^2$ 和 $(\boldsymbol{v}_i-\boldsymbol{v}_j)(\boldsymbol{v}_k-\boldsymbol{v}_i)$ 的函数，这是散射理论的结果，它来自关于伽利略变换群的不变性，即旋转不变性，以及这样的事实，即物理定律不依赖于非加速参考系的特殊选择。为了使涨落项中的表示式具有简单性，我们决定利用标积来保持这种函数关系。]

a) 克雷默斯-莫耶尔展式

现在我们按下式用积分代替求和

$$(\xi^3\lambda^3)\sum_j \to \int \mathrm{d}^3\boldsymbol{r}_j\mathrm{d}\boldsymbol{v}_j \tag{8.5.57}$$

并作变量变换

$$\boldsymbol{v}_1 = \boldsymbol{v}_i$$

$$\boldsymbol{v}_2 = \frac{1}{2}(\boldsymbol{p}+\boldsymbol{q}) \tag{8.5.58}$$

$$\boldsymbol{v}_i = \frac{1}{2}(\boldsymbol{p}-\boldsymbol{q})$$

对碰撞项作了某些变换后，[从 (8.5.3, 41)]得到下面形式的决定性方程

$$\begin{aligned} \mathrm{d}f(\boldsymbol{r},\boldsymbol{v}) = \Big\{ & -\boldsymbol{v}\cdot\nabla f(\boldsymbol{r},\boldsymbol{v}) - \boldsymbol{A}\cdot\nabla_{\boldsymbol{v}} f(\boldsymbol{r},\boldsymbol{v}) \\ & + \int \mathrm{d}^3\boldsymbol{v}_1 \int \frac{\mathrm{d}^3\boldsymbol{q}}{|\boldsymbol{q}|}\,\delta(|\boldsymbol{q}| - |\boldsymbol{v}-\boldsymbol{v}_1|)\,\sigma \\ & \times [\boldsymbol{q}^2,\ \boldsymbol{q}\cdot(\boldsymbol{v}-\boldsymbol{v}_1)] \\ & \times \Big[f\Big[\boldsymbol{r},\ \frac{1}{2}(\boldsymbol{v}+\boldsymbol{v}_1-\boldsymbol{q})\Big] f\Big[\boldsymbol{r},\ \frac{1}{2}(\boldsymbol{v}+\boldsymbol{v}_1+\boldsymbol{q})\Big] \\ & - f(\boldsymbol{r},\boldsymbol{v}) f(\boldsymbol{r},\boldsymbol{v}_1) \Big] \Big\} \mathrm{d}t \end{aligned} \tag{8.5.59}$$

这个来自克雷默斯-莫耶尔展式的随机微分方程，加上随机项 $\mathrm{d}W(\boldsymbol{r},\boldsymbol{v},t)$ 后就得到了，此随机项满足下式

$$\begin{aligned} \mathrm{d}W(\boldsymbol{r},\boldsymbol{v},t)\mathrm{d}W(\boldsymbol{r}',\boldsymbol{v}',t) = {} & \delta(\boldsymbol{r}-\boldsymbol{r}')\mathrm{d}t \\ & \times \Big\{ \delta(\boldsymbol{v}-\boldsymbol{v}') \int \mathrm{d}^3\boldsymbol{v}_1 \frac{\mathrm{d}^3\boldsymbol{q}}{|\boldsymbol{q}|}\,\delta(|\boldsymbol{q}| - |\boldsymbol{v}-\boldsymbol{v}_1|) \\ & \times \sigma[\boldsymbol{q}^2, \boldsymbol{q}\cdot(\boldsymbol{v}-\boldsymbol{v}_1)] f\Big[\boldsymbol{r},\ \frac{1}{2}(\boldsymbol{v}+\boldsymbol{v}_1-\boldsymbol{q})\Big] \\ & \times f\Big[\boldsymbol{r},\ \frac{1}{2}(\boldsymbol{v}+\boldsymbol{v}_1+\boldsymbol{q})\Big] + f[(\boldsymbol{r},\boldsymbol{v}) f(\boldsymbol{r},\boldsymbol{v}_1)] \\ & - 2\int \mathrm{d}^3\boldsymbol{k}\,\delta[(\boldsymbol{v}_a-\boldsymbol{v}_b)\cdot\boldsymbol{k}]\,\sigma\Big[(\boldsymbol{v}_a-\boldsymbol{v}_b)^2 + \frac{1}{4}\boldsymbol{k}^2, \\ & -(\boldsymbol{v}_a-\boldsymbol{v}_b)^2 + \frac{1}{4}\boldsymbol{k}^2\Big] \\ & \times [f(\boldsymbol{v}_a) f(\boldsymbol{v}_b-\boldsymbol{k}) + f(\boldsymbol{v}_b) f(\boldsymbol{v}_a-\boldsymbol{k})] \Big\} \end{aligned} \tag{8.5.60}$$

利用这种克雷默斯-莫耶尔方法，范、德、勃洛克等人[8.7]已可能应用查普曼-恩斯喔克(Chapman-Enskog)展式，并从它求得了涨落动力学。

关于克雷默斯-莫耶尔展式的适用性，它取决于格子大的程度；而如生灭过程中的流模型的适用性，则取决于格子大小的程度，所以这种方法的适用性自然是值得怀疑的。但是，由于查普曼-爱恩斯喔克方法很可能等价于由玻耳兹曼碰撞项控制的变量的绝热消去法，而绝热消去法的结果对这种算子的精密形式多半不是非常敏感的。因此，克雷默斯-莫耶尔近似确实是充分正确的，即使对非常小的格子。

b) 泊松表示

泊松表示的随机微分方程可以类似地从 (8.5.45, 46) 得出，我们使用

$$\phi(\boldsymbol{r}, \boldsymbol{v}) = a(\boldsymbol{r}, \boldsymbol{v})/(\lambda^3\xi^3) \tag{8.5.61}$$

定义的符号 $\phi(\boldsymbol{r}, \boldsymbol{v})$，就会发现

$$\begin{aligned} \mathrm{d}\phi(\boldsymbol{r}, \boldsymbol{v}) = \mathrm{d}t\Big\{ & -\boldsymbol{v}\cdot\nabla\phi(\boldsymbol{r}, \boldsymbol{v}) - \boldsymbol{A}\cdot\nabla_v\phi(\boldsymbol{r}, \boldsymbol{v}) \\ & + \int \mathrm{d}^3v_1 \frac{\mathrm{d}^3q}{|\boldsymbol{q}|}\,\delta(|\boldsymbol{q}| - |\boldsymbol{v}-\boldsymbol{v}_1|) \\ & \times\sigma[\boldsymbol{q}^2, \boldsymbol{q}\cdot(\boldsymbol{v}-\boldsymbol{v}_1)]\Big[f\left(\boldsymbol{r}, \frac{1}{2}(\boldsymbol{v}+\boldsymbol{v}_1-\boldsymbol{q})\right) \\ & \times f\left(\boldsymbol{r}, \frac{1}{2}(\boldsymbol{v}+\boldsymbol{v}_1+\boldsymbol{q})\right) \\ & - f(\boldsymbol{r}, \boldsymbol{v})f(\boldsymbol{r}, \boldsymbol{v}_1)\Big]\Big\} + \mathrm{d}W(\boldsymbol{r}, \boldsymbol{v}, t) \end{aligned} \tag{8.5.62}$$

式中

$$\begin{aligned} \mathrm{d}W(\boldsymbol{r}, \boldsymbol{v}, t)\mathrm{d}W(\boldsymbol{r}', \boldsymbol{v}', t) = {} & \delta(\boldsymbol{r}-\boldsymbol{r}')\mathrm{d}t \\ & \times\Big(\delta(\boldsymbol{v}-\boldsymbol{v}')\int \mathrm{d}^3v_1 \frac{\mathrm{d}^3q}{|\boldsymbol{q}|}\,\delta(|\boldsymbol{q}| - |\boldsymbol{v}-\boldsymbol{r}_1|) \\ & \times\sigma[\boldsymbol{q}^2, \boldsymbol{q}\cdot(\boldsymbol{v}-\boldsymbol{v}_1)]\,\{\phi[\boldsymbol{r}, \frac{1}{2}(\boldsymbol{v}+\boldsymbol{v}_1-\boldsymbol{q})] \\ & \times\phi[\boldsymbol{r}, \frac{1}{2}(\boldsymbol{v}+\boldsymbol{v}_1+\boldsymbol{q})] - \phi(\boldsymbol{r}, \boldsymbol{v})\phi(\boldsymbol{r}, \boldsymbol{v}_1)\} \\ & + \int \frac{\mathrm{d}^3q}{|\boldsymbol{q}|}\,\delta(|\boldsymbol{q}| - |\boldsymbol{v}-\boldsymbol{v}_1|)\sigma[\boldsymbol{q}^2, \boldsymbol{q}\cdot(\boldsymbol{v}-\boldsymbol{v}_1)] \end{aligned}$$

$$\times\{\phi[r,\ \frac{1}{2}(r+r'-q)]\,\phi[r,\ \frac{1}{2}(v+v'+q)]$$

$$-\phi(r,v)\phi(r,v')\}\Big) \tag{8.5.63}$$

(8.5.63)中的那些项相应于 (8.5.46) 中的前二项。在$\xi^3\lambda^3\to 0$的极限下，最后一项给出零分布。

我们总是强调，这种泊松表示的随机微分方程，精确地等价于加有流项的玻耳兹曼主方程的小格子其大小取极限的情况。方程 (8.5.62, 63) 先前没有被写出过，至今也没有被应用过。应用查普曼-爱恩斯渥克的或类似的技巧，我们很可能导出精确的涨落动力学方程。

9. 双稳定性、亚稳定性和逃逸问题

本章致力于可能至少存在两个稳定态的系统的渐近行为和某些密切有关问题的研究。这类系统具有巨大的实际意义，例如计算机中的开关和存储系统就是具有这种性质的系统。某些可以分解为各个部分的分子也是这种系统，最近又研究了大量显示各种各样有关性质的电子学系统、化学和物理系统。

令人关心的问题是：

i) 相互关联的各种态是如何稳定的？

ii) 一个系统自动地从一个态转换到另一个态需要多长时间？

iii) 输运是怎样进行的？即相应的态空间中的路径是怎样的？

iv) 系统从不稳定态发生的弛豫怎样进行？

对于一维扩散过程，这些问题都比较容易回答。但是除了很低的维数之外，推广到高维系统仅是最近的事了。当推广到无限多个变量时，把我们带入液气相变和类似的相变的领域之中，这时系统可以处于两个相中的一个，并在空间中任意地分布。这是个以随机观点的描述尚未就绪的领域，这里不处理这个问题。

本章基本分三部分：单变量双稳态扩散过程，一步生灭双稳态系统和多变量扩散过程。最后所得结果，定性上都是相似的，但要得到定量的精确性必须作巨大的努力。

9.1 在双井势中的扩散（单变量）

我们再次考虑5.2.7节的模型，其中粒子的概率密度$P(x, t)$

遵循福克-普朗克方程

$$\partial_t p(x,\ t) = \partial_x [U'(x) p(x,\ t)] + D\partial_x^2 p(x,\ t) \qquad (9.1.1)$$

$U(x)$的形状如图 9.1 所示。在 a 和 c 点有两个极小值，而在两个极小值之间有一个局部极大。平稳分布为

$$p_s(x) = \mathcal{N}\exp[-U(x)/D] \qquad (9.1.2)$$

正是它显示了双稳态性。相应于 a、c 和 b 是二个极大和一个中央极小，如图 9.1 所示，于是系统最可能在 a 和 c 处找到。

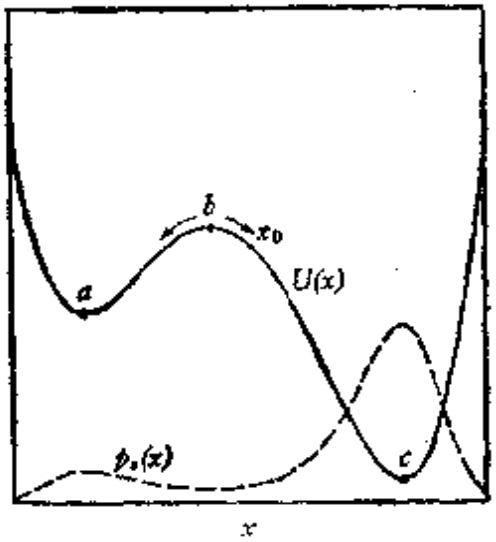


图 9.1　$p_s(x)$和双井势 $U(x)$的曲线

9.1.1　$D=0$ 时的行为

在这种情况下，$x(t)$服从微分方程

$$\frac{dx}{dt} = -U'(x),\ x(0) = x_0 \qquad (9.1.3)$$

由于

$$\frac{dU(x)}{dt} = U'(x)\ \frac{dx}{dt} = -[U'(x)]^2 < 0$$

所以 $x(t)$总是以这种方式运动：极小化 $U(x)$，并只停留在 $U'(x)$为零的地方。因此，其运动决定于 x_0 是否大于或小于 b。粒子分别停在 c 或 a，运动沿图中箭头方向进行。

一旦粒子处在 a 或 c，就继续停在那儿。如果开始时它精确

地位于 b，则它也将呆在那儿，虽然稍有扰动就趋向 a 或 c 运动。因此 b 是不稳定的平稳点，而 a 和 c 是稳定点。这里不存在关于 a 和 c 的相对稳定性问题。

9.1.2 D 非常小时的行为

随着噪声的加入，情况发生了变化。平稳态可能是如下的渐近近似态。设$U(x)$处处是充分光滑的，我们可以写成

$$U(x)\simeq U(a)+\frac{1}{2}U''(a)(x-a)^2 \quad |x-a|\text{是小量}$$

$$\simeq U(c)+\frac{1}{2}U''(c)(x-c)^2 \quad |x-c|\text{是小量} \tag{9.1.4}$$

如果D很小，则我们可以得到近似

$$p_s(x)\simeq \mathcal{N}\exp\left[-U(a)/D-\frac{1}{2}U''(a)(x-a)^2/D\right]$$

$$|x-a|\text{是小量}$$

$$\simeq \mathcal{N}\exp\left[-U(c)/D-\frac{1}{2}U''(c)(x-c)^2/D\right]$$

$$|x-c|\text{是小量}$$

$$\simeq 0 \quad (\text{其他情况}) \tag{9.1.5}$$

以致

$$\mathcal{N}^{-1}\simeq e^{-U(a)/D}\sqrt{2\pi D/U''(a)}+e^{-U(c)/D}\sqrt{2\pi D/U''(c)} \tag{9.1.6}$$

如图所示，设

$$U(a)>U(c) \tag{9.1.7}$$

则对于足够小的 D，(9.1.6) 中第二项比第一项大得多而占压倒优势，并且 $\mathcal{N}^{-1}$ 可以单独由第二项近似得出。将它代入(9.1.5)，就得到

$$P_s(x)=\sqrt{\frac{U''(c)}{2\pi D}}\exp\left[-\frac{1}{2}U''(c)(x-c)^2/D\right]$$

$$|x-c|\sim\sqrt{D}$$

$$\simeq 0 \quad (\text{其他情况}) \tag{9.1.8}$$

这意味着在D很小的范围内，使$U(x)$取绝对极小值的决定性平稳态，是在随机平稳态的意义下更为稳定的状态，$p_s(x)$除了在它的最紧接的近邻外，到处都非常小。

当然，这个结果与前面的不同，前面的结果说明每个态都是同样稳定的。区别在于时间，而实际上我们将证明，如果开始时令系统处在x_0，并考虑对任何有限的t，$p(x,t)$中的$D\to 0$时的极限，则决定性行为将被随机地重新产生。6.3节的方法证明了这一点，只要关于决定性方程的展式是适用的。方程(6.3.10)表明，只要$U'(x_0)$不为零，或者在任何D为有限的条件下$U'(x_0)$与D^0有相同的阶(这里，D代替6.3节中的ε^2)，则情况就是这样。如果x_0不在a、c或b的$D^{1/2}$阶宽度领域内，则这就是正确的。这意味着在a和c的情况下，涨落盛行起来，而运动由随机微分方程在a或c附近的线性化近似给出。在b附近，线性化随机微分方程是不稳定的。所以粒子遵循奥恩斯坦-乌伦贝克过程运动，直到它离开$x=b$的邻近区域，在这阶段$\sqrt{D}$的渐近展式占了优势。

但是，对于$t\to\infty$的情况，渐近展式不再适用，或者，换句话说，小噪声微扰理论的$t\to\infty$极限复现不了平稳态的$D\to 0$极限。

能够出现的过程是越过中心势垒的逃逸过程，噪声$\mathrm{d}W(t)$可以促使粒子爬上b处的势垒而达到另一边。噪声包含$\exp(-常数/D)$的几倍的量级。因为当$D\to 0$时它们趋于0的速率比D的任何幂次都快，所以它们对以D的幂展开的渐近式不起作用。

9.1.3 出口时间

这个问题已在5.2.7c节研究过了。最初粒子在a附近，后来到达中点b，其所用时间为

$$T(a\to b)=\pi\left[|U''(b)|U''(a)\right]^{-1/2}\exp\{[U(b)-U(a)]/D\} \tag{9.1.9}$$

(只要粒子到达b右边势井中一点所需时间的一半)。对于$D\to 0$的情况，上式按指数律变大，因此系统到达平稳态所需时间也将按指

数律增大，在这种时间尺度基础上，用 $D^{1/2}$ 幂表示相应的随机微分方程解的时间演化，不可能被期望是适用的。

9.1.4 分裂概率

设想把粒子放在 x_0 处；则它在 c 之前达到 a 或 a 之前达到 c 的概率是什么？这可以跟通过区间的特殊一端的出口问题联系起来，这个问题已在 5.2.8 节中研究过。我们在 $x=a$ 及 $x=c$ 处放置吸收壁，并利用那一节的结果得到最早到达 a 或 c 的“分裂概率”π_a 和 π_c。它们是(注意扩散系数为 D，因而与 x 无关)：

$$\pi_a(x_0)=\left[\int_{x_0}^{c}\mathrm{d}x p_s(x)^{-1}\right]\Big/\left[\int_a^c \mathrm{d}x p_s(x)^{-1}\right]$$

$$\pi_c(x_0)=\left[\int_a^{x_0}\mathrm{d}x p_s(x)^{-1}\right]\Big/\left[\int_a^c \mathrm{d}x p_s(x)^{-1}\right] \qquad (9.1.10)$$

这个分裂概率 π_a 和 π_c 可以更一般地看成仅是这样的概率，即开始时处在 x_0 的粒子将落入左边或右边的势井中，因为已到达 a 的粒子将留在该井的那一边。其停留时间与 b 的平均出口时间有相同的量级。

现在我们讨论当 $D\to 0$ 时两种可能的渐近形式。

a) x_0 为距 b 的有限距离

我们先求值

$$\int_a^c \mathrm{d}x p_s(x)^{-1}=\int_a^c \mathrm{d}x \mathcal{N}^{-1}\exp[U(x)/D] \qquad (9.1.11)$$

这是由位于 $x\sim b$ 内的行为所支配的。设

$$U(x)\simeq U(b)-\frac{1}{2}|U''(b)|/(b-x)^2 \qquad (9.1.12)$$

则渐近值能正确地求得。当 $D\to 0$ 时，在 $x=a$、c 处，极限分散成 $\pm\infty$，我们并且发现

$$\int_a^c \mathrm{d}x p_s(x)^{-1}\sim \mathcal{N}^{-1}\sqrt{\frac{2\pi D}{|U''(b)|}}\exp[U(b)/D] \qquad (9.1.13)$$

现在设 $x_0<b$，则通过代换

$$y=U(x) \qquad (9.1.14)$$

及其反变换 $x=W(y)$，我们可以求出 $\int_a^{x_0} dx\,p(x)^{-1}$ 的值，它的渐近式为

$$\mathcal{N}^{-1}\int_{-\infty}^{U(x_0)} e^{y/D}W'(y)dy \sim \mathcal{N}^{-1}De^{U(x_0)/D}W'[U(x_0)]$$

$$= \mathcal{N}^{-1}\frac{De^{U(x_0)/D}}{U'(x_0)} \tag{9.1.15}$$

因此，

$$\pi_c \sim \frac{1}{U'(x_0)}\sqrt{\frac{|U''(b)|}{2\pi}D}\exp\left[\frac{U(x_0)-U(b)}{D}\right] \tag{9.1.16}$$

以及

$$\pi_a = 1-\pi_c \tag{9.1.17}$$

这里我们看到分裂概率只与 x_0 和 b 有关，因此，在这极限下达到 c 点的概率完全由在 b 点跃过势垒的概率给出。a 和 c 两点实际上有无限的距离。

b) x_0 为距 b 无穷小距离

设

$$x_0 = b - y_0\sqrt{D} \tag{9.1.18}$$

在这种情况下，我们可在两个积分中求出近似值。定义[9.1]

$$\mathrm{erf}(x) = \sqrt{\frac{\pi}{2}}\int_0^x dt e^{-t^2} \tag{9.1.19}$$

我们发现

$$\pi_c = 1-\pi_a \sim \frac{1}{2}\left(1-\mathrm{erf}\left[y_0\sqrt{|U''(b)|}\right]\right) \tag{9.1.20}$$

$$= \frac{1}{2}\left\{1-\mathrm{erf}\left[(b-x_0)\sqrt{\frac{|U''(b)|}{D}}\right]\right\} \tag{9.1.21}$$

如果用 $U(x)$ 在整个区域上的二次近似式(9.1.12)来替代它，结果得到方程(9.1.21)。

c) 两个区域的比较

两个区域给出不同的结果，而且我们发现随机微分方程的简单线性化[它是用 $U(x)$ 的二次近似代替 $U(x)$ 得到的]，仅在 D 大的限制下、以及在极大值 b 的附近 $\sqrt{D}$ 量级的范围内，才给出正确

的结果。

9.1.5 来自非稳态的衰变

一个位于势场中一点的粒子，到达这个或那个势井中所需要的平均时间是一个可以用实验加以测量的对象。如果我们利用(9.1.1)作为过程，那末从 b 到达 a 或 c 的平均时间，可以利用5.2.8节的公式精确地算出。从 b 到 a 的平均时间是下式的解：

$$-U'(x)\partial_x[\pi_a(x)T(a,x)]+D\partial_x^2[\pi_a(x)T(a,x)]\equiv-\pi_a(x) \tag{9.1.22}$$

其边界条件为

$$\pi_a(a)T(a,a)=\pi_a(c)T(a,c)=0 \tag{9.1.23}$$

而 $\pi_a(x)$ 由(9.1.10)给出。

通过直接积分来求解(9.1.22)是非常直截了当的，但它也是相当麻烦的。解题技巧完全与(5.2.158)所用的相同，而且结果也相似。在(5.2.158)中，我们没有区分左、右出口，即使这样，情况也是非常复杂的。

但是为了记录在案，我们写下

$$T(a,x)=\frac{\left[\begin{array}{l}\pi_a(x)\int_x^c dx' p_s(x')^{-1}\int_a^{x'}\pi_a(z)p_s(z)dz\\ -\pi_a(x)\int_a^x dx' p_s(x')^{-1}\int_c^{x'}\pi_a(z)p_s(z)dz\end{array}\right]}{D\pi_a(x)} \tag{9.1.24}$$

在本式中，我们认为 $\pi_a(x)$ 由(9.1.10)给出，$p_s(z)$ 由(9.1.2)给出。也可以看到，即使对于最简单的可能情况，即

$$U(x)=-\frac{1}{2}kx^2 \tag{9.1.25}$$

该式几乎不可能被理解。渐近处理方法大概是需要的。幸好，当 $p_s(z)$ 在 a 和 c 处是尖锐的峰，而在 b 处是尖锐的极小。这种情况，问题实质上简约为从位于 b 的势壁反射后到 a 或 c 的弛豫问题。

为了看清这点，请注意

1) 关于 $\pi_a(x)$(9.1.10)的显式解意味着

$$\pi_a(x)\simeq 1 \qquad (x<b)$$
$$\simeq \frac{1}{2} \qquad (x \quad b) \tag{9.1.26}$$
$$\simeq 0 \qquad (x>b)$$

并且从 1 到 0 的输运发生在量级为 $\sqrt{D}$ 的距离，即 $p_s(x)^{-1}$ 中峰的宽度上。

ii) 在被积函数 $\pi_a(z)p_s(z)$ 的积分中，我们区分二种情况 $x'>b$：在这种情况下，我们估计有

$$\int_a^{x'} \pi_a(z)p_s(z)\mathrm{d}z \simeq n_0/2 \tag{9.1.29}$$

式中 $n_0=\int_{-\infty}^{b} p_s(z)\mathrm{d}z$，并代表处在左边的势井中的概率。

然而，当 $x'<a$ 时，我们仍可以把 $\pi_a(z)$ 近似为 1，所以我们得

$$\int_0^{x'} \pi_a(z)p_s(z)\ \mathrm{d}z \simeq \int_a^{x'} p_s(z)\,\mathrm{d}z = \frac{n_0}{2} - \int_{x'}^{b} p_s(z)\mathrm{d}z \tag{9.1.30}$$

把这些估计值代入(9.1.24)，我们得

$$T(a,\ b)\simeq D^{-1}\int_a^b \mathrm{d}x' p_s(x)^{-1}\int_{x'}^{b} p_s(z)\mathrm{d}z \tag{9.1.31}$$

当 b 处存在反射势垒时，它是从 b 到 a 的精确平均出口时间。类似地

$$T(c,\ b)\simeq D^{-1}\int_b^c \mathrm{d}x' p_s(x)^{-1}\int_b^{x'} p_s(z)\mathrm{d}z \tag{9.1.32}$$

9.2 在每个势井中的粒子数平衡

假定开始时粒子系统处在左边势井中某个位置 x_i，以致

$$p(x,\ 0)=\delta(x-x_i) \tag{9.2.1}$$

如果 D 很小，从 x 到达中心处的时间很长，而同首次出口时间相比却要小几倍，故将不受到来自 c 处所存在的势井的影响。实际上我们可以假设在 b 处存在一个反射势垒。

在左边势井中的运动将用小噪声展式来描述，因此典型的弛豫时间将与决定性运动的弛豫时间相同。让我们作近似

$$U\ (x)\simeq U(a)+\frac{1}{2}U''(a)x^2$$

因而现在系统是奥恩斯坦-乌伦贝克过程，而典型的时间标度是 $[U''(a)]^{-1}$ 阶次的。

这样，我们期望两种时间标度描述。在短期内，系统弛豫到开始所在的势井中的准稳态，然后在较长的时间标度上，它可以跳越过 b 处的极大值，而趋近于长时间的双峰平稳分布。

9.2.1 克雷默斯方法

1940 年克雷默斯[9.2] 从分子输运的观点讨论了逃逸问题。他引入了所谓的克雷默斯方程(5.3.6 a 节)，在该方程中他考虑在双井势影响下的运动。在阻尼很大的情况下，他证明了在(6.4.18)中建立的、相应的斯莫路乔夫斯基方程形式是可以利用的。因此逃逸问题基本上简化为目前正在讨论中的问题。

克雷默斯方法已被人们重新发现、并多次被表述[9.3]。这里我们将介绍一种形式，使它的精确适用范围更为合理而清楚。

利用图 9.1 中的符号，定义

$$M\ (x,\ t) = \int_{-\infty}^{x} \mathrm{d}x'\, p(x',\ t) \tag{9.2.2}$$

$$N_a(t) = 1 - N_c(t) = M(b,\ t) \tag{9.2.3}$$

及

$$N_0(t) = (c-a)\, p(x_0,\ t)$$

进而由下式定义相应的平衡态量

$$n_a = 1 - n_c = \int_{-\infty}^{a} p_s(x')\mathrm{d}x' \tag{9.2.4}$$

$$n_0 = (c-a)\, p_s(x_0)$$

根据福克-普朗克方程(9.1.1)及(9.1.2) 给出的 $p_s(x)$的形式，我们可以写出

$$\mathrm{d}_t M(x,\ t) = D p_s(x) \partial_x [p(x,\ t)/p_s(x)] \tag{9.2.5}$$

它可以被积分，并给出

$$\mathrm{d}_t \int_a^{x_0} \mathrm{d}x M(x,t)/p_s(x) = D[p(x_0,t)/p_s(x_0) - p(a,\ t)/p_s(a)] \tag{9.2.6}$$

这个方程没有包含任何近似。我们想要引进某种类型的、在长时间情况下适用的近似。

为了介绍这种方法的实质，我们不得不引入并非我们希望的某些不太严格的论证。由于我们相信每个势井中的弛豫是相当快的，所以我们期望在每个势井中的分子的分布(当 $D\to 0$时，时间的数量级为有限)趋向于与平稳态分布相同的形状，但是两个峰的相对权重是不同的。这可用以下公式表示

$$\begin{aligned} p(x,\ t) &= p_a(x)N_a(t)/n_a \qquad x<b \\ &= p_c(x)N_c(t)/n_c \qquad x>b \end{aligned} \tag{9.2.7}$$

这个式子除了在 b 附近$\sqrt{D}$量级的范围以外，将会精确到含 D 的最低级。

如果我们将这些式子代入 (9.2.6)，则得

$$\boxed{\begin{aligned} \kappa(x_0)\dot{N}_a(t) &= D[N_0(t)/n_0 - N_a(t)/n_a] \\ \mu(x_0)\dot{N}_c(t) &= D[N_0(t)/n_0 - N_c(t)/n_c] \end{aligned}} \tag{9.2.8}$$

以及

$$\begin{aligned} \kappa(x_0) &= \int_a^{x_0} p_a(x)^{-1}[1-\psi(x)]\,dx \\ \mu(x_0) &= \int_{x_0}^{c} p_c(x)^{-1}[1-\psi(x)]\,dx \end{aligned} \tag{9.2.9}$$

和

$$\begin{aligned} \psi(x) &= n_a^{-1}\int_x^b p_a(z)dz \qquad x<b \\ &= n_c^{-1}\int_b^x p_c(z)dz \qquad x>b \end{aligned} \tag{9.2.10}$$

注意，如果 x 与 a 或 c 相差有限，则如同直接从 $p_a(x)$ 的显式可看出那样，当 $D\to 0$ 时 $\psi(x)$ 指数地衰减。因此，由于 x 在 (9.2.9) 的两个积分中满足整个积分区的条件，我们可以令

$$\psi(x)=0$$

以及应用

$$\boxed{\begin{aligned}\kappa(x_0) &= \int_a^{x_0} p_s(x)^{-1}dx \\ \mu(x_0) &= \int_{x_0}^{c} p_s(x)^{-1}\,dx\end{aligned}} \tag{9.2.11}$$

a) 三态解释

方程(9.2.8)对应于可能写成如下类型的化学反应过程

$$X_a \rightleftarrows X_0 \rightleftarrows X_c \tag{9.2.12}$$

只是该反应中不存在 x_0 的数目 N_0 的方程。注意到 $N_a+N_c=1$，我们发现

$$V_0(t)=n_0[\mu(x_0)N_a(t)+\kappa(x_0)N_c(t)]/[\kappa(x_0)+\mu(x_0)] \tag{9.2.13}$$

这是与从(9.2.8)中绝热消去 $N_0(t)$ 所得到的方程及 $N_0(t)$ 的进一步的方程

$$\begin{aligned}\dot{N}_0(t) &= D\{N_a(t)/[n_a\kappa(x_0)]+N_c(t)/[n_c\mu(x_0)]\} \\ &\quad -N_0(t)\{[n_0\kappa(x_0)]^{-1}+[n_0\mu(x_0)]^{-1}\}\end{aligned} \tag{9.2.14}$$

相同的。

由于

$$n_0=p_s(x_0)(c-a)=\mathcal{N}\exp[-U(x_0)/D]\,(c-a) \tag{9.2.15}$$

我们看到，$D\to 0$ 的极限相当于 $n_0\to 0$，因此(9.2.14)中速率常数乘$N_0(t)$变得指数式地增大。所以，绝热消去法是适用的。

这三态解释本质上是爱林(Eyring)[9.4]所提出的化学反应过渡态理论。

b) 中间态的消去

把两个方程相加而从(9.2.8)中消去 $N_0(t)$之后，我们得

$$\dot{N}_a(t)=-\dot{N}_c(t)=r_aN_a(t)+\gamma_cN_c(t) \tag{9.2.16}$$

其中

$$\gamma_a=D\left[n_a\int_a^c dx\,p_s(x)^{-1}\right]^{-1},\ \gamma_c=D\left[n_c\int_a^c dx\,p_s(x)^{-1}\right]^{-1} \tag{9.2.17}$$

式中 γ_a 及 γ_c 与 x_0 无关。于是，x_0 的精确选择并不影响中心峰的弛豫。

由于 $N_a+N_c=1$，弛豫时间常数 τ_r 为

$$\tau_r^{-1}=\gamma_a+\gamma_c=\frac{D}{n_a n_c\int_a^c \mathrm{d}x\, p_s(x)^{-1}} \tag{9.2.18}$$

c) 单位时间的逃逸概率：最初在 a 附近向 x_0 运动的一个粒子，其单位时间的逃逸概率，就是在 x_0 处存在一吸收势垒的条件下 $N_a(t)$ 的衰减率。这意味着在 (9.2.8) 式中令 $N(t)=0$ [但应注意，$p_s(x)$ 由(9.1.2)定义]。同理我们得

$$\dot{N}_a(t)=-DN_a(t)/n_a\kappa(x_0) \tag{9.2.19}$$

以致平均首次通过时间由下式给出

$$\tau_a=n_c D^{-1}\int_a^{x_0}\mathrm{d}x\, p_s(x)^{-1} \tag{9.2.20}$$

这个结果实质上就是用更严格的推理在(5.2.166)中所得的结果。

d) 弛豫时间对峰值粒子数的依赖关系

式(9.2.18)初看起来好象是使弛豫时间与 n_a 及 $n_c=1-n_a$ 有联系的一个简单公式。我们可以认为所有其他因子都与 n_a 和 $\tau_r\propto n_a(1-n_a)$ 无关。然而，更仔细的计算表明，情况并非如此。如果我们使用渐近计算(9.1.13)，就会发现

$$\tau_r=\frac{n_a n_c}{\mathcal{N}D}\sqrt{\frac{2\pi D}{|U''(b)|}}\exp[U(b)/D] \tag{9.2.21}$$

同理，$\mathcal{N}$ 也可通过采用来自每个峰的贡献而作渐近计算。我们得

$$\mathcal{N}^{-1}=\sqrt{2\pi D}\{[U''(a)]^{-1/2}\exp[-U(a)/D]+[U''(c)]^{-1/2}\exp[-U(c)/D]\} \tag{9.2.22}$$

并根据 n_a 和 n_c 的定义(9.2.4)，类似地得

$$n_a/n_c=\sqrt{U''(c)/U''(a)}\exp\left[\frac{U(c)-U(a)}{D}\right] \tag{9.2.23}$$

经某些代数运算后，我们可以重新把(9.2.21)写成

$$\tau_r=2\pi H(b;\ a,\ c)[n_a n_c]^{1/2} \tag{9.2.24}$$

式中 $H(b;\ a,\ c)$ 是一个函数，它依赖于 $U(b)$ 与 $U(a)$ 和 $U(b)$ 的平

均值之相对大小，明确地讲就是

$$H(b;a,c)=[U''(b)^{-1/2}U''(a)^{-1/4}U''(c)^{-1/4}]\times\exp\left[\frac{2U(b)-U(a)-U(c)}{2D}\right] \tag{9.2.25}$$

9.2.2 例子：糜蛋白酶的可逆质变

糜蛋白酶是一种蛋白质，对它加压至几千个大气压，它可以转变成一种性质已变化的形式。这已为霍莱(Hawley) [9.5] 所证实。据估计，如果所加压力充分大，则分子可能突然断裂。

这种过程的某种虽非真实但是简单的模型由下面的方程给出：

$$\mathrm{d}x=\frac{-U'(x)}{\gamma}\,\mathrm{d}t+\sqrt{\frac{2KT}{\gamma}}\,\mathrm{d}W(t) \tag{9.2.26}$$

式中 x 是分子体积，$U(x)$ 是每个分子的吉布斯自由能，KT/γ 起着 D 的作用。这里 γ 是摩擦系数。它是由绝热消去法引起的。就象用于导出 6.4 节中的斯莫洛乔夫斯基方程那样，因此平稳分布为

$$p_s(x)=\mathcal{N}\exp[-U(x)/KT] \tag{9.2.27}$$

和统计力学所要求的相同。

压力变化的明显影响包含在下式中

$$U(x)=U_0(x)+x\delta p \tag{9.2.28}$$

式中 δp 是相对于 n_a 和 n_c 相等的那个态的压力差，项 $x\delta p$ 改变了相对稳定性，并且等价于反抗压力 δp 所做的功。按(9.2.23)这就要求 $U_0(x)$满足

$$\sqrt{U_0''(a)}\exp[U_0(a)/KT]=\sqrt{U_0''(c)}\exp[U_0(c)/KT] \tag{9.2.29}$$

$U(x)$ 的极大和极小与 $U_0(x)$ 的极大和极小稍有不同。如果我们假设 $U_0(x)$ 的高阶微商可以忽略，则 $U(x)$ 的极大和极小位于 $U'(x)=0$ 的点，并由 $a+\delta a$，$b+\delta b$ 和 $c+\delta c$ 点之值给出。式中

$$\left.\begin{aligned}\delta a &= -\delta p/U_0''(a) = \beta_a \delta p\\ \delta b &= \delta p/|U_0''(b)| = \beta_b \delta p\\ \delta c &= -\delta p/U_0''(c) = \beta_c \delta p\end{aligned}\right\} \tag{9.2.30}$$

这样我们把 β_a 和 β_c 标记为态 a 和 c 的压缩系数 $\frac{\partial x}{\partial p}$，按稳定性的要求，它是一个负值。量 β_b 是过渡态的某种不可压缩性，由于这个态是不稳定的，所以 β_b 是个正值。

这些极小的 $U(a)$、$U(b)$ 和 $U(c)$ 的值为

$$\left.\begin{aligned}U(a+\delta a) &= U_0(a+\delta a) + (a+\delta a)\delta p\\ &= U_0(a) + a\delta p + \frac{1}{2}\beta_a(\delta p)^2\\ U(b+\delta b) &= U_0(b) + b\delta p + \frac{1}{2}\beta_b(\delta p)^2\\ U(c+\delta c) &= U_0(c) + c\delta p + \frac{1}{2}\beta_c(\delta p)^2\end{aligned}\right\} \tag{9.2.31}$$

而根据(9.2.23)，我们得

$$\boxed{\frac{n_a}{n_c} = \exp\left[\frac{-(a-c)}{kT}\delta p - \frac{(\beta_a-\beta_c)}{2kT}(\delta p)^2\right]} \tag{9.2.32}$$

这个公式是精确的，它由热力学推理得出，并且与实验数据符合得很好。

到平稳分布的弛豫时间 τ_r 也是可以测量的。我们利用(9.2.24, 25)来计算它，我们发现

$$\boxed{\begin{aligned}\tau_r(0) &= \gamma^{-1}\pi\beta_b^{1/2}\exp\left[\frac{U(b)}{kT}\right]\\ \tau_r(\delta p) &= (n_a n_c)^{1/2}\tau_r(0)\exp\left[-\frac{a+c-2b}{2kT}\delta p\right.\\ &\quad \left.+\frac{\delta p^2}{2kT}\left(\beta_b - \frac{1}{2}\beta_a - \frac{1}{2}\beta_c\right)\right]\end{aligned}} \tag{9.2.33}$$

注意，原则上两个态的体积 a 和 c 以及它们的压缩系数 β_a 和 β_c 都是可直接测量的量。过渡态的特性 b，$U(b)$ 和 β_b 留作自由参量，它直接由寿命的测量确定。当然，平方项只适用于 δp 充分小的情

况，而为了应用，可能有必要使用更复杂的方法。

在图 9.2 中，$\tau_r(\delta p)$ 和 n_a/n_c 按一组可能参量值作了图，并与(9.2.33)所作的计算相同。

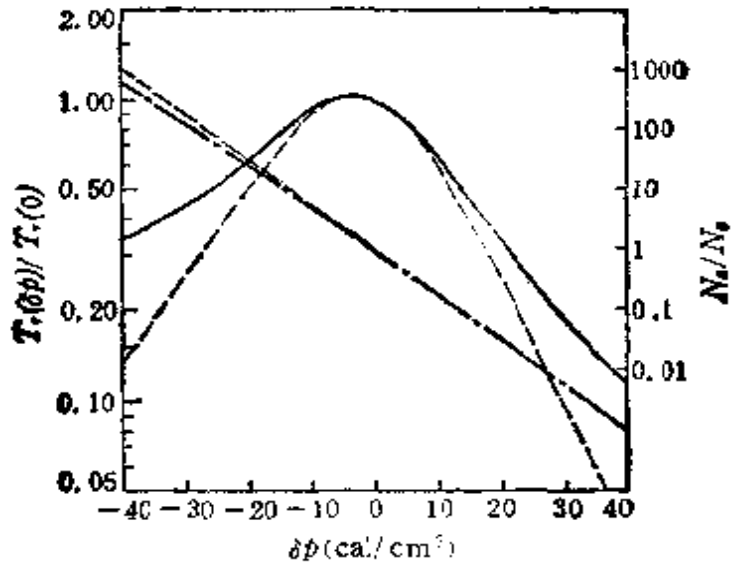


图 9.2　按(9.2.33)式描述弛豫速率 $\tau_r(\delta p)$ 和糜蛋白酶的天然和变性形式的浓度比率。虚线省略了压缩系数修正

注意，当系统接近使自然和变性形式具有相等浓度的点时，平衡时间达到极大值。但是，曲线的不对称可能是由造成势井不对称的因素引起的。事实上，霍莱注意到了在这个范围的测量受到"仪器稳定性和研究者耐心"的限制。

最后，为了作比较，图中给出了零压缩系数的曲线。在图的两翼，两曲线相差如此之大，这清楚表明，平方修正法对 $\tau_r(\delta p)$ 曲线是不适用的。然而，对 n_a/n_c 曲线修正几乎为零。包含更多变量的一种现实处理方法，是保留了这种描述的定性性质，但是它同样允许 $\beta_b<0$ 的可能性，并如在[9.9]中指出的那样，这在单变量的情况下是不可能的。

9.2.3　生灭主方程(单变量)的双稳定性

由单步生灭主方程给出的双稳态系统的定性行为，几乎与福克-普朗克方程给出的相同。

我们来考虑转移概率为 $t^+(x)$ 和 $t^-(x)$ 的单步过程，故主方程可以写成

$$\frac{\mathrm{d}p(x,\ t)}{\mathrm{d}t} = J(x+1,\ t) - J(x,\ t) \tag{9.2.34}$$

且

$$J(x,\ t) = t^-(x)p(x,\ t) - t^+(x-1)p(x-1,\ t) \tag{9.2.35}$$

设平稳分布在 a、c 具有极大值，而在 b 具有极小值，用 9.2.1 节的类似方法，定义

$$M\ (x,\ t) = \sum_{z=0}^{x-1} p(z,\ t) \tag{9.2.36}$$

$$N_a(t) = 1 - N_c(t) = M(b,\ t) \tag{9.2.37}$$

并且如果 x_0 是邻近 b 的一点，则

$$N_0(t) = p(x_0,\ t) \tag{9.2.38}$$

相应的平稳态量为

$$n_a = 1 - n_c = \sum_{z=0}^{b-1} p_s(z) \tag{9.2.39}$$

$$n_0 = p_s(x_0) \tag{9.2.40}$$

现在对(9.2.34)从 0 到 $x-1$ 求和，就得到

$$\frac{\mathrm{d}M(x,\ t)}{\mathrm{d}t} = J(x,\ t) \tag{9.2.41}$$

[由于 $J(0,\ t) = 0$]。为了导出(9.2.41)中的"积分因子"，现在利用一个事实，即处在一步过程的平稳态解是从细致平衡方程

$$t^-(x)p_s(x) = t^+(x-1)p_s(x-1) \tag{9.2.42}$$

得到的，即若定义

$$\beta(x,\ t) = p(x,\ t)/p_s(x) \tag{9.2.43}$$

则(9.2.41)可写成

$$\frac{\mathrm{d}M(x,\ t)}{\mathrm{d}t} = p_s(x)t^-(x)[\beta(x,\ t) - \beta(x-1,\ t)] \tag{9.2.44}$$

以致

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{z=a+1}^{x_0}\left[\frac{M(z,\ t)}{p_s(z)t^-(z)}\right] = \beta(x_0,\ t) - \beta(a,\ t)$$

$$= \frac{p(x_0, t)}{p_s(x_0)} - \frac{p(a, t)}{p_s(a)} \qquad (9.2.45)$$

现在方程(9.2.45)的形式与相应的福克-普朗克方程(9.2.6)的形式几乎严格相同。它取决于解是通过细致平衡方法(9.2.42)得到的。

我们作与克雷默斯相同的假设，即假设只有两个峰之间的驰豫是有关的，并写成

$$\begin{aligned} p(x, t) &= p_s(x) N_a(t)/n_a \qquad x < b \\ &= p_s(x) N_c(t)/n_c \qquad x > b \end{aligned} \qquad (9.2.46)$$

所得到的驰豫方程完全相似于(9.2.8)中的方程：

$$\begin{aligned} \kappa(x_0)\dot{N}_a(t) &= N_0(t)/n_0 - N_a(t)/n_a \\ \mu(x_0)\dot{N}_c(t) &= N_0(t)/n_0 - N_c(t)/n_c \end{aligned} \qquad (9.2.47)$$

式中

$$\begin{aligned} \kappa(x_0) &= \sum_{x=a+1}^{x_0} [p_s(x) t^-(x)]^{-1} [1 - \psi(x)] \\ \mu(x_0) &= \sum_{x=x_0+1}^{c} [p_s(x) t^-(x)]^{-1} [1 - \psi(x)] \end{aligned} \qquad (9.2.48)$$

而

$$\begin{aligned} \psi(z) &= n_a^{-1} \sum_{y=z}^{b} p_s(y) \qquad z < b \\ &= n_a^{-1} \sum_{y=b+1}^{z+1} p_s(y) \qquad z > b \end{aligned} \qquad (9.2.49)$$

唯一重要的区别在于，在 $\kappa(x_0)$ 和 $\mu(x_0)$ 的定义中，D 出现在(9.2.8)的右边，而这里它被 $t^-(z)^{-1}$ 代替。

可以作出所有相同的近似，唯一的困难是重新阐明 $D \to 0$ 这个极限。这里，该极限必定与大数极限相对应，在大数极限中当 x 变化 ± 1 时，所有函数将光滑地发生变化。这恰巧是系统大小展式的极限，在这个展式中，福克-普朗克的描述总是可以使用的。

我们将不再详细地深入下去，只是提一下：精确的平均出口时间是由7.4节的方法得到的。把5.2.8节的方法用于这个系统，我们发现，开始处在 x_0 的系统达到 a 和 c 的分裂概率为

$$\pi_a = \left\{ \sum_{z=x_0+1}^{b} [p_s(z)t^-(z)]^{-1} \right\} \Big/ \left\{ \sum_{a+1=z}^{a} [p_s(z)t^-(z)]^{-1} \right\}$$

$$\pi_c = \left\{ \sum_{z=a+1}^{x_0} [p_s(z)t^-(z)]^{-1} \right\} \Big/ \left\{ \sum_{z=a+1}^{c} [p_s(z)t^-(z)]^{-1} \right\} \tag{9.2.50}$$

这样，对于所有实际的考虑，我们还不如利用福克-普朗克描述的模型。完全搞清楚基本的机理是什么，这是很罕见的，所以任何写出的方程可能不会比经训练的猜测更多，在这种场合，最简单的就是最适用的。

9.3 多变量系统中的双稳定性

当我们涉及多变量系统时，就存在各种各样的可能性，如果系统用主方程来描述，则各种可能的跃迁和态空间是如此丰富，以致我们很难想象从何着手。但是，由于我们在9.2.4节已看到，主方程描述与福克-普朗克描述并不是很不同的，所以我们限于讨论这些问题似乎是合理的，也已证明这些问题是非常复杂的。

这些问题的启发式的处理方法，主要是由物理学家兰格(Langer)、兰道(Landauer)及斯旺森(Swanson)[9.6]所发展起来的，现在由舒斯(Schuss)及马特科夫斯基(Matkowsky)及其他人[9.7]的数学处理而处在严格化的过程中。第一个严格处理的是由文切尔(Ventsel)和费雷德林(Freidlin)[9.8]给出的，但是，这似乎没有引起从事应用工作者的注意，这是由于这种严格只是用于进一步证实长期以来猜测到的估计，而不是给出精确的渐近表示式，如新近的处理方法所做的那样。

这里我们将考虑由福克-普朗克方程在 l 维空间中描述的系统，福克-普朗克方程写成下列形式是方便的：

$$\partial_t p = \nabla \cdot [-\boldsymbol{v}(\boldsymbol{x})p + \varepsilon \underset{\sim}{\boldsymbol{D}}(\boldsymbol{x}) \cdot \nabla p] \tag{9.3.1}$$

该方程的平稳解称为 $p_s(x)$，并且把它看成在下列这样的条件下是已知的。当然，用6.3.3节的方法可以在小的 ε 极限下作出渐近估计。

9.3.1 出口点的分布

这里我们将只处理方程 9.3.1 简化后的情况，在其中 $D(x)$ 是恒等变换，

$$D(x)=1 \tag{9.3.2}$$

这个等式的确隐藏着一些特点，它们可以从强烈变化的 D，但几乎没有大的变化的结果中得到。

我们假设，系统限制在边界为 s 的区域 R 内，而且速度场 $v(x)$ 是向内指向平稳点 a 的。现在的问题是如何对边界上从它们逃离区域 R 的点 b 的分布作渐近估计。我们利用关于 $\pi(b, x)$ 的(5.4.49)(从点 x 出发后，s 上逃逸点的分布)，在这种情况下(5.4.49)取下面这种形式

$$v(x)\cdot\nabla\pi(b, x)+\varepsilon\nabla^2\pi(b, x)=0 \tag{9.3.3}$$

具有边界条件

$$\pi(b, u)=\delta_s(b-u) \qquad (u\in S) \tag{9.3.4}$$

根据马特科夫斯基和舒斯的方法 [9.7]，构造出了适合 $\varepsilon\to 0$ 的渐近解。

a) 接近 $x=u$ 及在 R 内部的解

首先我们构造一个适合 R 内部的解，对 $\varepsilon\to 0$ 的情况，我们有

$$v(x)\cdot\nabla\pi(b, x)=0 \tag{9.3.5}$$

它意味着 $\pi(b, x)$ 沿 $v(x)$ 的流线是常数。因为简单地说，沿着这些线 $\pi(b, x)$ 为零。由于我们假定所有的流线都通过 a，所以对内部所有的 x 我们得

$$\pi(b, x)=\pi(b, a) \tag{9.3.6}$$

然而这种论证因 $v(a)=0$ 这个事实而无效，并因此 (9.3.5) 不再是合适的近似。

因此，我们考虑(9.3.3)在离原子 $\sqrt{\varepsilon}$ 量级距离内的解。为了有助于做到这点，我们引进新的坐标 (z, y_r)，它们是这样选取的，z 度量距 a 的距离，而 y_r 是一组测量相对于 a 的取向的 $l-1$ 个切向变量。

更精确地讲，要这样选取 $z(x)$ 和 $y_r(x)$，使得

$$\begin{aligned} &v(x)\cdot\nabla z(x) = -z(x)\\ &v(x)\cdot\nabla y_r(x) = 0\\ &z(a) = 0 \end{aligned} \tag{9.3.7}$$

这组方程中第一个方程中的负号是考虑到这样的事实：假定 a 是稳定的，以致 $v(x)$ 指向 a，这样 $z(x)$ 随 x 距 a 越远而越增加。

因此，对任何函数 f，我们发现

$$\nabla f = \nabla z(x)\frac{\partial f}{\partial z} + \sum_r \nabla y_r(x)\frac{\partial f}{\partial y_r} \tag{9.3.8}$$

因而

$$v(x)\cdot\nabla\pi = -z\frac{\partial\pi}{\partial z} \tag{9.3.9}$$

以及

$$\begin{aligned} \nabla^2\pi = {}& \nabla z(x)\cdot\nabla z(x)\frac{\partial^2\pi}{\partial z^2} + 2\sum_r \nabla z(x)\cdot\nabla y_r(x)\frac{\partial^2\pi}{\partial z\partial y_r}\\ &+ \sum_{r,s}\nabla y_r(x)\cdot\nabla y_s(x)\frac{\partial^2\pi}{\partial y_r\partial y_s} + \nabla^2 z(x)\frac{\partial\pi}{\partial z}\\ &+ \sum_r \nabla^2 y_r(x)\frac{\partial\pi}{\partial y_r} \end{aligned} \tag{9.3.10}$$

现在我们通过变换为经尺度调整(或延伸)过的变量 ξ 来渐近地求解 π。变量 ξ 定义为

$$z = \xi\sqrt{\varepsilon} \tag{9.3.11}$$

将(9.3.8～11)代入 (9.3.3)，我们发现对 ε 的最低阶来说，

$$-\xi\frac{\partial\pi}{\partial\xi} + H\frac{\partial^2\pi}{\partial\xi^2} = 0 \tag{9.3.12}$$

式中

$$H = \nabla z(a)\cdot\nabla z(a)$$

现在解出这个方程，得到

$$\pi(b, x) = C_1\int_0^{z/\sqrt{\varepsilon}} d\xi\exp(\xi^2/2H) + \pi(b, a) \tag{9.3.13}$$

因为 H 是正的量，所以我们只能得到 π 的这样的解，即若 $C_1 = 0$，则沿着 $x \neq a$ 的流线 $\pi(b, x)$ 是恒定的。因此，对 R 内的所有 x，

$$\pi(b,\ x) = \pi(b,\ a) \tag{9.3.14}$$

[注意，如果 $v(x)$ 指向其他方向，即若是不稳定的，则我们略去(9.3.9)中的负号，并发现 $\pi(b,\ x)$ 由 $H \to -H$ 时的 (9.3.3) 给出，因此，在 a 的 $\sqrt{\varepsilon}$ 阶的距离内，$\pi(b,\ x)$ 改变其值。]

b) 边界 S 附近的解

着眼于以后的应用，我们来考虑稍微更一般些的方程

$$v(x)\cdot\nabla f(x) + \varepsilon\nabla^2 f(x) = 0$$

的解，其中

$$f(u) = g(u) \qquad (u \in S) \tag{9.3.15}$$

边界值问题(9.3.4)是上式的特殊情况。存在两种情况

i) 在 R 内的 S 上或除 $x = a$ 为稳定点以外的其他任何地方，$\nu\cdot v(x) \neq 0$；显然，在任何渐近方法中，边界条件(9.3.15) 与常数解并不相容。因此，在边界上必产生快速变化。

靠近 S 上的 u 点，我们写出

$$v(u)\cdot\nabla f(x) + \varepsilon\nabla^2 f(x) = 0 \tag{9.3.16}$$

近 S 处引入 $\nu(u)$，即在 u 处指向 S(外向)的法线是方便的，并且定义变量 ρ 为：

$$x = u - \varepsilon\rho\nu(u) \tag{9.3.17}$$

而其它变量 y_r 与 S 平行。

此外，对于 ε 的最低阶，(9.3.16)(在靠近 S 上的 u 点) 简化为

$$[\nu\cdot v\ (u)]\frac{\partial f}{\partial \rho} + H(u)\frac{\partial^2 f}{\partial \rho^2} = 0 \tag{9.3.18}$$

其中

$$H(u) = v^2 = 1$$

因此解为

$$f(x) = g(u) + C_1(u)\{1 - \exp[-\nu\cdot v(u)\rho]\} \tag{9.3.19}$$

当 $\rho \to \infty$ 时，我们向 R 内部的有限距离靠近，于是

$$f(x) \longrightarrow g(u) + C_1(u) = C_0 \tag{9.3.20}$$

根据(a)中的分析，得

$$C_1(u) = C_0 - g(u) \tag{9.3.22}$$

现在我们必须确定 C_0，实际上它是要寻求的主要量。

这可以由格林定理得到，因为 $p_s(x)$ 是前向福克-普朗克方程的通常的平稳解，所以我们知道，它可以写成

$$p_s(x) = \exp\left\{-\frac{1}{\varepsilon}[\phi(x) + O(\varepsilon)]\right\} \tag{9.3.23}$$

它和6.3.3节中所表明的相同。取(9.3.16)，用 $p_s(x)$ 相乘，并在整个 k 上对它积分。利用 $p_s(x)$ 满足前向福克-普朗克方程的事实，这个积分可以化成面积分

$$0 = \int_R dx p_s(x)[v(x)\cdot\nabla f(x) + \varepsilon\nabla^2 f(x)] \tag{9.3.24}$$

$$= \int_S ds\{p_s(x)\nu\cdot v(x) f(x) + \varepsilon[p_s(x)\nu\cdot\nabla f(x) - f(x)\nu\cdot\nabla p_s(x)]\} \tag{9.3.25}$$

注意，到 ε 的最低阶

$$\nu\cdot\nabla f(x) = -\frac{1}{\varepsilon}\frac{\partial f}{\partial \rho} = -\nu\cdot v(x)[C_0 - g(x)] \tag{9.3.26}$$

并且

$$\nu\cdot\nabla p_s(x) = -\frac{1}{\varepsilon}\nu\cdot\nabla\phi(x)\exp[-\phi(x)/\varepsilon] \tag{9.3.27}$$

我们导得

$$\boxed{C_0 = \frac{\int_S dS e^{-\phi(x)/\varepsilon}[2\nu\cdot v(x) + \nu\cdot\nabla\phi(x)]g(x)}{\int_S dS e^{-\phi(x)/\varepsilon}\nu\cdot v(x)}} \tag{9.3.28}$$

对此问题，回顾 $g(u) = \delta_\varepsilon(u-b)$，我们发现，如果 x 是 R 内的势井，则

$$\boxed{\pi(x, b) = C_0 = \frac{e^{-\phi(b)/\varepsilon}[2\nu\cdot v(b) + \nu\cdot\nabla\phi(b)]}{\int_S dS e^{-\phi(x)/\varepsilon}\nu\cdot v(x)}} \tag{9.3.29}$$

这里我们看到，出口分布实质上为 $\exp[-\phi(b)/\varepsilon]$，即近似于

平稳分布。如果福克-普朗克方程具有势函数解，则

$$v(b) = -\nabla\phi(b) \tag{9.3.30}$$

而且

$$\pi(x, b) = e^{-\phi(b)/\varepsilon} \nu \cdot v(b) \Big/ \left[\int_S dS e^{-\phi(s)/\varepsilon} \nu \cdot v(x)\right]$$

我们就有了一类平均流结果。

ii) 在 S 面上 $\nu \cdot v(x) = 0$：这个问题与双稳定性更直接有关，因为在两个稳定点 a 和 c 中间划分成两个区域的、被称为分界线的曲线 $\nu \cdot v(x) = 0$，是期望之中的。

该方法除了靠近 δ 上的 u 外是极为相似的。我们预期

$$\nu \cdot v(x) \sim \nu \cdot (x - u)\kappa(u) \tag{9.3.31}$$

式中 $\kappa(u)$ 是依赖于 $v(x)$ 的系数，并且假定不等于零。

现在情况就象 $x = a$ 的那样，并且作如下代换是合适的：

$$x = u - \sqrt{\varepsilon}\rho\nu(u) \tag{9.3.32}$$

而对 ε 的最低阶，(9.3.16)（在接近 S 面上的 u 处）简化为

$$\kappa(u)\rho\frac{\partial f}{\partial \rho} + \frac{\partial^2 f}{\partial \rho^2} = 0 \tag{9.3.33}$$

以致

$$f(x) = g(u) + C_1\int_0^\rho d\rho \exp\left[-\frac{1}{2}\kappa(u)\rho^2\right] \tag{9.3.34}$$

再取极限 $\rho \to \infty$，我们得到

$$C_1 = [C_0 - g(u)]\sqrt{\frac{2\kappa(u)}{\pi}} \tag{9.3.35}$$

现在关于 $\pi(x, b)$ 的结果就象以前那样地完成了：可得

$$\boxed{\pi(x, b) = \frac{e^{-\phi(b)/\varepsilon}\left[\left(1 - \sqrt{\frac{2\kappa(b)}{\pi}}\right)\nu \cdot v(b) + \nu \cdot \nabla\phi(b)\right]}{\int dS e^{-\phi(x)/\varepsilon}\sqrt{\frac{2\kappa(x)}{\pi}}\,\nu \cdot v(x)}} \tag{9.3.36}$$

9.3.2 平均出口时间的渐近分析

根据我们在一维系统中的经验，我们预期R内的平均出口时间对某个$K>0$而言，当$\varepsilon\to 0$时为$\exp[K/\varepsilon]$的量级。因此，我们定义

$$\tau(x)=\exp(-K/\varepsilon)T(x) \tag{9.3.37}$$

式中$T(x)$为开始时在x处的，从R中逸出的平均逃逸时间，而$\tau(x)$满足(根据5.4节)

$$v(x)\cdot\nabla\tau(x)+\varepsilon\nabla^2\tau(x)=-e^{-K/\varepsilon}$$

$$\tau(u)=0 \qquad u\in S \tag{9.3.38}$$

如果上式标度正确，则$\tau(x)$按ε的幂所作的任何展式将会看到指数，所以直至ε的最低阶的方程，本质上就是(9.3.16)式。

和那种情况相同，我们证明在R内部$\tau(x)$实际上是常数，所以在S附近可以写成[在S上$\nu\cdot v(x)\neq 0$的条件下]

$$\tau(x)\sim C_0\{1-\exp[-\nu\cdot v(u)\rho]\} \tag{9.3.39}$$

用$p_s(x)=\exp[-\phi(x)/\varepsilon]$乘(9.3.38)，并利用格林定理得到[使用的方法与(9.3.25)大体相同，但在S上$\tau(x)=0$。]

$$-e^{-K/\varepsilon}\int_R dx\, e^{-\phi(x)/\varepsilon}=-\int_S dS e^{-\phi(x)/\varepsilon}[C_0\nu\cdot v(x)] \tag{9.3.40}$$

即

$$C_0=\int_R dx\, e^{-[K+\phi(x)]/\varepsilon}\Big/\int_S dS e^{-\phi(x)/\varepsilon}\nu\cdot v(x) \tag{9.3.41}$$

根据假设，C_0并象$\exp(A/\varepsilon)$那样作指数变化。在(9.3.41)式的分子中，主要的贡献来自$\phi(x)$在a点出现时的极小值，而在分母中，$\phi(x)$的极小值的点位于S上，这些点称为x_0。于是，分子与分母的比值具有

$$\exp\{[\phi(a)-\phi(x_0)-K]/\varepsilon\}$$

的性质。因此对C_0而言，它渐近于常数

$$K=\phi(a)-\phi(x_0) \tag{9.3.42}$$

当x完全位于R内时，有

$$\tau(x)=\int_R \mathrm{d}x\mathrm{e}^{\phi(x)/\varepsilon}\Big/\int_S \mathrm{d}S\left[\mathrm{e}^{-\phi(x)/\varepsilon}\nu\cdot v(x)\right] \tag{9.3.43}$$

在 S 上的所有点 $\nu\cdot v(x)=0$ 的场合，有

$$\tau(x)\sim C_0\int_0^{\rho_0}\mathrm{d}\rho\,\exp\left[-\frac{1}{2}K(u)\rho^2\right] \tag{9.3.44}$$

因此，在内部

$$\tau(x)\sim C_0\sqrt{\frac{\pi}{2K(u)}} \tag{9.3.45}$$

继续进行类似的分析，我们发现，对完全在 R 内的所有 x，

$$\tau(x)\sim\sqrt{\frac{\pi}{2K(u)}}\int_R \mathrm{d}x\,\mathrm{e}^{\phi(x)/\varepsilon}\Big/\int_S \mathrm{d}S\mathrm{e}^{-\phi(x)/\varepsilon} \tag{9.3.46}$$

9.3.3 多维空间中的克雷默斯方法

克雷默斯方法的推广相对来讲是简单的。我们考虑一个十分普遍的 l 维福克-普朗克方程[为了记号上的方便，我们用 $p(x)$ 作为概率密度]

$$\partial_t p=\nabla\cdot[-v(x)p+\varepsilon D(x)\cdot\nabla p] \tag{9.3.47}$$

其平稳解应为 $p_s(x)$，而且只有当(9.3.47)满足势条件时，$p_s(x)$ 才能表示成显式。设 $p_s(x)$ 在 a 和 c 有两个完全确定的极大，在 b 有一个完全确定的鞍点（图 9.3）。假定 $p_s(x)$ 在鞍点的值远小于在 a 和 c 的值，我们引入一族 $(l-1)$ 维平面 $S(w)$，这里 w 是标记平面的参量。我们选取平面 $S(a)$ 通过 a，$S(b)$ 和 $S(c)$ 分别通过 b 和 c。我们设想 $S(w)$ 这样取向，使得当其中任一个平面受到限制时 $p_s(x)$ 具有唯一一个极大值。与 9.2.1 节相似，我们定义

$$M[S(w)]=\int_{L(w)}\mathrm{d}xP(x) \tag{9.3.48}$$

式中 $L(w)$ 是平面 $S(w)$ 左边的空间区域；因此

$$\dot{M}[S(w)]=\int_{S(w)}\mathrm{d}S\cdot[-v(x)P+\varepsilon\underset{\sim}{D}(x)\cdot\nabla P] \tag{9.3.49}$$

平稳态下的流定义为

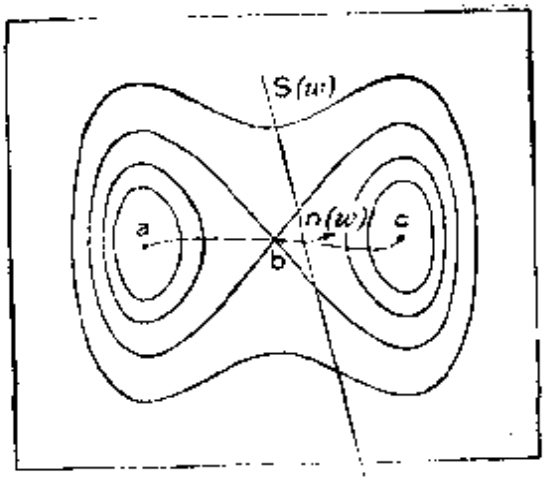


图 9.3　平稳分布函数 $p_s(x)$ 的等值线
这里平面 $S(w)$的取向使得 $P_s(x)$ 具有唯一的一个极大值，
而曲线 $x=u(w)$ (虚线)是这些极大值的轨迹

$$J_s = -v(x)P_s + \varepsilon \underset{\sim}{D}(x)\cdot\nabla P_s \tag{9.3.50}$$

假设 I　我们排除 P_s 很小的地方出现有限流 J_s 的情况，由于 $\nabla\cdot J_s=0$，所以我们可写成

$$J_s = -\varepsilon\nabla\cdot(\underset{\sim}{A}P_s) \tag{9.3.51}$$

式中 $\underset{\sim}{A}$ 是一个反对称张量。要求 $\underset{\sim}{A}$ 的大小具有像 $\underset{\sim}{D}(x)$ 那样或更小的数量级。

弛豫方程分两步导出。定义量 $\beta(x)$ 为

$$\begin{aligned}\beta(x) = P(x,\ t)/P_s(x) &\simeq N_a(t)/n_a \quad (x \text{ 靠近 } a)\\ &\simeq N_c(t)/n_c \quad (x \text{ 靠近 } c)\end{aligned} \tag{9.3.52}$$

这个假设表明，峰内的所有弛豫都已停止。现将它代入(9.3.49)，通过分部积分除掉无穷大的项，得

$$\dot{M}[S(w)] = \varepsilon\int_{S(w)} dS\cdot[\mathscr{D}(x)\cdot\nabla\beta]P_s(x) \tag{9.3.53}$$

其中

$$\mathscr{D}(x) = \underset{\sim}{D}(x) + \underset{\sim}{A}(x) \tag{9.3.54}$$

假设 II　在 $\underset{\sim}{S}(w)$上 $\underset{\sim}{P}_s(x)$是一个陡峭的单峰，所以我们可以构造一个近似解

$$\dot{M}[S(w)] = \{\varepsilon[u(w)\cdot\mathscr{D}(x)\cdot\nabla\beta]_{u(w)} + \delta(w)\}$$

$$\times \left| \int_{S(w)} \mathrm{d}S P_s(x) \right| \tag{9.3.55}$$

式中 $\delta(w)$ 预期比方括号内的项小得多。这里 $u(w)$ 是这样的空间位置，当限制 $S(w)$ 时，$P_s(x)$ 具有其极大值，而且 $n(w)$ 与 $S(w)$ 垂直。

假设 III $u(w)$ 的方向可以这样选取，使得 $\mathscr{D}^T(x)\cdot n(w)$ 平行于曲线 $x=u(w)$ 在 w 处的切线，这并不违背其他假设。因此，

$$\mathscr{D}^T[u(w)]\cdot n(w)=\mathrm{d}(w)\partial_w u(w) \tag{9.3.56}$$

现在定义

$$P(w)=\left| \int_{S(w)} \mathrm{d}S P_s(x) \right| \tag{9.3.57}$$

它是(达到慢变因子)位于平面 $S(w)$ 上的粒子的概率密度，并预期它在 $w=a$ 和 $w=c$ 处具有极大的双峰形状，而在 $w=b$ 处具有极小值。

假设 IV 锐极大和锐极小的假设。忽略 $\delta(w)$，作(9.3.56)的选择，并注意

$$\partial_w u(w)\cdot \nabla\beta[u(w)]=\partial_w \beta[u(w)] \tag{9.3.58}$$

我们发现

$$\frac{1}{\varepsilon}\int_a^{w_0} \mathrm{d}w\{\dot{M}[S(w)]/[P(w)\mathrm{d}(w)]\}=\beta(w_0)-\beta(a) \tag{9.3.59}$$

利用 $p(w)^{-1}$ 具有锐峰的性质，现在(9.3.59)可以近似地取作峰值，利用(9.3.52)及

$$N(a,\ t)=M[S(b),\ t] \tag{9.3.60}$$

与下面的定义相同

$$\kappa(w_0)=\int_a^{w_0}[p(w)]^{-1}\mathrm{d}w \tag{9.3.61}$$

$$\mu(w_0)=\int_{w_0}^{b}[p(w)]^{-1}\mathrm{d}w \tag{9.3.62}$$

就得弛豫方程

$$\kappa(w_0)\dot{N}_a(t)=\varepsilon\mathrm{d}(w_0)[N_0(t)/n_0-N_a(t)/n_a] \tag{9.3.63}$$

$$\mu(w_0)\dot{N}_c(t)=\varepsilon\mathrm{d}(w_0)\ [N_0(t)/n_0-N_c(t)/n_c] \tag{9.3.64}$$

这些表示式与单变量的情况具有完全相同的形式，而且可以作出

完全相同的解释。

9.3.4 例子：双势井中的布朗运动

如5.3.6节所概述的那样，用速度和位置讨论布朗运动。于是考虑福克-普朗克方程

$$\frac{\partial P(x,\ p,\ t)}{\partial t} = -p\frac{\partial P}{\partial x} + U'(x)\frac{\partial P}{\partial p} + \gamma\left[\frac{\partial P}{\partial p}pP + \frac{\partial^2 P}{\partial p^2}\right] \tag{9.3.65}$$

用上述的符号得：

$$\left.\begin{aligned} &\boldsymbol{x} = (x,\ p)\\ &\boldsymbol{v}(\boldsymbol{x}) = (p,\ -U'(x) - \gamma(p))\\ &\boldsymbol{e} = 1\\ &\underset{\sim}{\boldsymbol{D}}(\boldsymbol{x}) = \begin{bmatrix} 0 & 0\\ 0 & \gamma \end{bmatrix}\\ &P_s(\boldsymbol{x}) = \mathcal{N}_2\exp\left[-\frac{1}{2}p^2 - U(x)\right]\\ &\mathcal{N}_2 = (2\pi)^{-1/2}\mathcal{N}_1\\ &\mathcal{N}_1 = \left\{\int_{-\infty}^{\infty} dx\exp[-U(x)]\right\}^{-1} \end{aligned}\right\} \tag{9.3.66}$$

因此，可以写成

$$\boldsymbol{v}(\boldsymbol{x}) = \begin{bmatrix} 0 & -1\\ 1 & \gamma \end{bmatrix}\cdot\nabla(\log P_s) \tag{9.3.67}$$

而平稳态的流为

$$\boldsymbol{J}_s = -\boldsymbol{v}P_s + \underset{\sim}{\boldsymbol{D}}\cdot\nabla P_s = -\nabla\cdot\left\{\begin{bmatrix} 0 & -1\\ 1 & 0 \end{bmatrix}P_s\right\} \tag{9.3.68}$$

所以 $\underset{\sim}{\boldsymbol{A}}$ 存在且为

$$\underset{\sim}{\boldsymbol{A}} = \begin{bmatrix} 0 & -1\\ 1 & 0 \end{bmatrix} \tag{9.3.69}$$

因而假设 I 满足。

平面 $\boldsymbol{S}(w)$ 可以写成如下形式

$$\lambda x + p = w \tag{9.3.70}$$

假设 II 要求在该平面上 $P_s(x)$达到极大值，即在该平面上 $-\frac{1}{2}p^2 - U(x)$为极大。利用标准方法，我们发现极大值一定位于沿曲线 $u(w)$的地方，$u(w)$由下式给出

$$u(w) = \begin{bmatrix} x(w) \\ p(w) \end{bmatrix} = \begin{bmatrix} x(w) \\ w - \lambda x(w) \end{bmatrix} \tag{9.3.71}$$

式中 $x(w)$满足

$$U'[x(w)] + \lambda^2 x(w) - \lambda w = 0 \tag{9.3.72}$$

不管 $P_s(X)$是否是陡峭的峰值，它取决于 $U(x)$的本性。

现在对假设 III 作些补充。

参量 λ 是满足(9.3.56)的那些平面上的特殊位置 w 的函数。$u(w)$的切线平行于

$$\left[\frac{dx}{dw},\ 1 - \lambda\frac{dx}{dw} - x\frac{d\lambda}{dw}\right] \tag{9.3.73}$$

而对(9.3.73)求微商得

$$\frac{dx}{dw} = (U'' + \lambda^2)^{-1}\left[\lambda - \frac{d\lambda}{dw}(2\lambda x - w)\right] \tag{9.3.74}$$

(9.3.70)的法线平行于$(\lambda, 1)$。因此

$$\mathscr{D}^T n = (1+\lambda^2)^{-1/2}\begin{bmatrix} 0 & 1 \\ -1 & \gamma \end{bmatrix}\begin{bmatrix} \lambda \\ 1 \end{bmatrix} = (1+\lambda^2)^{-1/2}\begin{bmatrix} 1 \\ \gamma - \lambda \end{bmatrix} \tag{9.3.75}$$

若

$$\frac{dx}{dw} \Big/ 1 = \left[1 - \lambda\frac{dx}{dw} - x\frac{d\lambda}{dw}\right] \Big/ (\gamma - \lambda) \tag{9.3.76}$$

则(9.3.75)平行于(9.3.73)。现在可以立即求解 (9.3.74, 76)，得

$$\frac{dx}{dw} = \frac{1}{\gamma} - \frac{x}{\gamma^2}\left[\frac{U'' - \gamma\lambda + \lambda^2}{x(U'' + \lambda^2) - (2\lambda x - w)}\right] \tag{9.3.77}$$

$$\frac{d\lambda}{dw} = \frac{1}{\gamma}\left[\frac{U'' - \lambda\gamma + \lambda^2}{x(U'' + \lambda^2) - (2\lambda x - w)}\right] \tag{9.3.78}$$

鞍点位于$(x, p)=(0, 0)$，于是$w=0 \Longleftrightarrow x=0$。把这结果用于(9.3.77)，我们知道必然有

$$x \simeq w/\gamma \qquad \text{当 } w \simeq 0 \text{ 时} \tag{9.3.79}$$

靠近$x=0$处，我们近似写成

$$U[x] \simeq -\frac{1}{2}U_2 x^2 \tag{9.3.80}$$

将(9.3.79)代入(9.3.72)，可得

$$\lambda^2 - \gamma\lambda + U''(0) = 0 \tag{9.3.81}$$

由上式确定

$$\lambda(0) = \frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} + U_2} \tag{9.3.82}$$

(9.3.78)告诉我们，当$w=0$时，$\frac{\mathrm{d}\lambda}{\mathrm{d}w}=0$。于是，根据(9.3.82)，在鞍点附近，$\lambda$将不会有显著的变化，因此，从现在开始，我们将以(9.3.82)作为λ的近似。

只有一个根是合用的。物理上，在高摩擦极限情况下，λ将趋向于∞，这种极限情况将给出克雷默斯的结果，并要求它具有正号。另一个根相应于取这种平面，它上面的$p_s(x)$具有极小值。

现在对(9.3.57)积分，并确定$\mathrm{d}(w)$。注意，$\mathrm{d}(w)$必须用$n(w)$的单位向量来确定。直接代入(9.3.75)，并利用(9.3.79)，得

$$(1+\lambda^2)^{-1/2} = \frac{\mathrm{d}x}{\mathrm{d}w}(w=0)\mathrm{d}(0) = \frac{1}{\gamma}\mathrm{d}(0) \tag{9.3.83}$$

所以

$$\mathrm{d}(0) = \gamma(1+\lambda^2)^{-1/2} \tag{9.3.84}$$

进而得

$$P(w) = \int_{S(w)} |\mathrm{d}S P_s(x)| = \int_{S(w)} \sqrt{\mathrm{d}x^2 + \mathrm{d}p^2}\, P_s(x, p)$$

$$= \mathcal{N}_2 \frac{[1+\lambda^2]^{1/2}}{\lambda} \int \mathrm{d}p \exp\left[-\frac{p^2}{2} - U\left(\frac{w-p}{\lambda}\right)\right] \tag{9.3.85}$$

精确解决定于$U(x)$的选取。取近似

$$U(x) \simeq U_0 - \frac{1}{2}U_2 x^2 \tag{9.3.86}$$

并解得类似高斯结果为

$$p(w) = \frac{(1+\lambda^2)^{1/2}}{\lambda} \mathcal{N}_2 e^{-U_0} \exp\left[\frac{U_2 w^2}{2(\lambda^2 - U_2)}\right] \tag{9.3.87}$$

因而

$$\kappa(0) = \int_{-\infty}^{0} p(w)^{-1} dw = \frac{1}{2} \mathcal{N}_2^{-1} \frac{\lambda\gamma}{(1+\lambda^2)^{1/2}} \frac{e^{U_0}}{\sqrt{U_2}} = \mu(0) \tag{9.3.88}$$

于是，根据适合于多变量理论的(9.2.19)式，我们得到关于从一个势井到点 $x=0$ 的平均首次通过时间

$$\tau_0 = \kappa(0) d(0)^{-1} = \frac{\lambda}{2} e^{U_0} \mathcal{N}_1^{-1} \sqrt{\frac{2\pi}{U_2}} \tag{9.3.89}$$

也就是，

$$\boxed{\tau_0 = \frac{1}{2}\left(\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + U_2}\right) e^{U_0} \mathcal{N}_1^{-1} \sqrt{\frac{2\pi}{U_2}}} \tag{9.3.90}$$

与其它结果相比较

a) 精确的一维平均首先通过时间(斯莫洛乔夫斯基方程)

在大的摩擦极限情况下，我们把克雷默斯方程简化成关于

$$\hat{P}(x, t) = \int dv P(x, v, t) \tag{9.3.91}$$

的斯莫洛乔夫斯基方程，即

$$\frac{\partial \hat{P}(x, t)}{\partial t} = \frac{1}{\gamma} \frac{\partial}{\partial x}\left[U''(x)\hat{P} + \frac{\partial \hat{P}}{\partial x}\right] \tag{9.3.92}$$

这个近似方程的、从 $x=a$ 到 $x=0$ 的平均首先通过时间的精确解为

$$\tau_1 = \gamma \int_a^0 dx \exp[U(x)] \int_{-\infty}^{x} dz \exp[-U(z)] \tag{9.3.93}$$

这个结果可以用数值方法求得。

b) 克雷默斯结果

把我们的方法应用于一维斯莫洛乔夫斯基方程 (9.3.92)，并对所有积分作高斯近似，可以得到克雷默斯结果，其为

$$\tau_2 = \frac{1}{2}\gamma e^{U_0} \mathcal{N}_1^{-1} \sqrt{\frac{2\pi}{U_2}} \tag{9.3.94}$$

它与作 $\lambda \to \gamma$ 代换后的(9.3.90)中的 τ_0 不同，上式显然当限制 γ 很大时成立。在这限制下，

$$\tau_0 \simeq (1 + U_2\gamma^{-2})\tau_2 \tag{9.3.95}$$

c) 修正的斯莫洛乔夫斯基方程

比斯莫洛乔夫斯基方程(9.3.1)更精确的方程是修正的斯莫洛乔夫斯基方程 (6.4.108)：

$$\frac{\partial \hat{P}}{\partial t} = \frac{1}{\gamma}\frac{\partial}{\partial x}\left\{[1 + \gamma^{-2}U''(x)]\left[U''(x)\hat{P} + \frac{\partial \hat{P}}{\partial x}\right]\right\} \tag{9.3.96}$$

现在我们应用标准理论来计算该方程精确的平均首先遍过时间，它为

$$\tau_3 = \gamma\int_a^0 dx[1 + \gamma^{-2}U''(x)]\exp\,[U(x)]\int_{-\infty}^{x} dz\,\exp[-U(z)] \tag{9.3.97}$$

然而应注意，对 x 积分的主要贡献来自 $x=0$ 附近，所以小修正项 $\gamma^{-2}U''(x)$ 与(9.3.97)中令

$$U''(x) \simeq U''(0) = -U_2 \tag{9.3.98}$$

完全等价。这样我们就得到修正的斯莫洛乔夫斯基结果，

$$\tau_3 = (1 - \gamma^{-2}U_2)^{-1}\tau_1 \simeq (1 + \gamma^{-2}U_2)\tau_1 \tag{9.3.99}$$

注意在这极限下，

$$\frac{\tau_3}{\tau_1} = \frac{\tau_0}{\tau_2} \tag{9.3.100}$$

上式意味着，在该极限下所有积分都可作为具有陡峰的高斯分布来计算，我们的结果与修正后的斯莫洛乔夫斯基结果相符。

d) 计算机模拟

通过等价的随机微分方程

$$dx = p\,dt \tag{9.3.101}$$

$$dp = -[\gamma p + U'(x)]dt + \sqrt{2\gamma}dW(t) \tag{9.3.102}$$

的计算机模拟，我们可以估计对平面 S_0，即对直线

$$p = -\lambda x \tag{9.3.103}$$

的平均首次通过时间。对于势场中的给定集来讲，其结果必须用计算机运算。为了估计峰的锐度的影响，我们讨论不同的温度T，即考虑

$$\mathrm{d}x = p\mathrm{d}t \tag{9.3.104}$$

$$\mathrm{d}p = -[\gamma p + U'(x)]\mathrm{d}t + \sqrt{2\gamma T}\,\mathrm{d}w(t) \tag{9.3.105}$$

代入

$$\left.\begin{aligned} p &\to pT^{1/2} \\ x &\to xT^{1/2} \end{aligned}\right\} \tag{9.3.106}$$

我们得

$$\left.\begin{aligned} \mathrm{d}x &= p\mathrm{d}t \\ \mathrm{d}p &= -[\gamma p + V'(x, T)] + \sqrt{2\gamma}\mathrm{d}W(t) \end{aligned}\right\} \tag{9.3.107}$$

式中

$$V(x, T) = U(xT^{1/2}) \tag{9.3.108}$$

用

$$U(x) = \frac{1}{4}(x^2 - 1)^2 \tag{9.3.109}$$

进行计算机模拟，结果如图 9.4 所示。它们自然地分成两组：一组曲线和一组直线。对这个问题的最好回答是：修正的斯莫洛乔夫斯基结果在所有温度下与计算机模拟结果相符，在低温下与用我们的方法所得的结果一致。这样，我们就证实，由于低温相应于陡峰分布，所以在预期适用的范围内该方法是成功的。

也应注意，作为分隔界面 S_0 的选取，就其范围而言也是适当的。对于邻近 $x=0$、$p=0$ 的区域，我们可以写出

$$\mathrm{d}x = p\mathrm{d}t \tag{9.3.110}$$

$$\mathrm{d}p = (-\gamma p + U_2 x)\mathrm{d}t + \sqrt{2\gamma T}\,\mathrm{d}W(t) \tag{9.3.111}$$

$(\mathrm{d}x, \mathrm{d}p)$的决定性部分，即$(p, -\gamma p + U_2 x)$位于点$(x, p)$到原点联线方向的条件是

$$\frac{p}{-\gamma p + U_2 x} = \frac{x}{p} \tag{9.3.112}$$

代入 $p = -\lambda x$，得

$$\lambda^2 - \lambda\gamma - U^2 = 0 \tag{9.3.113}$$

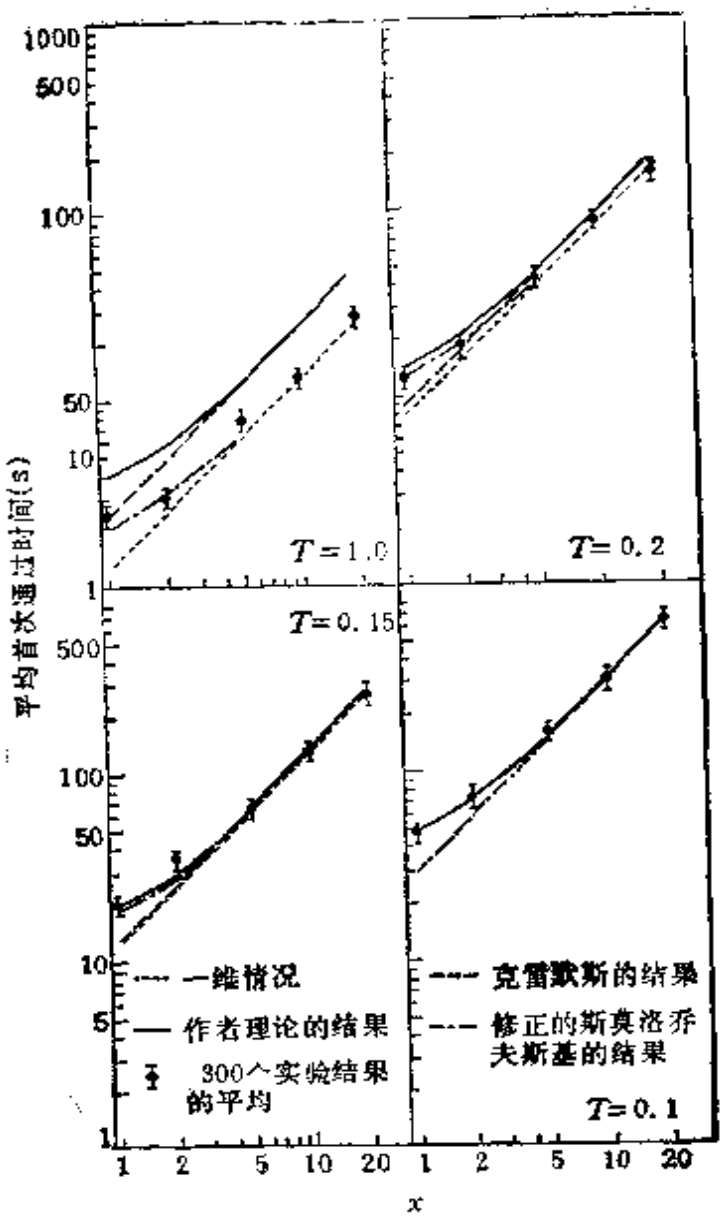


图 9.4　9.3.4节中双井势的平均出口时间的各种估计的比较

在 $x=0$ 的附近，上式与(9.3.81)相同，上式的二个解相应于指向原点的确定性运动（正根）及从原点指向远处的决定性运动（负根）。

因此，当粒子在分界面上运动时，在下一个时间间隔 dt，只有随机项 $dW(t)$会使它离开这个分界面，而且该粒子将以相同的概率向右或向左运动，即，这意味着，在这个平面上向左或向右的分裂概率应为 1:1。

这个分界面的定义也与第 9.1.2节的定义相一致，那里，$v(x)$ 应垂直于 S 的法线。

10. 量子力学马尔可夫过程

二十世纪20年代初以来，人们已认识到量子力学是对包含本质上为统计性质的世界的描述。因此，必须把整个量子力学看成是某种随机过程。用复几率幅(*)作出描述本质上是量子力学所独有的，几率幅模的平方给出一个事件出现的实际概率。

本章的目的不在于系统地陈述完全的量子力学概率理论，或在这推广了的概率理论中用适当定义的随机过程所表达的量子力学。我们感兴趣的是把读者引导到量子理论和经典概率论的交界处，这是一个相当迷人的世界。这个世界是量子光学和量子电子学的领域，那里存在的统计性质，就象涨落来自热效应一样，来自于系统固有的量子特性。我们将证明，光学系统的量子力学怎样才能与形式适当推广了的马尔可夫跳过程密切联系起来。这类马尔可夫过程本身，往往通过称为P表示或象相空间方法那样的其他方法与复平面上的扩散过程密切相联系。这些扩散过程可以描述成取负值或复值的拟概率，或者它们也可以定义地道的正概率。这种情况与第7.7节的泊松表示的情况非常类似，实际上，泊松表示是P表示的一种局限形式。

我们将以如下方式阐述本章。首先略述谐振子的量子力学，并介绍相干态的概念，这是我们的核心任务。然后定义量子马尔可夫过程，并证明广义主方程如何才能用类似于第6章的绝热消去方法推导出来。从这些广义的主方程，有时我们可以演化出普通的生-灭主方程，而且有时利用P表示，可以演化出福克-普朗克

(*) "Probability"一词在数学中都译为"概率"，但在物理学中常常译为"几率"，本章因内容接近物理学，故有时袭用物理学通用译名。——译注

方程。这两种方法都允许我们把经典随机过程中的所有工具全部应用到这些量子力学系统中去。

10.1 谐振子的量子力学

我们用产生和湮灭算子 a^+ 和 a 来描述谐振子。它们满足交换关系

$$[a,\ a^+]\equiv aa^+ - a^+a = \mathbb{1} \tag{10.1.1}$$

并根据这个交换关系可以证明存在 a^+a 如下的本征态 $|n\rangle$，使得

$$a^+a|n\rangle = n|n\rangle \qquad n=0,\ 1,\ 2,\ 3,\ \cdots \tag{10.1.2}$$

$$\left.\begin{aligned} a|n\rangle &= \sqrt{n}\,|n-1\rangle \\ a^+|n\rangle &= \sqrt{n+1}\,|n+1\rangle \end{aligned}\right\} \tag{10.1.3}$$

而

$$\langle n|m\rangle = \delta_{m,n} \tag{10.1.4}$$

算子 N 定义为

$$N = a^+a \tag{10.1.5}$$

而且根据(10.1.2)，N 称为数算子，它的本征值是整数 n。

谐振子本身由其哈密顿函数来定义

$$H = \left(a^+a + \frac{1}{2}\right)\hbar\omega \tag{10.1.6}$$

式中

$$h = 2\pi\hbar \tag{10.1.7}$$

是普朗克常数，ω 是频率。H 的本征态当然是 $|n\rangle$。H 的本征值为

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \tag{10.1.8}$$

谐振子的动力学行为由薛定谔方程引入，该方程确定任何物理态 $|A,\ t\rangle$ 的时间演化。

众所周知，薛定谔方程的形式为

$$H|A,\ t\rangle = i\hbar\partial_t|A,t\rangle \tag{10.1.9}$$

正交性(10.1.4)意味着我们可以利用能量本征态 $|n\rangle$ 来展开任意

的态。

$$|A,\ t\rangle = \sum_n |n\rangle\langle n|A,\ t\rangle \tag{10.1.10}$$

因此，

$$\left.\begin{aligned} \mathrm{i}\hbar\partial_t|A,\ t\rangle &= \mathrm{i}\hbar\sum_n |n\rangle\partial_t\langle n|A,t\rangle \\ &= \sum_n H|n\rangle\langle n|A,\ t\rangle \\ &= \sum_n\left(n+\frac{1}{2}\right)\hbar\omega|n\rangle\langle n|A,\ t\rangle \end{aligned}\right\} \tag{10.1.11}$$

所以

$$\begin{aligned} \langle n|A,\ t\rangle &= \mathrm{e}^{-\mathrm{i}E_n t/\hbar}\langle n|A,\ 0\rangle \\ &= \mathrm{e}^{-\mathrm{i}\left(n+\frac{1}{2}\right)\omega t}\langle n|A,\ 0\rangle \end{aligned} \tag{10.1.12}$$

因而，一个任意态的时间演化就被完全确定了。

10.1.1 与外场的相互作用

谐振子与外场的相互作用是一个半经典问题。这里我们不去涉及太深的物理概念，仅指出将(10.1.6)的哈密顿量修改成

$$H(\alpha) = \left[\left(a^+a+\frac{1}{2}\right) - (\alpha a^+ + \alpha^* a) + |\alpha|^2\right]\hbar\omega \tag{10.1.13}$$

就行了，这里 α 是复数。式中的三部分可以分别看成是谐振子的能量，驱动场 α 与谐振子的相互作用能以及驱动场的（恒定）能量。

将 $H(\alpha)$ 改写成

$$H(\alpha) = \hbar\omega\left[(a-\alpha)^+(a-\alpha)+\frac{1}{2}\right] \tag{10.1.14}$$

是有利的。显然，由于 α 是一个复常数，所以 $(a-\alpha)^+$ 和 $(a-\alpha)$ 与 a^+ 和 a 遵循相同的交换关系。既然态 $|n\rangle$ 的存在是从唯一的交换关系得到的，因此，能量本征态具有相同的形式。

方程(10.1.3)可以用来定义 a^+a 的基态或真空态 $|0\rangle$，其为

$$a|0\rangle = 0 \tag{10.1.15}$$

移位算子 $(a-\alpha)^+$，$(a-\alpha)$ 的相应方程为

$$a|\alpha\rangle = \alpha|\alpha\rangle \tag{10.1.16}$$

可以由(10.1.3)加以核实，$|\alpha\rangle$的解为

$$|\alpha\rangle = \exp\left(-\frac{1}{2}|\alpha|^2\right)\sum_{n=0}^{\infty}\frac{2n}{\sqrt{n!}}|n\rangle \tag{10.1.17}$$

式中精确因子 $\exp\left(-\frac{1}{2}|\alpha|^2\right)$选取成使

$$\langle\alpha|\alpha\rangle = 1 \tag{10.1.18}$$

态$|\alpha\rangle$是由格劳伯(Glauber)[10.1]所提出的，称为相干态。

当谐振子被看成为单模系统的辐射场模型时，相干态可以看成为接近经典态的量子力学态。

现在哈密顿量的能量本征态是 $|n;\alpha\rangle$，并且与用移位算子 $(a-\alpha)^+$ 和 $(a-\alpha)$改写成的(10.1.2～4)式具有相同的性质。

10.1.2 相干态的性质

我们列出最重要的性质；如果结果是代入定义就可证明的简单情况，就不给予证明了。

a) 定义：

$$|\alpha\rangle = \exp\left(-\frac{1}{2}|\alpha|^2\right)\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle \tag{10.1.19}$$

b) 标积：

$$\langle\alpha|\beta\rangle = \exp\left(\alpha^*\beta - \frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2\right) \tag{10.1.20}$$

$$|\langle\alpha|\beta\rangle|^2 = \exp(-|\alpha-\beta|^2) \tag{10.1.21}$$

c) 完备性公式

$$1 = \frac{1}{\pi}\int d^2\alpha|\alpha\rangle\langle\alpha| \tag{10.1.22}$$

这里

$$\alpha = \alpha_x + i\alpha_y$$

$$d^2\alpha = d\alpha_x\, d\alpha_y \tag{10.1.23}$$

面积分对整个复平面进行。我们来证明这一点。若$|A\rangle$是一个任意矢量，则可写成

$$|A\rangle = \sum_n A_n |n\rangle \tag{10.1.24}$$

以致

$$\frac{1}{\pi}\int \mathrm{d}^2\alpha\,|\alpha\rangle\langle\alpha|A\rangle = \frac{1}{\pi}\sum_n \int A_n |\alpha\rangle\langle\alpha|n\rangle \mathrm{d}^2\alpha \tag{10.1.25}$$

代入(10.1.19)，并用极坐标

$$\alpha = r\mathrm{e}^{i\theta} \tag{10.1.26}$$

$$\mathrm{d}^2\alpha = r\mathrm{d}r\mathrm{d}\theta \tag{10.1.27}$$

作变换，因此

$$\begin{aligned}(10.1.25) &= \frac{1}{\pi}\sum_{m,n}\int A_n\,\mathrm{e}^{-r^2} r^{n+m}\mathrm{e}^{i(m-n)\theta}(n!\,m!)^{-1/2} \\ &\quad \cdot |m\rangle r\mathrm{d}r\mathrm{d}\theta \\ &= 2\sum_n \int A_n\,\mathrm{e}^{-r^2} r^{2n+1}(n!)^{-1}|n\rangle \mathrm{d}r \end{aligned} \tag{10.1.28}$$

式中用了

$$2\pi\delta_{n,m} = \frac{1}{\pi}\int \mathrm{d}\theta \mathrm{e}^{i(m-n)\theta} \tag{10.1.29}$$

现在注意

$$\int_0^\infty \mathrm{d}r\mathrm{e}^{-r^2} r^{2n+1} = n!/2 \tag{10.1.30}$$

我们得到

$$(10.1.25) = \sum_n A_n |n\rangle = |A\rangle \tag{10.1.31}$$

公式(10.1.21, 22)一起指出，相干态对不同的 α 和 β 不是正交的态，而且由于积分(10.1.22)存在因子 $\frac{1}{\pi}$，所以相干态是过完备态。[事实上，对任何 r，(10.1.19)表明，我们可以写成

$$|n\rangle = \exp\left(\frac{1}{2}r^2\right) r^{-n}\sqrt{n!}\int \mathrm{d}\theta\,\mathrm{e}^{-in\theta}|\alpha\rangle \tag{10.1.32}$$

上式指出，对任何固定的 $r = |\alpha|$ 的态是完备的]。然而，由于相干态和经典物理态之间存在非常简单的联系，又因为伯格曼(Bargmann)态是 α 的解析函数这个事实，所以过完备性并非是缺点。伯格曼态定义为

$$|\alpha\rangle = \exp\left(\frac{1}{2}|\alpha|^2\right)|\alpha\rangle = \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle \tag{10.1.33}$$

伯格曼态的这个性质在下面的讨论中非常重要。

d) 任意态用相干态来展开

考虑一个任意态$|f\rangle$，然后用完备性关系(10.1.32)得

$$|f\rangle = \frac{1}{\pi}\int d^2\alpha |\alpha\rangle f(\alpha^*)\exp\left(-\frac{1}{2}|\alpha|^2\right) \tag{10.1.34}$$

式中

$$f(\alpha^*) = \langle\alpha|f\rangle\exp\left(\frac{1}{2}|\alpha|^2\right) = \langle\alpha\|f\rangle \tag{10.1.35}$$

是α^*的解析函数。用这些条件，展开式(10.1.34)是唯一的。如果允许存在α^*和α两个量的函数，则正如格劳伯所证明的那样，展开式不再是唯一的。

两个态f和g的标积可以直截了当地表示为

$$\langle g|f\rangle = \frac{1}{\pi}\int d^2\alpha [g(\alpha^*)]^* f(\alpha^*)e^{-|\alpha|^2} \tag{10.1.36}$$

显然，该标积是解析函数希尔伯特空间。实际上，它为谐振子研究、产生湮灭算子和本章所有公式体系提供了有效的数学出发点。

e) 算子用相干态来展开

考虑量子希尔伯特空间中的一个任意算子T，用两次单位分解就得到

$$\begin{aligned} T &= 1\cdot T\cdot 1 = \frac{1}{\pi^2}\int d^2\alpha\, d^2\beta |\alpha\rangle\langle\alpha|T|\beta\rangle\langle\beta| \\ &= \frac{1}{\pi^2}\int d^2\alpha\, d^2\beta |\alpha\rangle\langle\beta| T(\alpha^*, \beta) \\ &\quad \cdot\exp\left(-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2\right) \end{aligned} \tag{10.1.37}$$

式中

$$\begin{aligned} T(\alpha^*, \beta) &= \exp\left(\frac{1}{2}|\alpha|^2 + \frac{1}{2}|\beta|^2\right)\langle\alpha|T|\beta\rangle \\ &= \langle\alpha|T\|\beta\rangle \end{aligned} \tag{10.1.38}$$

而根据态$\|\alpha\rangle$，$\|\beta\rangle$的解析性，我们知道$T(\alpha^*, \beta)$是α^*和β的解析函数，而且以此条件，它是唯一的。注意，例如若

$$T = (a^+)^m(a^n) \tag{10.1.39}$$

则

$$\left.\begin{aligned} T(\alpha^*, \beta) &= \langle\alpha|(a^+)^m(a^n)|\beta\rangle \exp\left(\frac{1}{2}|\alpha|^2 + \frac{1}{2}|\beta|^2\right) \\ &= (\alpha^*)^m(\beta)^n\langle\alpha|\beta\rangle \exp\left(\frac{1}{2}|\alpha|^2 + \frac{1}{2}|\beta|^2\right) \\ &= (\alpha^*)^m(\beta)^n \exp(\alpha^*\beta) \end{aligned}\right\} \tag{10.1.40}$$

f) 任意算子T由它在所有相干态的期待值决定

由于

$$\langle\alpha|T|\alpha\rangle = \sum_{n,m}\langle n|T|m\rangle e^{|\alpha|^2}(\alpha^*)^n(\alpha)^m/\sqrt{n!m!}$$

所以

$$\langle n|T|m\rangle = \sqrt{n!m!}\,\frac{\partial^n}{\partial\alpha^{*n}}\,\frac{\partial^m}{\partial\alpha^m}\left(e^{\alpha\alpha^*}\langle\alpha|T|\alpha\rangle\right) \tag{10.1.41}$$

该导数是形式上的导数，和解析函数理论中的一样，它们可以解释为

$$\left.\begin{aligned} \alpha &= x + iy \\ \frac{\partial}{\partial\alpha} &= \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \\ \frac{\partial}{\partial\alpha^*} &= \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) \end{aligned}\right\} \tag{10.1.42}$$

在这种情况下，$\langle n|T|m\rangle$是实变量x和y的幂级数的系数组合。

g) 相干态是a的本征态

$$\left.\begin{aligned} a|\alpha\rangle &= \alpha|\alpha\rangle \\ \text{及} \\ \langle\alpha|a^+ &= \langle\alpha|\alpha^* \end{aligned}\right\} \tag{10.1.43}$$

上两式可以根据定义直接证明，而定义是研究它们的原始基础。在等价的矩阵元中，算子的正规乘积是很有用的。在正规乘积中，所有湮灭算子都排在产生算子的右面。这样，

$$\left.\begin{aligned}\langle\alpha|a^+aa^+|\beta\rangle &= \langle\alpha|a^+a^+a+a^+[a,\ a^+]|\beta\rangle\\ &= \langle\alpha|a^+a^+a+a^+|\beta\rangle\\ &= (\alpha^{*2}\beta+\alpha^*)\langle\alpha|\beta\rangle\end{aligned}\right\} \tag{10.1.44}$$

在表示式两端的符号：　：意味着把它看成是正规乘积，于是

$$:(a+a^+)(a+a^+):=a^{+2}+a^2+2a^+a \tag{10.1.45}$$

b) 相干态的泊松分布

由于态$|n\rangle$的能量为$n\hbar\omega$，比真空态$|0\rangle$或零量子态的能量大，所以把它称为"第n个量子态"。因此，量子力学中，在相干态$|\alpha\rangle$中观察到第n个量子的概率为

$$P_\alpha(n)=|\langle n|\alpha\rangle|^2=\left|\exp\left(-\frac{1}{2}\alpha^2\right)\frac{\alpha^n}{\sqrt{n!}}\right|^2 \tag{10.1.46}$$

$$=\frac{e^{-|\alpha|^2}|\alpha|^{2n}}{n!} \tag{10.1.47}$$

它是用$|\alpha|^2$表示的泊松分布。

由于数n与数算子N的本征值相对应，我们得

$$\langle N\rangle=\langle\alpha|N|\alpha\rangle=\sum_n nP(n)=|\alpha|^2 \tag{10.1.48}$$

$$\begin{aligned}\langle N^2\rangle &= \langle\alpha|a^+aa^+a|\alpha\rangle\\ &= \langle\alpha|a^+a^+aa+a^+[a,\ a^+]a|\alpha\rangle\\ &= |\alpha|^4+|\alpha|^2\end{aligned}$$

因此，

$$\langle N(N-1)\rangle=|\alpha|^4=\langle N\rangle^2 \tag{10.1.49}$$

是泊松表示所需要的。相干态中量子的概率分布的泊松特性提供了与泊松表示的联系。

10.2　密度矩阵和概率

量M在态$|\psi\rangle$的平均值由

$$\langle M\rangle=\langle\psi|M|\psi\rangle \tag{10.2.1}$$

给出，这个常见的量子力学公式仅为这些实验作准备，即其中的相同量子态被重复测量的情况。更常见的可能性是，热效应或者不过是不适当的制备装置都引起状态制备的随机性，由于这种随机性，

我们每次所做的测量测得的是不同状态$|\psi_a\rangle$的平均值。因此，在这种情况下，如果每个态以概率$P(a)$出现，而这种概率不是来自量子效应而仅来自状态制备的随机性，则测得的平均值为

$$\langle M\rangle = \Sigma P(a)\langle\psi_a|M|\psi_a\rangle \tag{10.2.2}$$

现在引入密度矩阵(或算子)ρ

$$\rho = \sum_a P(a)|\psi_a\rangle\langle\psi_a| \tag{10.2.3}$$

由上式得

$$\langle M\rangle = \mathrm{Tr}\{\rho M\} \tag{10.2.4}$$

这里对任意算子B定义求迹运算为

$$\mathrm{Tr}\{B\} = \sum_n \langle n|B|n\rangle \tag{10.2.5}$$

所以得

$$\begin{aligned}\mathrm{Tr}\{\rho M\} &= \sum_{n,a} P(a)\langle n|\psi_a\rangle\langle\psi_a|M|n\rangle \\ &= \sum_{na} P(a)\langle\psi_a|M|n\rangle\langle n|\psi_a\rangle \\ &= \sum_a P(a)\langle\psi_a|M|\psi_a\rangle \\ &= \langle M\rangle\end{aligned}$$

迹的重要性质是其诸因子的循环置换不变性，故有

$$\mathrm{Tr}\{\rho M\} = \mathrm{Tr}\{M\rho\} \tag{10.2.6}$$

$$\mathrm{Tr}\{ABCD\} = \mathrm{Tr}\{BCDA\}\text{等等。} \tag{10.2.7}$$

重要的密度矩阵性质：

i) $\mathrm{Tr}\{\rho\} = 1$ (10.2.8)

因为 $\mathrm{Tr}\{\rho\} = \sum_a P(a)\langle\psi_a|\psi_a\rangle = \sum_a P_a = 1$ (10.2.9)

ii) ρ 是半正定的。因为对任何态$|A\rangle$

$$\langle A|\rho|A\rangle = \sum P_a|\langle A|\psi_a\rangle|^2 \geqslant 0 \tag{10.2.10}$$

iii) 如果 ρ 对应于一个纯态，则 $\rho^2 = \rho$；相反，对于纯态，$P_a = \delta_{a,a_0}$

仅对某些 a_0 成立，所以

$$\rho^2 = |\psi_{a_0}\rangle\langle\psi_{a_0}|\psi_{a_0}\rangle\langle\psi_{a_0}| = \rho \tag{10.2.11}$$

相反地考虑

$$\rho^2 = \Sigma P_a P_b |\psi_a\rangle\langle\psi_b|\langle\psi_a|\psi_b\rangle \tag{10.2.12}$$

上式仅在下列条件才等于 ρ，即不仅

$$P_a P_b \langle\psi_a|\psi_b\rangle = 0 \qquad (\text{当 } a \neq b \text{ 时})$$

$$P_a^2 \langle\psi_a|\psi_a\rangle = P_a$$

而且

$$\langle\psi_a|\psi_a\rangle = 1$$

因此，

$$P_a^2 = P_a \Longrightarrow P_a = 1 \text{ 或 } 0$$

但因 $\Sigma P_a = 1$，所以只有一个 P_a 能够为 1，其余的均为 0，它相应于一个纯态。

iv) $\mathrm{Tr}\{\rho^2\} \leqslant \mathrm{Tr}\{\rho\}$，等号仅对纯态成立。

因为

$$\mathrm{Tr}\{\rho^2\} = \sum_{a,b} P_a P_b |\langle\psi_a|\psi_b\rangle|^2 \tag{10.2.13}$$

并由于

$$|\langle\psi_a|\psi_b\rangle|^2 \leqslant 1 \tag{10.2.14}$$

$$\sum_b P_b |\langle\psi_a|\psi_b\rangle|^2 \leqslant 1 \qquad (\text{对任意 } a) \tag{10.2.15}$$

因此，

$$\mathrm{Tr}\{\rho^2\} \leqslant \sum_a P_a = \mathrm{Tr}\{\rho\} \tag{10.2.16}$$

而且显然只有对某个 a_0

$$P_b = \delta_{b,a_0}$$

时，即 ρ 为纯态时，式(10.2.15)才取等号。

10.2.1 冯·诺伊曼方程

任意态的薛定谔方程为

$$H|\psi\rangle = i\hbar\partial_t|\psi\rangle \tag{10.2.17}$$

可以导出相应的密度算子方程。因为

$$\partial_t\rho = \sum_a P_a[(\partial_t|\psi_a\rangle)\langle\psi_a| + |\psi_a\rangle(\partial_t\langle\psi_a|)]$$

$$= \frac{1}{i\hbar}(H\rho - \rho H)$$

即

$$\boxed{[H,\ \rho]=i\hbar\partial_t\rho} \tag{10.2.18}$$

这就是冯·诺伊曼方程或量子刘维方程，冯·诺伊曼方程可以表示成指数形式以给出形式解：

$$\rho(t)=\exp(-iHt/\hbar)\cdot\rho(0)\exp(iHt/\hbar) \tag{10.2.19}$$

10.2.2 格劳伯-苏达香P表示

格劳伯[10.1]和苏达香[10.2] 引进了现在称为格劳伯-苏达香P表示的密度矩阵(表示)。我们假定ρ可写成

$$\rho=\int d^2\alpha P(\alpha,\ \alpha^*)\,|\alpha\rangle\langle\alpha| \tag{10.2.20}$$

式中$P(\alpha,\ \alpha^*)$起着拟概率的作用。我们把P表示的存在问题留到和矩一起讨论。

注意，由于

$$1=\langle\alpha|\alpha\rangle=\sum_n\langle\alpha|n\rangle\langle n|\alpha\rangle$$

所以

$$\mathrm{Tr}\{\rho\}=1=\int d^2\alpha P(\alpha,\ \alpha^*)\sum_n\langle n|\alpha\rangle\langle\alpha|n\rangle$$

即

$$\boxed{1=\int d^2\alpha P(\alpha,\ \alpha^*)} \tag{10.2.21}$$

进而对任何正规乘积$(a^+)^r a^s$，

$$\mathrm{Tr}\{(a^+)^r a^s\rho\}=\mathrm{Tr}\{a^s\rho(a^+)^r\}$$

$$=\int d^2\alpha P(\alpha,\ \alpha^*)\sum_n\langle n|a^s|\alpha\rangle\langle\alpha|(a^+)^r|n\rangle$$

即

$$\boxed{\langle(a^+)^r a^s\rangle=\int d^2\alpha(\alpha^*)^r\alpha^s P(\alpha,\ \alpha^*)} \tag{10.2.22}$$

于是量 $P(\alpha, \alpha^*)$ 对变量 α 和 α^* 起着一类概率密度的作用，这种变量的量子算子的正规序列乘积的平均值就是 $P(\alpha, \alpha^*)$ 的简单矩。

格劳伯-苏达香 P 表示的这个存在条件是值得怀疑的。格劳伯和苏达香已证明[10.3]，只要选用充分奇异的广义函数，这种表示总是存在的。对一个正函数，它确实并不总是存在，就是对一个光滑函数它也不肯定存在。

我们将把这些问题留待和矩阵一起讨论，这个问题的答案是与第7章时这些问题的各种泊松表示非常相似的。

10.2.3 算子对应

我们知道，

$$\left.\begin{aligned} &a|\alpha\rangle = \alpha|\alpha\rangle \\ &\text{及} \\ &\langle\alpha|a^+ = \alpha^*\langle\alpha| \end{aligned}\right\} \tag{10.2.23}$$

对于用 a、a^+ 的其他可能的作用方式，利用伯格曼态 $|\alpha\rangle$ 是方便的。结果

$$\begin{aligned} a^+|\alpha\rangle &= \sum_n \frac{\alpha^n}{\sqrt{n!}}\sqrt{n+1}\,|n+1\rangle \\ &= \frac{\partial}{\partial\alpha}|\alpha\rangle \end{aligned} \tag{10.2.24}$$

类似地

$$\langle\alpha|a = \frac{\partial}{\partial\alpha^*}\langle\alpha|$$

因此，对于可以用便于变换的 P 表示描述的 ρ 来说，我们可以为此目的而把它写成形式

$$\rho = \int d^2\alpha |\alpha\rangle\langle\alpha| e^{-\partial\alpha^*} P(\alpha, \alpha^*)$$

简单方程

$$a^+\rho = \int d^2\alpha \frac{\partial}{\partial\alpha}(|\alpha\rangle)\langle\alpha| e^{-\partial\alpha^*} P(\alpha, \alpha^*)$$

经分部积分后为

$$=\int d^2\alpha |\alpha\rangle\langle\alpha| e^{-\partial\alpha^*}\left(\alpha^*-\frac{\partial}{\partial\alpha}\right)P(\alpha,\alpha^*) \tag{10.2.25}$$

于是我们能构成 a^+ 和$\left(\alpha^*-\frac{\partial}{\partial\alpha}\right)$之间的一个算子对应。对 a 也存在一个类似的公式。将来自 (10.2.23, 24) 的明显的对应作一概括，得

$$\boxed{\begin{aligned}
&a\rho\longleftrightarrow\alpha P(\alpha,\alpha^*)\\
&a^+\rho\longleftrightarrow\left(\alpha^*-\frac{\partial}{\partial\alpha}\right)P(\alpha,\alpha^*)\\
&\rho a\longleftrightarrow\left(\alpha-\frac{\partial}{\partial\alpha^*}\right)P(\alpha,\alpha^*)\\
&\rho a^+\longleftrightarrow\alpha^*P(\alpha,\alpha^*)
\end{aligned}} \tag{10.2.26}$$

10.2.4 对受迫谐振子的应用

考虑哈密顿量

$$H=\hbar\omega\left(a^+a+\frac{1}{2}\right)+(\lambda a^++\lambda^*a) \tag{10.2.27}$$

对于该哈密顿量量子刘维方程为

$$i\hbar\frac{\partial\rho}{\partial t}=\hbar\omega[a^+a,\rho]+\lambda[a^+,\rho]+\lambda^*[a,\rho] \tag{10.2.28}$$

现在利用算子对应(10.2.26)，将上式转换成为 $P(\alpha,\alpha^*)$ 的方程。于是，

$$a^+a\rho\longrightarrow\left(\alpha^*-\frac{\partial}{\partial\alpha}\right)\alpha P \tag{10.2.29}$$

$$-\rho a^+a\longrightarrow-\left(\alpha-\frac{\partial}{\partial\alpha^*}\right)\alpha^*P \tag{10.2.30}$$

[注意在(10.2.30)中算子的次序颠倒了，因为对 ρ 的作用是从右面进行的，而对 P 的作用是从左面进行的]：

$$[a^+,\rho]\rightarrow\left[\left(\alpha^*-\frac{\partial}{\partial\alpha}\right)-\alpha^*\right]P$$

$$= -\frac{\partial}{\partial\alpha}P \tag{10.2.31}$$

类似地

$$[a,\ \rho]\to\frac{\partial}{\partial\alpha^*}P \tag{10.2.32}$$

所以我们得到

$$\frac{\partial P}{\partial t}=i\left(-\omega\frac{\partial}{\partial\alpha}\alpha+\omega\frac{\partial}{\partial\alpha^*}\alpha^*-\frac{\lambda}{\hbar}\frac{\partial}{\partial\alpha}+\frac{\lambda^*}{\hbar}\frac{\partial}{\partial\alpha^*}\right)P \tag{10.2.33}$$

这相应于各变量的刘维方程，必须小心，别把 α 和 α^* 处理为独立变量，那是不严格的，在写出上述对应关系时，我们其实应该写成

$$\left.\begin{aligned}\alpha&=x+iy\\ \alpha^*&=x-iy\\ \frac{\partial}{\partial\alpha}&=\frac{1}{2}\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)\\ \frac{\partial}{\partial\alpha^*}&=\frac{1}{2}\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\end{aligned}\right\} \tag{10.2.34}$$

及

$$\lambda=\mu+i\nu$$

并利用这些实变数，(10.2.33)变成

$$\frac{\partial P}{\partial t}=\left[\frac{\partial}{\partial x}(\omega y+\nu/\hbar)-\frac{\partial}{\partial y}(\omega x+\mu/\hbar)\right]P \tag{10.2.35}$$

上式是刘维方程，与下列微分方程等价

$$\left.\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t}&=-\omega y-\nu/\hbar\\ \frac{\mathrm{d}y}{\mathrm{d}t}&=\omega x+\mu/\hbar\end{aligned}\right. \tag{10.2.36}$$

上式又等价于方程

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t}=i(\omega\alpha+\lambda/\hbar) \tag{10.2.37}$$

式(10.2.37)的解为

$$\alpha=-\lambda/\hbar\omega+\xi e^{i\omega t} \tag{10.2.38}$$

假设一个确定的初条件后，P 的解为

$$P(\alpha, \alpha^*, t) = \delta\left(x - \mathrm{Re}\left\{-\frac{\lambda}{\hbar\omega} + \xi e^{i\omega t}\right\}\right)$$

$$\cdot \delta\left(y - \mathrm{Im}\left\{-\frac{\lambda}{\hbar\omega} + \xi e^{i\omega t}\right\}\right) \tag{10.2.39}$$

$$= \delta^2(\alpha + \lambda/\hbar\omega - \xi e^{i\omega t}) \tag{10.2.40}$$

式中复 δ 函数理解为对(10.2.39)的平均。

注意 λ 可能与时间有关，在这种情况下，(10.2.37)变成

$$\frac{d\alpha}{dt} = i[\omega\alpha + \lambda(t)/\hbar] \tag{10.2.41}$$

该方程的解为

$$\alpha(t) = \alpha(0)e^{i\omega t} + i\int_0^t dt' e^{i\omega(t-t')}\lambda(t')/\hbar \tag{10.2.42}$$

而相应的 P 为

$$P(\alpha, \alpha^*, t) = \delta^2[\alpha - \alpha(t)] \tag{10.2.43}$$

10.2.5 量子特征函数

$P(\alpha, \alpha^*)$ 的傅里叶变换将提供一个象 2.6 节引进的那种固有特征函数。我们定义

$$\chi(\lambda, \lambda^*) = \int d^2\alpha \exp(\lambda\alpha^* - \lambda^*\alpha)P(\alpha, \alpha^*) \tag{10.2.44}$$

注意，如果

$$\alpha = x + iy$$

$$\lambda = \mu + i\nu$$

则

$$\lambda\alpha^* - \lambda^*\alpha = 2i(\nu x - y\mu) \tag{10.2.45}$$

所以(10.2.44)是一个两个实变量的傅里叶变换，它也可以写成

$$\boxed{\chi(\lambda, \lambda^*) = \mathrm{Tr}\{\rho \exp(\lambda a^+)\exp(-\lambda^* a)\}} \tag{10.2.46}$$

上式能够起到量子特征函数的一般定义作用。注意，对任何算子 A，

$$\mathrm{Tr}\{A\} = \sum_n \langle n|A|n\rangle = \frac{1}{\pi}\int d^2\alpha \sum_n \langle n|\alpha\rangle\langle\alpha|A|n\rangle$$
$$= \frac{1}{\pi}\int d^2\alpha \sum_n \langle\alpha|A|n\rangle\langle n|\alpha\rangle$$

即

$$\boxed{\mathrm{Tr}\{A\} = \frac{1}{\pi}\int d^2\alpha \langle\alpha|A|\alpha\rangle} \qquad (10.2.47)$$

现在我们引入贝克-豪斯道夫(Baker-Hausdorff)公式[10.4]。对于任意两个算子A和B, 若它们的换位子$[A, B]$与它们本身相互对易,则我们可以写出

$$\boxed{\begin{aligned}\exp(A+B) &= \exp(A)\exp(B)\exp\left(-\frac{1}{2}[A, B]\right)\\ &= \exp(B)\exp(A)\exp\left(\frac{1}{2}[A, B]\right)\end{aligned}} \qquad (10.2.48)$$

注意,

$$[\lambda a^+, -\lambda^* a] = |\lambda|^2$$

我们知道

$$\exp(\lambda a^+)\exp(-\lambda^* a) = \exp(-\lambda^* a)\exp(\lambda a^+)\exp(|\lambda|^2)$$

因此,

$$\mathrm{Tr}\{\rho\exp(\lambda a^+)\exp(-\lambda^* a)\}$$
$$= \exp(|\lambda|^2)\mathrm{Tr}\{\rho\exp(-\lambda^* a)\exp(\lambda a^+)\}$$
$$= \exp(|\lambda|^2)\mathrm{Tr}\{\exp(\lambda a^+)\rho\exp(-\lambda^* a)\}$$
$$= \frac{\exp(|\lambda|^2)}{\pi}\int d^2\alpha\langle\alpha|\exp(\lambda a^+)\rho\exp(-\lambda^* a)|\alpha\rangle$$

所以得

$$\boxed{\chi(\lambda, \lambda^*) = \frac{\exp(|\lambda|^2)}{\pi}\int d^2\alpha\exp(\lambda\alpha^* - \lambda^*\alpha)\langle\alpha|\rho|\alpha\rangle} \qquad (10.2.49)$$

由于

$$\langle\alpha|\rho|\alpha\rangle>0 \tag{10.2.50}$$

以及

$$\mathrm{Tr}\{\rho\}=\frac{1}{\pi}\int d^2\alpha\langle\alpha|\rho|\alpha\rangle=1 \tag{10.2.51}$$

所以(10.2.49)是函数$\langle\alpha|\rho|\alpha\rangle$的傅里叶变换，它满足概率函数的条件。因此，傅里叶变换$\chi(\lambda,\lambda^*)\exp(-|\lambda|^2)$是它的相应的特征函数，而根据2.6节，它以概率1决定$\langle\alpha|\rho|\alpha\rangle$。

由10.1.2节，我们知道$\langle\alpha|\rho|\alpha\rangle$确定着$\rho$，因此，$\chi(\lambda,\lambda^*)$以概率1决定$\rho$。

10.3 量子马尔可夫过程

现在我们简略地展开一种阻尼的量子理论的简化形式，它需要某些量子统计的知识。

10.3.1 热槽

实际上，阻尼是因为系统与另一个称为热槽的非常大的系统相作用而产生的。在热槽内系统是耗散的。但是，由于热槽把本身的一些能量分还给系统，所以也产生了噪声。

作为热槽的一个模型，考虑具有算子Γ_i和哈密顿量

$$H_B=\sum_i\hbar\omega_i\left(\Gamma_i^+\Gamma_i+\frac{1}{2}\right) \tag{10.3.1}$$

的大量独立谐振子。事实上，这个系统的确具有稳定的密度矩阵。确实，H_B的任意正函数都是满足的。统计力学允许我们去选取一个正则系综，其中密度算子为

$$\rho(T)=\exp(-H_B/kT)\mathrm{Tr}\{\exp(-H_B/kT)\} \tag{10.3.2}$$

式中T是热槽的温度。显然，

$$[H_B,\ \rho(T)]=i\hbar\partial_t\rho=0 \tag{10.3.3}$$

因为H_B是相互对易的各项之和，所以我们可以写成

$$\rho=\prod_i\rho_i(T)$$

其中

$$\rho_i(T)=\frac{\exp\left[-\frac{\hbar\omega_i}{kT}\left(\Gamma_i^{+}\Gamma_i+\frac{1}{2}\right)\right]}{\mathrm{Tr}\left\{\exp\left[-\frac{\hbar\omega_i}{kT}\left(\Gamma_i^{+}\Gamma_i+\frac{1}{2}\right)\right]\right\}} \tag{10.3.4}$$

进而定义

$$\begin{aligned}Z_i&=\mathrm{Tr}\left\{\exp\left[-\frac{\hbar\omega_i}{kT}\left(\Gamma_i^{+}\Gamma_i+\frac{1}{2}\right)\right]\right\}\\&=\sum_n{}'\exp\left\{-\frac{\left(n+\frac{1}{2}\right)\hbar\omega_i}{kT}\right\}\end{aligned}$$

即

$$\boxed{Z_i(T)=\frac{\exp(\hbar\omega_i/2kT)}{1-\exp(-\hbar\omega_i/kT)}} \tag{10.3.5}$$

平均粒子数$\langle n_i(T)\rangle$也是有用的，定义为

$$\langle n_i(T)\rangle=\langle\Gamma_i^{+}\Gamma_i\rangle=\sum_n{}'n\exp\left[-\frac{\left(n+\frac{1}{2}\right)\hbar\omega_i}{kT}\right]\Big/Z_i(T)$$

即

$$\boxed{\langle n_i(T)\rangle=[\exp(\hbar\omega_i/kT)-1]^{-1}} \tag{10.3.6}$$

供进一步应用的项是槽相关函数，为了得到这些函数，我们定义

$$\begin{aligned}\Gamma_i(t)&=\exp\left(\frac{iH_Bt}{\hbar}\right)\Gamma_i\exp\left(-\frac{iH_Bt}{\hbar}\right)\\&=\exp(-i\omega_it)\Gamma_i\end{aligned} \tag{10.3.7}$$

所以槽的相关为

$$\boxed{\begin{aligned}&\langle\Gamma_i^{+}(t)\Gamma_j\rangle=e^{i\omega_it}\langle\Gamma_i^{+}\Gamma_j\rangle=e^{i\omega_it}\delta_{ij}\langle n_i(T)\rangle\\&\langle\Gamma_i(t)\Gamma_j\rangle=\langle\Gamma_i^{+}(t)\Gamma_j^{+}\rangle=0\\&\langle\Gamma_i(t)\Gamma_j^{+}\rangle=e^{i\omega_it}\langle\Gamma_j^{+}\Gamma_i+\delta_{ij}\rangle\\&\qquad\qquad=e^{i\omega_it}\delta_{ij}[\langle n_i(T)\rangle+1]\end{aligned}} \tag{10.3.8}$$

10.3.2 槽算子平滑函数的相关

考虑变量

$$y(t)=\Sigma\delta_i\Gamma_i(t)e^{i\omega_0 t} \tag{10.3.9}$$

(因子 $\exp(i\omega_0 t)$是考虑到 10.4.2 节的应用而加进去的。)

$$\begin{aligned}&\langle y(t)\rangle=0\\&\langle y(t)y(0)\rangle=\langle y^+(t)y^+(0)\rangle=0\\&\langle y^+(t)y(0)\rangle=\sum_i|S_i|^2\langle n_i(T)\rangle\exp[i(\omega_i-\omega_0)t]\\&\langle y(t)y^+(0)\rangle=\sum_i|S_i|^2\langle n_i(T)+1\rangle\exp[-i(\omega_i-\omega_0)t]\end{aligned} \tag{10.3.10}$$

现假定模式 i 是相互紧密地配置着的，所以 ω_i 是 i 的平滑函数。

因此我们把这些变量如此联系起来：

$$S_i\to S(\omega_i)\to S(\omega) \tag{10.3.11}$$

$$\langle n_i(T)\rangle\to\langle n(\omega_i,T)\rangle\to\langle n(\omega,T)\rangle \tag{10.3.12}$$

以及

$$\sum_i|S_i|^2(\cdots)\to\int d\omega S(\omega)(\cdots) \tag{10.3.13}$$

对于充分平滑的函数$\langle n(\omega,T)\rangle$及$S(\omega)$，相关函数是 t 的指数地迅速衰减的函数。例如，考虑

$$\langle y^+(t)y(0)\rangle=\int_0^\infty d\omega S(\omega)\langle n(\omega,T)\rangle\exp[i(\omega,\omega_0)t] \tag{10.3.14}$$

如果象通常所使用的那样，$S(\omega)\langle n(\omega,T)\rangle$是 ω 的平滑函数，则它的傅里叶变换是 t 的一个迅速衰减函数。因此，相关函数(10.3.14)将是 t 的迅速衰减函数的形式乘上 $\exp(-i\omega_0 t)$。

这使我们想起相关函数与 1.4.2 节中所介绍的谱之间的关系。但是这里有 $\exp[i(\omega-\omega_0)t]$ 而不是 $\cos\omega t$，而且相关函数是复数。

10.3.3 与热槽发生相互作用的系统的量子主方程

现在我们将证明，当一个量子系统与热槽发生相互作用时，为了求得量子主方程，我们可以应用6.6节的绝热消去法。这里所用的拉普拉斯变换方法并非最有用的方法，但它给出了非常迅捷和非常精确的回答。

我们假定系统利用可以服从任意交换关系的算子 A、A^+ 来描述，而槽由谐振子算子 Γ_i、Γ_i^+ 来描述。

考虑哈密顿量可以写成

$$H = \gamma^2 H_1 + \gamma H_2 + H_3 \tag{10.3.15}$$

H_3 仅是 A、A^+ 的函数，H_2 可写成

$$H_2 = \sum_i (C g_i \Gamma_i^+ + C^+ g_i^* \Gamma_i) \tag{10.3.16}$$

式中 C 是 A 和 A^+ 的函数。这里 g_i 可以是时间的函数。H_1 是槽的哈密顿量，且具有(10.3.1)的形式，尽管不是绝对必要的。系统和槽的算子相对易，并且每个作用都发生在它们自己的空间。这样，如果设 H_2 为零，则

$$\rho = \rho_B \times \rho_S \tag{10.3.17}$$

及

$$\begin{aligned} i\hbar \frac{\partial \rho_B}{\partial t} &= \gamma^2 [H_1, \rho_B] \\ i\hbar \frac{\partial \rho_S}{\partial t} &= [H_3, \rho_S] \end{aligned} \tag{10.3.18}$$

并且严格地讲，我们应写成

$$\begin{aligned} A &= 1 \times \hat{A} \\ \Gamma_i &= \hat{\Gamma}_i \times 1 \end{aligned} \tag{10.3.19}$$

用以指明在它们自身空间中的每个运算。参数 γ 的引进是为了用公式表示这样的事实，即当槽变量比系统变量具有快得多的时标时，这种步骤是合适的，并且为了给出有限的极限，必须把 g_i 变大这一点也考虑到。这种极限在实际情况下的严格执行，有赖于变大的各适当变量的知识。

因而密度矩阵的运动方程可以写成

$$\frac{g\rho}{\partial t} = -\frac{i}{\hbar}[H_t, \rho], \tag{10.3.20}$$

式中

$$L_t\rho = -\frac{i}{\hbar}[H_t, \rho], \tag{10.3.21}$$

算子 L_t 称为刘维算子，它是线性算子。只要我们能定义一个适当的投影算子，现在的方程(10.3.20)对于我们应用绝热消去技巧是完全正确的形式。我们选取如下类型的投影算子：

$$\boxed{P_\rho = \rho(T) \times \mathrm{Tr}_B\{\rho\} \equiv \rho(T) \times \beta} \tag{10.3.22}$$

式中 $\rho(T)$ 由(10.3.2)定义，而所谓 Tr_B，我们的意思是仅对槽的状态求迹。这就是说，对系统和槽的状态全部集合可以写成

$\langle n_B, n_S\rangle$ 及

$$\mathrm{Tr}_B\{A\} = \sum_B \langle n_B, n_S | A | n_B, n_S\rangle \tag{10.3.23}$$

P 显然是投影算子，而且由于 H_1 只对槽的状态作运算，所以

$$L_1 P_\rho = -\frac{i}{\hbar}[H_t, \rho(T) \times \mathrm{Tr}_B\rho] = 0 \tag{10.3.24}$$

下式也成立

$$PL_1\rho = -\frac{i}{\hbar}\rho(T) \times \mathrm{Tr}_B\{[H_1, \rho]\} = 0 \tag{10.3.25}$$

因此要求

$$PL_1 = L_1 P = 0 \tag{10.3.26}$$

得到满足。

现在我们核查一下 $PL_2\rho$ 是否为零，即

$$-\frac{i}{\hbar}\rho(T) \times \mathrm{Tr}_B\{[H_2, \rho(T) \times \beta]\} \tag{10.3.27}$$

是什么。如果我们把 H_2 中的一个典型项，即 $C\Gamma_i^+$ 代入上式得

$$-\frac{i}{\hbar}\rho(T) \times \mathrm{Tr}_B\{\Gamma_i^+\rho(T) \times C\beta - \rho(T)\Gamma_i^+ \times \beta C\} \tag{10.3.28}$$

显然，

$$\mathrm{Tr}_B\{\Gamma_i^+\rho(T)\}=\mathrm{Tr}_B\{\rho(T)\Gamma_i^+\}=0 \tag{10.3.29}$$

类似地，所有其他项都为零。

因此，我们可以进行象 6.6 节中那样的绝热消去步骤，以得到

$$\frac{\partial}{\partial t}[\rho(T)\times\beta]=L_3\rho(T)\times\beta-PL_2L_1^{-1}L_2\rho(T)\times\beta \tag{10.3.30}$$

上式简化为

$$\boxed{\frac{\partial\rho}{\partial t}=L_3\beta-\mathrm{Tr}_B\{L_2L_1^{-1}L_2\rho(T)\times\beta\}} \tag{10.3.31}$$

这个式子现在可以明显地简化成一个简单形式，但是，我们首先必须处理 L_1^{-1}。在 6.6 节我们曾用过关系

$$L_1^{-1}(1-P)=\int_0^\infty e^{L_1t}dt(1-P) \tag{10.3.32}$$

该式的正确性取决于

$$\lim_{t\to\infty}e^{L_1t}(1-P)=0$$

的正确性，而上式来源于 L_1 具有负本征值的假设。

事实上，这么强的结果在这里并非必要——这里所需要的是，当 t 很大时，Γ_i，Γ_i^+ 对稳定密度算子 $\rho(T)$ 中槽的特殊迹为零，并且如果出现的特殊线性组合，即 $\Sigma g_i\Gamma_i^+$ 等具有系数 g_i，而 g_i 满足10.3.2节所讨论的 $S(\omega)$的平滑条件，则这种特殊迹为零。我们把对这个问题的详细验算留给读者并沿用(10.3.32)式。

我们注意，对任意算子 A，可以定义

$$A(t)\equiv e^{-L_1t}A=\exp(iH_1t/\hbar)A\exp(-iH_1t/\hbar) \tag{10.3.33}$$

上式是一个与(10.3.7)相一致的符号，由显导数和 L_1 的定义可以证明：

$$\frac{dA(t)}{dt}=\frac{i}{\hbar}[H_1,\ A(t)]=-L_1A(t) \tag{10.3.34}$$

现在我们可以进行计算：

$$-\mathrm{Tr}_B\{L_2L_1^{-1}L_2\rho(T)\times\beta\}$$

$$= \frac{1}{\hbar^2}\int_0^\infty \mathrm{Tr}_B\{[H_2, e^{-iH_1t/\hbar}[H_2, \rho(T)\times\beta]e^{iH_1t/\hbar}]\}dt$$

$$= \frac{1}{\hbar^2}\int_0^\infty \mathrm{Tr}_B\{[e^{-H_1t/\hbar}H_2e^{-iH_1t/\hbar}, [H_2, \rho(T)\times\beta]]\}dt \tag{10.3.35}$$

$$= \frac{1}{\hbar^2}\int_0^\infty \mathrm{Tr}_B\{\sum_{i,j}[Cg_i\Gamma_i^+(t) + C^+g_i^*\Gamma_i(t), [Cg_j\Gamma_j^+ + C^+g_j^*\Gamma_j, \rho(T)\times\beta]]\}dt$$

在上式中有 16 项，只有那些含有 Γ 和 Γ^+ 的项将不为零。考虑特殊项

$$\frac{1}{\hbar^2}\int_0^\infty \mathrm{Tr}_B\{\sum_{i,j} Cg_i\Gamma_i^+(t)C^+g_j^*\Gamma_j\rho(T)\times\beta\}dt$$

利用(10.3.9)的符号，$g_i^* \to S_i$，$\omega_0 \to 0$，借助

$$y(t) = \Sigma g_i^*\Gamma_j(t) \tag{10.3.36}$$

我们可以写出上面的式子。我们引进类似于量子光学中常用的符号，即

$$\begin{aligned}
&\frac{1}{\hbar^2}\int_0^\infty dt\langle y^+(t)y(0)\rangle = \frac{1}{2}\bar{N}K + i\delta\\
&\frac{1}{\hbar^2}\int_0^\infty dt\langle y^+(0)y(t)\rangle = \frac{1}{2}K\bar{N} - i\delta\\
&\frac{1}{\hbar^2}\int_0^\infty dt\langle y(t)y^+(0)\rangle = \frac{1}{2}K(\bar{N}+1) - i\delta'\\
&\frac{1}{\hbar^2}\int_0^\infty dt\langle y(0)y^+(t)\rangle = \frac{1}{2}K(\bar{N}+1) + i\delta'
\end{aligned} \tag{10.3.37}$$

双换位子可以算出，我们发现主方程简化为

$$\begin{aligned}
\frac{\partial\beta}{\partial t} = &-\frac{i}{\hbar}[H_s, \beta] - i\delta[C^+C, \beta] + i\delta'[CC^+, \beta]\\
&+\frac{1}{2}K(H\bar{N})(2C\beta C^+ - C^+C\beta - \beta C^+C)\\
&+\frac{1}{2}K\bar{N}(2C^+\beta C - CC^+\beta - \beta CC^+) \equiv \hat{L}\beta
\end{aligned} \tag{10.3.38}$$

注解

i) 上面这种形式的主方程，通常不用拉普拉斯变换的方法而使用实质上是时域方程来导出。例如，请看罗塞耳 (Louisell)的著作[10.5]。我们的推导依赖于 L_3 比 $\gamma^2 L_1$ 小得多，而实际上，在光的情况下并非如此，因为 H_3 表示由于与热槽相互作用而衰减很慢的高速振荡。这通常可以用引入相互作用图景而得到重视，结果是(10.3.38)形式的主方程，其中出现了 g_i 明显的时间依赖关系，而 (10.3.9) 中的 ω_0 是自由振动的固有频率。这个方法在10.4.2节我们处理双能级原子时得到证实。这里对这问题作全面而详细的叙述并非我们的目的，关于完整的解释建议读者参考罗塞尔的著作。

ii) 主方程(10.3.38) 可以看作是量子马尔可夫过程的定义，该过程从数学的观点来看是比较简单的。表示自由振动和热槽相互作用的量 H_3 和 C 是非常任意的，就象它们的相互关系那样。但是，应注意，上述定义是不完全的，它没有详细说明如何确定多时平均，换言之，它没有详细说明如何去确定作为普通随机过程的基础的多时联合概率的相似物。

iii) 在温度为零的热槽中，$\rho_B = |0\rangle\langle 0|$，$\Gamma^+\Gamma$ 的平均全部为零。这意味着

$$T = 0 \Longrightarrow \bar{N} = 0 \tag{10.3.39}$$

而第三行没有贡献。因此，这一行与热噪声有关。

10.4 量子马尔可夫过程的例子及应用

10.4.1 谐振子

这里我们取

$$C \to a$$

$$C^+ \to a^+$$

$$[a,\ a^+] = 1 \tag{10.4.1}$$

$$H_3 = \hbar\omega\left(a^+a + \frac{1}{2}\right)$$

利用交换关系，我们发现(10.3.38)简化成（现在我们简单地把 $\dot{\rho}$ 写成 ρ。）

$$\frac{\partial\rho}{\partial t}=i\omega'[a^+a,\ \rho]$$
$$+\frac{1}{2}K(\bar{N}+1)(2a\rho a^+-a^+a\rho-\rho a^+a)$$
$$+\frac{1}{2}K\bar{N}(2a^+\rho a-aa^+\rho-\rho aa^+) \qquad (10.4.2)$$

式中

$$\omega'=\omega+\delta-\delta'$$

a) 对角矩阵元

对角矩阵元

$$\langle n|\rho|n\rangle\equiv P(n) \qquad (10.4.3)$$

代表系统中存在 n 个量子的概率。我们容易验证这点（利用 10.1 节中定义的 a^+ 和 a 的性质）。

$$\partial_t P(n)=K(\bar{N}+1)[(n+1)P(n+1)-nP(n)]$$
$$+K\bar{N}[nP(n-1)-(n+1)P(n)] \qquad (10.4.4)$$

这是一个普通的生灭主方程。注意，转移概率具有

$$t^+(n)=K\bar{N}(n+1)$$
$$t^-(n)=K(\bar{N}+1)n \qquad (10.4.5)$$

的形式。所以产生一个量子的概率具有与 $(n+1)$ 成正比的部分。形式为

$$B\rightleftharpoons X$$
$$A+X\longrightarrow 2X \qquad (10.4.6)$$

的化学反应将具有类似的主方程。

该方程在稳定态的解为

$$P_s(n)=\left(\frac{\bar{N}}{1+\bar{N}}\right)^n\frac{1}{1+\bar{N}} \qquad (10.4.7)$$

这是有用的玻耳兹曼分布，式中我们可以验明

$$\frac{\bar{N}}{1+\bar{N}}=\exp(-\hbar\omega/KT) \qquad (10.4.8)$$

它意味着

$$\bar{N} = 1/[\exp(\hbar\omega/KT') - 1] \tag{10.4.9}$$

上式中 $\bar{N}$ 决定于 T 或者相反，我们注意

$$\langle n\rangle_s = \bar{N}$$

$$\text{变量}\langle n\rangle_s = \bar{N}^2 \tag{10.4.10}$$

b) 从 P 表示得出福克-普朗克方程

我们利用10.2.2节格劳伯-苏达香的 P 表示及10.2.3节式(10.2.26)的算子相关，记住写在 ρ 右面的算子乘积，当形成相关时次序发生颠倒。我们导出关于 P 函数的福克-普朗克方程：

$$\boxed{\begin{aligned}\frac{\partial P}{\partial t} = \Big[&\frac{1}{2}K\Big(\frac{\partial}{\partial\alpha}\alpha + \frac{\partial}{\partial\alpha^*}\alpha^*\Big)\\ &- i\omega'\Big(\frac{\partial}{\partial\alpha}\alpha - \frac{\partial}{\partial\alpha^*}\alpha^*\Big) + K\bar{N}\frac{\partial^2}{\partial\alpha\partial\alpha^*}\Big]P\end{aligned}} \tag{10.4.11}$$

这是一个奥恩斯坦-乌伦贝克过程，所以我们可以写成

$$\alpha = x + iy$$

及

$$\partial_\alpha = \frac{1}{2}(\partial_x - i\partial_y)$$

并得到

$$\boxed{\begin{aligned}\frac{\partial P}{\partial t} = \frac{1}{2}K\Big[&\Big(\frac{\partial}{\partial x}x + \frac{\partial}{\partial y}y\Big)\\ &+ \omega'\Big(\frac{\partial}{\partial x}y - \frac{\partial}{\partial y}x\Big)\\ &+ \frac{\bar{N}}{2}\Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\Big)\Big]P\end{aligned}} \tag{10.4.12}$$

与总是包含非常数扩散项的 P 表示福克-普朗克方程相比，这个结果是非常简单的。与(10.4.12)等价的随机微分方程为

$$\boxed{\begin{aligned} dx &= -\left(\frac{1}{2}Kx+\omega y\right)dt+\sqrt{\frac{1}{2}\bar{N}}dW_1(t)\\ dy &= -\left(\frac{1}{2}Ky-\omega x\right)dt+\sqrt{\frac{1}{2}\bar{N}}dW_2(t)\end{aligned}} \tag{10.4.13}$$

这些方程表示一个阻尼振子的方程，与 5.3.6d 节所讨论的那些方程非常相似。

这些方程常写成一个复朗之万方程：

$$d\alpha=-\left(\frac{1}{2}K-i\omega\right)\alpha\,dt+\sqrt{\bar{N}}d\eta(t) \tag{10.4.14}$$

式中 $d\eta(t)$ 是复维纳过程的增量，它满足

$$\begin{aligned}&\langle d\eta(t)\rangle=\langle d\eta^*(t)\rangle=\langle d\eta(t)d\eta(t')\rangle\\&\qquad=\langle d\eta^*(t)d\eta^*(t')\rangle=0\\&\langle d\eta(t)d\eta^*(t)\rangle=dt\end{aligned} \tag{10.4.15}$$

并由下式清楚地给出：

$$d\eta(t)=[dW_1(t)+idW_2(t)]/\sqrt{2} \tag{10.4.16}$$

复数 $d\eta(t)$ 的使用很不容易被推广，所以以后我们将不用它们。但是，复变量的福克-普朗克方程是很有用的，并将继续使用它。

奥恩斯坦-乌伦贝克过程的解由 4.4 节给出。我们发现

$$\begin{aligned}&\langle\alpha(t)\rangle=\alpha(0)\exp[-(K/2+i\omega)t]\\&\langle\alpha^*(t)\alpha(t)\rangle=\langle\alpha^*(0)\alpha(0)\rangle e^{-Kt}+\bar{N}(1-e^{-Kt})\end{aligned} \tag{10.4.17}$$

$$\begin{aligned}&\langle\alpha^2\rangle_s=\langle\alpha^{*2}\rangle_s=0\\&\langle\alpha\alpha^*\rangle_s=\langle a^+a\rangle_s=\bar{N}\end{aligned} \tag{10.4.18}$$

仅当 $\bar{N}=0$ 时我们才得到 $\langle a^+a\rangle_s$ 为零，即温度为零度时的情况。奥恩斯坦-乌伦贝克过程的时间相关函数要求作 10.5 节所做的那种解释。

c) 包含驱动场的情况

假定我们讨论经选取

$$H_3=\hbar\omega\left(a^+a+\frac{1}{2}\right)+\hbar(\varepsilon a^++\varepsilon^*a) \tag{10.4.19}$$

所得到的驱动阻尼谐振子，上式表示与非量子化场(10.1.1 节) 相

互作用的一个量子化振子。

福克-普朗克方程发展出一个额外项

$$\left(i\varepsilon\frac{\partial}{\partial\alpha}-i\varepsilon^*\frac{\partial}{\partial\alpha^*}\right)P \tag{10.4.20}$$

它允许福克-普朗克方程写成原来的形式。

$$\begin{aligned}\frac{\partial p}{\partial t}=&\left[\left(\frac{1}{2}K-i\omega'\right)\frac{\partial}{\partial\alpha}\left(\alpha+\frac{i\varepsilon}{\frac{1}{2}K-i\omega'}\right)\right.\\&+\left(\frac{1}{2}K+i\omega'\right)\frac{\partial}{\partial\alpha^*}\left(\alpha^*-\frac{i\varepsilon}{\frac{1}{2}K+i\omega'}\right)\\&\left.+K\bar{N}\frac{\partial^2}{\partial\alpha\partial\alpha^*}\right]P\end{aligned} \tag{10.4.21}$$

上式仍表示奥恩斯坦-乌伦贝克过程，但是原点作了移动，以致

$$\langle\alpha\rangle_s=i\varepsilon\Big/\left(\frac{1}{2}K-i\omega'\right)=\langle\alpha\rangle_s \tag{10.4.22}$$

$$\langle\partial\alpha^*\rangle-\langle\alpha\rangle\langle\alpha^*\rangle=\bar{N}$$

$$\langle\alpha^2\rangle-\langle\alpha\rangle^2=\langle\alpha^{*2}\rangle-\langle\alpha^*\rangle^2 \tag{10.4.23}$$

而均值的时间相依解为

$$\begin{aligned}\langle\alpha(t)\rangle=&\langle\alpha(0)\rangle\exp\left[-\left(\frac{K}{2}-i\omega'\right)t\right]\\&-\frac{i\varepsilon}{\frac{1}{2}K+i\omega'}\left\{1-\exp\left[-\left(\frac{K}{2}-i\omega'\right)t\right]\right\}\end{aligned} \tag{10.4.24}$$

在P表示下的稳态分布是含有 (10.4.22) 给出的变量$\bar{N}$及平均值的高斯分布。如果$\bar{N}$很小，该式表示几乎与纯相干态 $|\alpha_s\rangle\langle\alpha_s|$ 的密度矩阵相对应的密度矩阵 ρ。于是，一个对相干态的很好近似是由经典驱动场提供的，该场与谐振子相互作用，而谐振子又与低温热槽发生相互作用。

d) 涨落场的驱动

假定(10.4.19)式中的 ε 是时间的函数 $\varepsilon(t)$，则很可能是随机的。α 和 α^4 的朗之万方程为

$$d\alpha(t)=\left[-\left(\frac{1}{2}K-i\omega'\right)\alpha(t)-i\varepsilon(t)\right]dt$$
$$+\sqrt{\frac{K\bar{N}}{2}}[dW_1(t)+idW_2(t)] \qquad (10.4.25)$$

由于 $\varepsilon(t)$ 线性地出现，并且被常数相乘，所以我们可以作出某种简单的随机假设。

最简单的假设是采用短相关时间极限，并用

$$\varepsilon(t)dt=\varepsilon_0 dt+\sqrt{\frac{f}{2}}[dW_3(t)+idW_4(t)] \qquad (10.4.26)$$

代替，上式给出每个分量中的独立涨落[由于 $\varepsilon(t)$ 的系数是常数，所以不存在伊藤-斯特拉托诺维奇分歧]，净效应是由

$$K\bar{N}\longrightarrow K\bar{N}+f$$

的代替修改了(10.4.21)式，也就是说额外的噪声只简单地使已经存在的热噪声得到了加强。

最后，关于 $\varepsilon(t)$ 的运动方程可以被采用，并连同随机微分方程(10.4.25)一起求解。

10.4.2 被驱动的双能级原子

原子系统的理想情况是双能级原子组成的系统，这些原子可以存在于两个状态中的一个。我们建立具有下述矩阵的矩阵公式

$$|-\rangle=\begin{bmatrix}0\\1\end{bmatrix} \quad \text{基态}$$
$$|+\rangle=\begin{bmatrix}1\\0\end{bmatrix} \quad \text{激发态} \qquad (10.4.27)$$

系统可以构造跃迁，这种跃迁由泡利(Pauli)矩阵

$$S^+=\begin{bmatrix}0&1\\0&0\end{bmatrix},\ S^-=\begin{bmatrix}0&0\\1&0\end{bmatrix},\ S_z=\begin{bmatrix}1&0\\0&-1\end{bmatrix} \qquad (10.4.28)$$

来描述。

于是，S^+ 促使原子上升至激发态，S^- 使原子下落至基态。与

辐射场和驱动场 $E(t)$ 相互作用的系统，其哈密顿量可以写成

$$H_3 = \frac{1}{2}\hbar\omega_0 S_z - \frac{1}{2}\hbar d[S^+E(t) + S^-E^*(t)]$$

$$H_2 = \sum_i [g_i S^+ a_i + g_i^* S^- a_i^+] \tag{10.4.29}$$

$$H_1 = \sum_i \hbar\omega_i \left(a_i^+ a_i + \frac{1}{2}\right)$$

在物理上，项 $\frac{1}{2}\hbar\omega_0 S_z$ 规定激发态具有能量 $\frac{1}{2}\hbar\omega_0$，规定较低能态具有能量 $-\frac{1}{2}\hbar\omega_0$，于是激发能为 $\hbar\omega_0$。H_3(*) 中的第二项表示通过驱动场的偶极矩 d 对原子的耦合，实质上它与 H_2 中的项是相同的，但是在 H_2 中的其他模式 i 是量子化的。模式 i 被认为代表向用 i 标示的各不同方向跃迁的、且能量为 $\hbar\omega_0$ 的光子，而驱动场是一个强经典场，它处在某个特殊的方向。

辐射场量子形成热槽，于是我们用 a_i 表示 Γ_i。原则上，热槽可以处在任何温度，但是在实际实验情况下，通常处在零度。我们将保持非零温度 T。

假定驱动场 $E(t)$ 与激发原子所要求的频率相同，为 ω_0，于是我们写成

$$E(t) = \varepsilon e^{-i\omega_0 t} \tag{10.4.30}$$

这里最好在消去了快速振荡的相互作用图景中工作。定义

$$\delta(t) = \exp\left(\frac{i\omega_0}{2} S_z t\right)\rho \exp\left(-\frac{i\omega_0}{2} S_z t\right)^{(**)} \tag{10.4.31}$$

因此，我们可以用下述哈密顿量来描述 $\tilde{\rho}(t)$ 的运动，

$$\tilde{H}_3 = -\frac{1}{2}\hbar d(S^+\varepsilon + \delta^-\varepsilon^*)$$

$$\tilde{H}_2 = \sum_i (g_i e^{i\omega_0 t} S^+ a_i + g_i^* e^{-i\omega_0 t} S^- a_i^+) \tag{10.4.32}$$

$$\tilde{H}_1 = \sum_i \hbar\omega_i \left(a_i^+ a_i + \frac{1}{2}\right)$$

(*) 原文将 H_2 和 H_3 印错。——译注

(**) 为了与后面一致，这里的 ρ 应为 $\hat{\rho}$。——译注

读者可以核实一下，所有的推导与10.33节中所用的方法几乎完全相同，所不同的是作了代换

$$g_i \to g_i e^{i\omega t} \tag{10.4.33}$$

这意味着现在的 $g(t)$ 与(10.3.9)中定义的完全相同，因此，这里明显的指数时间依赖关系是由原子的自由运动引起的。

对 β 来说，主方程变成简化的密度矩阵(在相互作用图景中)，使用了相应于显式矩阵(10.4.28)的代数后，主方程为：

$$\begin{aligned}\frac{\partial \beta}{\partial t} &= \frac{1}{2} i d[S^+ \varepsilon^* + S^- \varepsilon, \beta] \\ &+ \frac{1}{2} K(1+\bar{N})\Big[2S^- \beta S^+ - \frac{1}{2}(1+S_z)\beta \\ &- \frac{1}{2}\beta(1+S_z)\Big] \\ &+ \frac{1}{2} K\bar{N}\Big[2S^+ \beta S^- - \frac{1}{2}(1-S_z)\beta \\ &- \frac{1}{2}\beta(1-S_z)\Big]\end{aligned} \tag{10.4.34}$$

由于 S^+、S_z 仅在二维空间中起作用，并且不是谐振子算子，所以这里我们不能使用相干态。密度矩阵是迹为1的 2×2 矩阵，并且完全由量 $\mathrm{Tr}\{\rho S^\pm\}$、$\mathrm{Tr}\{\rho S_z\}$ 规定，这两个量是算子的期待值，因此，为了对 $\rho(t)$ 作完全而详细的说明，我们导出

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle S^+\rangle &= -\frac{1}{2} K(2\bar{N}+1)\langle S^+\rangle + id\varepsilon\langle S_z\rangle \\ \frac{\mathrm{d}}{\mathrm{d}t}\langle S^-\rangle &= -\frac{1}{2} K(2\bar{N}+1)\langle S^-\rangle - id\varepsilon^*\langle S_z\rangle \\ \frac{\mathrm{d}}{\mathrm{d}t}\langle S_z\rangle &= -K(2\bar{N}+1)\langle S_z\rangle - K \\ &+ \frac{1}{2} id(\varepsilon^*\langle S^+\rangle - \varepsilon\langle S^-\rangle)\end{aligned} \tag{10.4.35}$$

这些方程已经被卡万克尔(Carmichael, H. T.)和沃尔斯(Walls, D. E.)[10.6]作过详细的研究。

我们注意几点：

a) 具有 $\varepsilon = 0$ 的稳态解

这里

$$\langle S^+\rangle = \langle S^-\rangle = 0$$

$$\langle S_z\rangle = -\frac{1}{2\bar{N}+1} \tag{10.4.36}$$

设 $\bar{N} = (e^{\hbar\omega/KT}-1)^{-1}$ 给定了正确的玻耳兹曼分布。

b) 与温度有关的弛豫

与谐振子相比，弛豫率跟依赖于温度的量 $(2\bar{N}+1)$ 成正比。算子代数中的差导致了这个结果。

c) 与槽的温度相比较

如果我们利用(10.3.12)式和有关公式，直接求出量 K、$\bar{N}$、δ 和 δ'，得到

$$\frac{1}{2}K\bar{N} + i\delta = \hbar^2\int_0 dt\, e^{-i\omega t}\int_0^\infty d\omega S(\omega)\langle n(\omega, T)\rangle e^{i\omega t} \tag{10.4.37}$$

我们注意，恒等式

$$\int d\Omega \int_0^\infty f(\Omega)e^{\pm i\Omega t} = \pi f(0) \pm i\int d\Omega\, f(\Omega)/\Omega \tag{10.4.38}$$

式中$\int$表示包含 $1/\Omega$ 的积分的柯西主值。因此，

$$\frac{1}{2}K\bar{N} + i\delta = \pi\hbar^2 S(\omega_0)\langle n(\omega_0, T)\rangle + i\hbar\int_0 d\omega S(\omega)\langle n(\omega, T)\rangle(\omega-\omega_0)^{-1} \tag{10.4.39}$$

类似地，

$$\frac{1}{2}K(\bar{N}+1) + i\delta' = \pi\hbar^2 S(\omega_0)\langle n(\omega_0, T)+1\rangle + i\hbar^2\int_0^0 d\omega S(\omega)\langle n(\omega, T)+1\rangle(\omega-\omega_0)^{-1} \tag{10.4.40}$$

该式是对阻尼常数更有用的公式。

由这些方程，我们注意到

$$K = 2\pi\hbar^2 S(\omega_0) \tag{10.4.41}$$

$$\bar{N} = \langle n(\omega_0, T)\rangle = [\exp(\hbar\omega_0/KT) - 1]^{-1} \tag{10.4.42}$$

$$-\delta + \delta' = \hbar^2 \int_0^\infty d\omega S(\omega)(\omega - \omega_0)^{-1} \tag{10.4.43}$$

注意，(10.4.42)与(10.4.36)相一致，也即原子与槽达到相同的温度。

d) 非零场的稳态平均

我们可以解(10.4.35)而得到

$$\langle S_z\rangle = \frac{-K^2(2\bar{N}+1)}{K^2(3\bar{N}+1)^2 + 2d^2|\varepsilon|^2} \tag{10.4.44}$$

$$\langle S^+\rangle = \langle S^-\rangle^* = \frac{-2id\bar{K}\varepsilon}{K^2(2\bar{N}+1)^2 + 2d^2|\varepsilon|^2} \tag{10.4.45}$$

e) 与随机电报过程的关系：爱因斯坦方程

当 $E(t)=0$ 时，我们可以定义

$$P(+) = \langle + |\rho| + \rangle$$

$$P(-) = \langle - |\rho| - \rangle \tag{10.4.46}$$

并导出关于 $P(+)$和 $P(-)$的主方程：

$$\begin{aligned} \partial_t P(+) &= -K(1+\bar{N})P(+) + K\bar{N}P(-) \\ \partial_t P(-) &= K(1+\bar{N})P(+) - K\bar{N}P(-) \end{aligned} \tag{10.4.47}$$

这些都是 3.8.5 节中随机电报过程方程。因此，双能级原子可以看成是一个量子随机过程。但是，如果 $E(t)\neq 0$，则非对角矩阵元也存在，它们给出了真实的量子过程。

方程 (10.4.47) 首先由爱因斯坦引进[10.7]，所以现在称为爱因斯坦方程。它们显示二种效应，即

i) 受激发射和吸收：与 $\bar{N}$ 成正比的项依赖于辐射场中光子的数目，它们称为受激项。因此，激发过程(−→+)及退激过程(+→−)与光子的吸收和发射相对应，它们可以作为被已经在场中的光子数的激发过程。

ii) 自发发射：即退激(+→−)，它是因两个方程中存在项

$KP(+)$ 而出现的. 即使不存在光子，退激也可以出现。从现代观点来看，这并不是令人惊奇的，但是当初引进它们时，这是一个重要的创新。

10.5 量子马尔可夫过程中的时间相关函数

对一个量子系统，仅对相关函数$\langle A(t')B(t)\rangle$可以很容易地精确定义，这里我们简洁地说出其结果。如果 $\rho(t)$ 是时刻 t（在薛定谔表象中）的密度矩阵，H是哈密顿量，A 和 B 是被测变量的算子，则

$$\langle A(t)\rangle = \mathrm{Tr}\{A\rho(t)\rangle \tag{10.5.1}$$

及

$$\langle A(t+\tau)B(t)\rangle = \mathrm{Tr}\{\mathrm{e}^{iH\tau/\hbar}A\mathrm{e}^{-iH\tau/\hbar}B\rho(t)\} \tag{10.5.2}$$

[式(10.5.2)的事实来自海森伯图景的形式，在海森伯图景中该式变为

$$\langle A(t+\tau)B(t)\rangle = \mathrm{Tr}\{A_{\mathrm{H}}(t+\tau)B_{\mathrm{H}}(t)\rho_{\mathrm{H}}\} \tag{10.5.3}$$

式中海森伯密度矩阵 ρ_{H} 是与时间无关的.]

方程(10.5.2)虽是精确的，但是没有用处。在量子马尔可夫系统中，我们希望用简化系统的刘维方程来表示每个事物，在这种简化系统中热槽的变量已经求出。当这样做了以后，我们就有效地定义了多时联合概率，它可能使我们全面地详细说明量子马尔可夫过程。

这个目的可以相对简单地达到。设A和B仅是系统空间中而不是热槽空间中的算子，因而我们可以把(10.5.3)式重新写成

$$\langle A(t+\tau)B(t)\rangle = \mathrm{Tr}_{\mathrm{s}}\{A\mathrm{Tr}_{\mathrm{B}}\{\mathrm{e}^{-iH\tau/\hbar}B\rho(t)\mathrm{e}^{iH\tau/\hbar}\}\} \tag{10.5.4}$$

(由于A与槽空间中的单位元成正比)。

关于项

$$X(\tau,\ t) = \mathrm{e}^{-iH\tau/\hbar}B\rho(t)\mathrm{e}^{iH\tau/\hbar} \tag{10.5.5}$$

对 τ 的运动方程为

$$i\hbar\partial_\tau X(\tau,\ t) = [H,\ X(\tau,\ t)] \tag{10.5.6}$$

并用导出(10.3.38)的完全相同方法，可以把上式缩写成

$$\partial_t \rho(t) = \hat{L}\rho(t) \tag{10.5.7}$$

由于

$$\rho = \mathrm{Tr}_B\{\rho\} \tag{10.5.8}$$

所以我们可导出

$$\partial_\tau[\mathrm{Tr}_B\{X(\tau, t)\}] = \hat{L}[\mathrm{Tr}_B\{X(\tau, t)\}] \tag{10.5.9}$$

因此在绝热近似下，(10.5.4)式等价于

$$\boxed{\langle A(t+\tau)B(t)\rangle = \mathrm{Tr}_B\{A e^{\hat{L}\tau} B\rho(t)\}} \tag{10.5.10}$$

用类似的步骤，我们可以求出普遍的多时刻相关，即

$$\boxed{\langle A(t_3)B(t_2)C(t_1)\rangle = \mathrm{Tr}_s\{A e^{\hat{L}(t_3-t_2)} B e^{\hat{L}(t_2-t_1)} C\rho(t_1)\}} \tag{10.5.11}$$

(在上面两个方程中$\hat{L}$作用于它右边的每个算子) 等等。

由于各算子并不作交换，所以时间顺序不必与算子的次序相同。这就允许存在一个宽广的可能性范围，但不是所有的可能性都给出简单的公式。的确出现过特殊情况，其形式为

$$\begin{aligned}&\langle A(t)B(t+\tau)C(t+\tau)D(t)\rangle\\ &= \mathrm{Tr}\{A e^{iH\tau/\hbar} BC e^{-iH\tau/\hbar} D\rho(t)\}\\ &= \mathrm{Tr}_s\{BC\mathrm{Tr}_B\{e^{-iH\tau/\hbar} D\rho(t) A^{iH\tau/\hbar}\}\}\end{aligned}$$

以致

$$\langle A(t)B(t+\tau)D(t)\rangle = \mathrm{Tr}_s\{BC e^{\hat{L}\tau} D\rho(t) A\} \tag{10.5.12}$$

注意(10.5.10)是上式中设 $C(t) = D(t) = 1$ 的特殊情况，我们得

$$\boxed{\langle A(t)B(t+\tau)\rangle = \mathrm{Tr}_s\{B e^{\hat{L}\tau}\rho(t) A\}} \tag{10.5.13}$$

所有可能算子的相关函数的完备组（系）定义了所有可能的联合平均，因而可以把它取为一个量子系统的马尔可夫性质的定义。在该量子系统中 $\hat{L}$ 可以是维持 $\mathrm{Tr}_s\{\rho\}$ 以及 ρ 的正定性的任意算子。尤其象(10.3.8)所定义的那样，$\hat{L}$ 可以用作、而且实际上也是唯一一类确实存在的算子。

10.5.1 量子回归定理

在线性方程被发现是可行手段的情况下，我们可以发展一个量子回归定理，它与普通马尔可夫过程(3.7.4节)的回归定理相类似。这个结果首先是由拉克斯(Lax, M.)导出的[10.8]。

对某组算子 Y_i，假定可以证明对任意的初条件 β，主方程会产生出

$$\boxed{\partial_t\langle Y_i(t)\rangle = \Sigma G_{ij}(t)\langle Y_j(t)\rangle} \tag{10.5.14}$$

则我们断言

$$\boxed{\partial_\tau\langle Y_i(t+\tau)Y_j(t)\rangle = \sum_{ij} G_{ij}(\tau)\langle Y_j(t+\tau)Y_j(t)\rangle} \tag{10.5.15}$$

这是因为

$$\langle Y_i(t+\tau)Y_j(t)\rangle = \mathrm{Tr}_s\{Y_i e^{\hat{L}\tau}Y_i\beta(t)\} \tag{10.5.16}$$

所以上式右面部分是选取初始密度矩阵

$$\rho_{初} = Y_i\rho(t) \tag{10.5.17}$$

后 Y_i 在 $t+\tau$ 时刻的平均值。由于假定任意初始 β 都是允许的，而且假设是线性的，所以实际上我们可毫无问题地产生出任何初条件(可能认为 ρ 是厄密和半正定的要求会限止合适的初条件，但这是不会的。例如，在二维情况下，存在4个迹为1的线性独立的、半正定的矩阵，即

$$\begin{bmatrix}1&0\\0&0\end{bmatrix},\ \begin{bmatrix}0&0\\0&1\end{bmatrix},\ \begin{bmatrix}\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}\end{bmatrix},\ \begin{bmatrix}\frac{1}{2}&\frac{1}{2}\\-\frac{1}{2}i,&\frac{1}{2}\end{bmatrix} \tag{10.5.18}$$

这些矩阵提供了所有 2×2 矩阵的空间之基。这种性质可以直接推广到任意维空间)。因此，若像 (10.5.17) 所定义的那样选取 $\rho_{初}$，则假设(10.5.14)就产生(10.5.15)的结果，(10.5.15) 就是量子回归定理。

10.5.2 P 表示对谐振子的应用

我们先讨论常序时间相关函数

$$\langle a^+(t+\tau)a(t)\rangle = \mathrm{Tr}\left\{a^+ e^{\hat{L}\tau} a \int \mathrm{d}^2\alpha P(\alpha,t)|\alpha\rangle\langle\alpha|\right\}$$

$$= \mathrm{Tr}\left\{a^+ e^{\hat{L}\tau}\int \mathrm{d}^2\alpha\alpha P(\alpha,\ t)|\alpha\rangle\langle\alpha|\right\} \tag{10.5.19}$$

项

$$e^{\hat{L}\tau}\int \mathrm{d}^2\alpha\alpha P(\alpha,\ t)|\alpha\rangle\langle\alpha|$$

是推广后的主方程的解，此方程的初条件是P表示密度为$\alpha P(\alpha,t)$的算子。如果存在相应于$\hat{L}$的福克-普朗克方程[即(10.4.11)、(10.4.12)]，则该量是具有由

$$\int \mathrm{d}^2\alpha'\alpha' P(\alpha,\ t+\tau|\alpha',\ t)P(\alpha',\ t) \tag{10.5.20}$$

给出的P函数的密度矩阵，所以，利用迹的循环性质及$\langle\alpha|a^+ = \alpha^*\langle\alpha|$，得

$$\langle a^+(t+\tau)a(t)\rangle = \mathrm{Tr}\left\{a^+\int \mathrm{d}^2\alpha\mathrm{d}^2\alpha'\alpha' P(\alpha,\ t+\tau|\alpha',\ t)\right.$$

$$\left.\times P(\alpha',\ t)|\alpha\rangle\langle\alpha|\right\}$$

$$= \int \mathrm{d}^2\alpha\mathrm{d}^2\alpha'\alpha^*\alpha' P(\alpha,\ t+\tau;\alpha',\ t)$$

这意味着

$$\boxed{\langle a^+(t+\tau)a(t)\rangle = \langle\alpha^*(t+\tau)\alpha(t)\rangle} \tag{10.5.21}$$

式中$\alpha(t)$是P表示随机变量。类似地，对任何a^+及a的函数，下式成立

$$\boxed{\begin{aligned}&\langle F_1[a^+(t+\tau)]F_2[a(t)]\rangle\\ &\quad = \langle F_1[\alpha^*(t+\tau)]F_2[\alpha(t)]\rangle\end{aligned}} \tag{10.5.22}$$

对于非常序乘积，答案并非如此简单。

例如，考虑

$$\langle a^+(t+\tau)a(t+\tau)a^+(t)a(t)\rangle$$

$$=\mathbf{Tr}\{a^+ae^{\hat{L}\tau}a^+a\int d^2\alpha P(\alpha,\ t)\,|\alpha\rangle\langle\alpha|\} \tag{10.5.23}$$

利用 10.2.3 节(10.2.26)式的算子恒等式，上式

$$=\mathbf{Tr}\left\{a^+ae^{\hat{L}\tau}\int d^2\alpha\left[\left(\alpha^*-\frac{\partial}{\partial\alpha}\right)\alpha P(\alpha,\ t)\right]|\alpha\rangle\langle\alpha|\right\} \tag{10.5.24}$$

继续进行与上面相同的论证，上式

$$=\mathbf{Tr}\left\{a^+a\int d^2\alpha\int d^2\alpha' P(\alpha,\ t+\tau|\alpha',\ t)\right.$$

$$\left.\times\left[\left(\alpha^*-\frac{\partial}{\partial\alpha'}\right)\alpha' P(\alpha',\ t)\right]|\alpha\rangle\langle\alpha|\right\} \tag{10.5.25}$$

最后，利用同样的算子恒等式，上式

$$=\int d^2\alpha\ d^2\alpha'\left[\left(\alpha^*-\frac{\partial}{\partial\alpha}\right)\alpha P(\alpha,\ t+\tau|\alpha',\ t)\right]$$

$$\times\left[\left(\alpha^*-\frac{\partial}{\partial\alpha'}\right)\alpha' P(\alpha',\ t)\right] \tag{10.5.26}$$

为简单起见，我们可以除去 $\frac{\partial}{\partial\alpha}$，因为对它的积分为零，并对 $\frac{\partial}{\partial\alpha'}$ 取分部积分，以致

$$(10.5.26)=\int d^2\alpha d^2\alpha'\alpha'(\alpha^*\alpha')P(\alpha,\ t+\tau;\ \alpha',\ t) \tag{10.5.27}$$

$$+\int d^2\alpha'\ P(\alpha',\ t)\alpha'$$

$$\times\frac{\partial}{\partial\alpha'}\left[\int d^2\alpha(\alpha^*\alpha)P(\alpha,\ t+\tau|\alpha',\ t)\right] \tag{10.5.28}$$

上式可以写成

$$\boxed{\begin{aligned}&\langle Q^+(t+\tau)a(t+\tau)a^+(t)a(t)\rangle\\&\quad=\langle|\alpha(t+\tau)|^2|\alpha(t)|^2\rangle\\&\qquad+\langle\alpha'\frac{\partial}{\partial\alpha'}\langle|\alpha(t+\tau)|^2|[\alpha',\ t]\rangle\rangle\end{aligned}} \tag{10.5.29}$$

上式与泊松表示结果(7.7.76)非常相似。事实上，如果P函数仅仅与$|\alpha|^2$和$|\alpha'|^2$有关，则籍助于将$|\alpha|^2$(P表示)变到α（泊松表示)的代换，它们是同一的。

更令人感兴趣的测量结果是相关函数

$$\langle a^+(t)a^+(t+\tau)a(t+\tau)a(t)\rangle \equiv G^2(\tau,\ t) \tag{10.5.30}$$

格劳伯[10.1]已经证明，当我们把探测器在不同时刻接收到的光强联系起来时，上式就是所测得的量。

利用(10.5.12)把上式写成

$$G^2(\tau,\ t) = \mathrm{Tr}\{a^+ a e^{\hat{L}\tau} a\rho a^+\} \tag{10.5.31}$$

与上面所述相同的理由，上式成为

$$\boxed{\begin{aligned} G^2(\tau,\ t) &= \int \mathrm{d}^2\alpha \mathrm{d}^2\alpha' |\alpha^2|\,|\alpha'^2| P(\alpha, t+\tau;\ \alpha',\ t) \\ &= \langle |\alpha(t+\tau)|^2 |\alpha(t)|^2\rangle \end{aligned}} \tag{10.5.32}$$

这特别简单的结果意味着，测得的强度相关函数$G^2(\tau,\ t)$与P表示变量的相关函数是同一件事，而与(10.5.29)不同。

例子，驱动的阻尼振子　从10.4.1c节我们得到关于$\alpha-\alpha_s$的随机微分方程：

$$\mathrm{d}(\alpha-\alpha_s) = -\left(\frac{1}{2}K - i\omega'\right)(\alpha-\alpha_s)\mathrm{d}t + \sqrt{K\bar{N}}\,\mathrm{d}\eta(t)$$

(并有 $\mathrm{d}\eta^2 = \mathrm{d}\eta^{*2} = 0$, $\mathrm{d}\eta\mathrm{d}\eta^* = \mathrm{d}t$)。

因此，定义

$$\tilde{I} = |(\alpha-\alpha_s)^2|$$

所以

$$\mathrm{d}\tilde{I} = -K\tilde{I}\mathrm{d}t + \sqrt{K\bar{N}}\,[(\alpha-\alpha_s)^*\mathrm{d}\eta + (\alpha-\alpha_s)\mathrm{d}\eta^*]$$

及

$$\frac{\mathrm{d}\langle\tilde{I}\rangle}{\mathrm{d}t} = -K\langle\tilde{I}\rangle$$

这样，$\langle\tilde{I}\rangle$服从线性方程，而且可以应用量子回归定理，即在平稳态下

$$\langle \tilde{I}(t)\tilde{I}(0)\rangle_s = e^{-Kt}\langle \tilde{I}(0)^2\rangle$$

我们利用 $\alpha(t)$ 及 $\alpha^*(t)$ 中平稳分布的高斯性质，注意

$$\langle \tilde{I}(0)^2\rangle_s = \langle (x^2+y^2)^2\rangle = \langle x^4 + 2x^2y^2 + y^4\rangle$$

（这里 $\alpha = x + iy$）

而 x 和 y 在平稳态是无关的[根据(10.4.23)]，所以

$$\langle \tilde{I}(0)^2\rangle = 2\langle \tilde{I}(0)\rangle_s^2 = 2\bar{N}^2$$

即

$$\boxed{\langle \tilde{I}(t)\tilde{I}(0)\rangle_s = G^{(2)}(t)_s = 2\bar{N}^2 e^{-Kt}} \tag{10.5.33}$$

更简单的是

$$\boxed{\langle \alpha^*(t)\alpha(0)\rangle_s = G^{(1)}(t)_s = \bar{N}\exp\left[-\left(\frac{1}{2}K + i\omega\right)t\right]} \tag{10.5.34}$$

10.5.3 双能级原子的时间相关

这些时间相关提供了回归定理的完美应用，并能很快地从10.4节导出。为简单起见，我们只考虑平稳相关函数：

$$\underset{\sim}{G}(t) = \begin{bmatrix} \langle S^+(t)S^+(0)\rangle_s & \langle S^-(t)S^-(0)\rangle_s & \langle S^+(t)S_z(0)\rangle_s \\ \langle S^-(t)S^+(0)\rangle_s & \langle S^-(t)S^-(0)\rangle_s & \langle S^-(t)S_z(0)\rangle_s \\ \langle S_z(t)S^+(0)\rangle_s & \langle S_z(t)S^-(0)\rangle_s & \langle S_z(t)S_z(0)\rangle_s \end{bmatrix} \tag{10.5.35}$$

因而

$$\frac{d\underset{\sim}{G}(t)}{dt} = \begin{bmatrix} -\frac{1}{2}K(2\bar{N}+1) & 0 & id\varepsilon \\ 0 & -\frac{1}{2}K(2\bar{N}+1) & -id\varepsilon^* \\ \frac{1}{2}id\varepsilon^* & -\frac{1}{2}id\varepsilon & -K(2\bar{N}+1) \end{bmatrix} \underset{\sim}{G}(t) \tag{10.5.36}$$

可以对这些矩阵方程进行积分并给出全部解。初始条件可以由矩阵(10.4.28)的代数简化为

$$\underset{\sim}{G}(0)=\begin{bmatrix} 0 & \frac{1}{2}\langle 1+S_z\rangle_s & -\langle S^+\rangle_s \\ \frac{1}{2}\langle 1-S_z\rangle_s & 0 & \langle S^-\rangle_s \\ \langle S^+\rangle_s & -\langle S^-\rangle_s & 1 \end{bmatrix} \tag{10.5.37}$$

而平稳值已由 10.4.2 节给出，相关函数和谱的问题已由卡迈克尔和沃尔斯[10.6]给出。

10.6 广义的 P 表示

与在格劳伯-苏达香 P 表示中的情况一样，把密度矩阵表示成相干态密度矩阵的线性组合的思想，在用玻色算子 a^+ 和 a 描述的量子系统中特别有用，因为借助普通随机过程对量子主方程作出解释是唯一可能的方法。这与仅将生灭随机过程转变成扩散过程的泊松表示相反。

除了与经典随机分布的形式相似外，函数 $P(\alpha)$ 并不是真正的随机分布，而是属于准经典的分布类型。虽然对于热光场(是高斯分布——参看 10.4.1 节)和相干激光场(δ 函数分布)来说 $P(\alpha)$ 存在，但是对于具有非经典的光子统计场来说，$P(\alpha)$ 并不像性能良好的正函数那样必定存在（虽然克劳德和苏达香[10.3]根据非常奇异的分布已经证明它的确存在)。这些非经典场已经在原子荧光实验中观察到。还存在另一种准随机分布，它避免了 P 表示的某些问题。维格纳 (Wigner) 函数是第一种准随机方法，它可以从 P 表示经下面的积分而得到：

$$W(\alpha)=\frac{2}{\pi}\int P(\beta)\exp(-2|\beta-\alpha|^2)\mathrm{d}^2\beta \tag{10.6.1}$$

虽然维格纳函数总是作为非奇异函数存在，但可以假定它为负值。维格纳分布简化了对称有序算子乘积的平均运算，但是对求常用的常序算子积的平均来说，它是不方便的。

恒正的另一个表示是 Q 表示或根据相干态得到的密度算子的

对角矩阵：

$$Q(\alpha, \alpha^*) = \langle\alpha|\rho|\alpha\rangle \tag{10.6.2}$$

虽然该表示是正的，但它有一个缺点，即并非每个正 Q 函数都对应一个半正定的厄密密度算子。另外，只有在反常序算子乘积的 Q 表示中求矩值才是简单的。

格劳伯利用 10.1.2e 节中给出的算子展式定义了 R 表示：

$$R(\alpha^*, \beta) = \langle\alpha|\rho|\beta\rangle \exp\left(\frac{1}{2}|\alpha|^2 + \frac{1}{2}|\beta|^2\right)$$

$$\rho = \frac{1}{\pi^2}\int |\alpha\rangle R(\alpha^*, \beta)\langle\beta| \exp\left[-\frac{1}{2}|\alpha|^2 + |\beta|^2\right] d^2\alpha d^2\beta \tag{10.6.3}$$

而该表示对 α^*、β 是解析的(因此是非奇异的)，根据定义，它也是非正的，并且具有含高斯权重因子的归一化。根据这个理由，它不可能具有福克-普朗克方程或任何关于准概率性的解释。尽管如此，这表示的存在确实证明，任何 ρ 的常序可观察量的计算，可能与非奇异表示相同。

10.6.1 广义 P 表示的定义

为了处理产生非经典光子统计的非线性量子光学中的问题，德拉蒙德(Drummond)和加德纳用展开成非对角相干态投影算子的办法，引进了一类广义 P 表示[10.9]。所用的方法非常类似于7.7节中定义各种泊松表示所使用的方法。广义 P 表示定义如下。设

$$\rho = \int_{\mathscr{D}} \Lambda(\alpha, \beta) P(\alpha, \beta) d\mu(\alpha, \beta) \tag{10.6.4}$$

式中

$$\Lambda(\alpha, \beta) = \frac{|\alpha\rangle\langle\beta^*|}{\langle\beta^*|\alpha\rangle} \tag{10.6.5}$$

而 $d\mu(\alpha, \beta)$ 是积分测度，为了定义不同类型的可能表示，可以选取这个测度，而 $\mathscr{D}$ 是积分范围。投影算子 $\Lambda(\alpha, \beta)$ 对 (α, β) 是解析的。

积分测度的如下选取是有用的。

a) 格劳伯-苏达香 P 表示

$$d\mu(\alpha,\beta) = \delta^2(\alpha^* - \beta)d^2\alpha d^2\beta \tag{10.6.6}$$

这个测度与格劳伯-苏达香的 P 表示相对应。

b) 复 P 表示

$$d\mu(\alpha, \beta) = d\alpha d\beta \tag{10.6.7}$$

这里，(α, β) 被看成是复变量，它们在各自的环路 C、C' 上积分。在某些情况下这种表示的存在问题留到下节去证明。尤其是，对于以数态有限基展开的算子来说，这种表示存在。这是一种特殊的情况，其中非经典光子统计(光子反聚束)可能产生，格劳伯-苏达香 P 表示将是奇异的。由于 $P(\alpha, \beta)$ 出现复值，所以这个 P 表示称为复 P 表示。可以证明，所得出的 P 表示满足福克-普朗克方程，只是方程是在福克-普朗克方程的普通格劳伯-苏达香类型中用 (α, β) 代替 (α, α^*) 而得到的。

在某种情况下，福克-普朗克方程的精确解不可能象格劳伯-苏达香的 P 函数那样归一化。这些解可以通过在 (α, β) 的复相空间内选取 C、C' (积分路径)用现在的表示来处理。

c) 正 P 表示

$$d\mu(\alpha, \beta) = d^2\alpha d^2\beta$$

该表示允许 (α, β) 在整个复平面上独立地变化。在下节我们将证明，对一个物理密度算子来说，$P(\alpha, \beta)$ 总是存在的，并且总可以选取为正，在这种情况下，我们称为正 P 表示。这就意味着 $P(\alpha, \beta)$ 完全具有真正的随机性质。也可以证明，对于格劳伯-苏达香表示中的时间发展来说，所提供的任何福克-普朗克方程都存在，对于正 P 表示，相应的福克-普朗克方程伴随半正定的扩散系数。这使得我们能够导出一个随机微分方程，以及量子马尔可夫过程和普通扩散方程之间的对应关系。

当然，在所有的表示中，可观察矩由

$$\langle(a^+)^m a^n\rangle = \int_{\mathcal{D}} d\mu(\alpha, \beta)\beta^m\alpha^n P(\alpha, \beta) \tag{10.6.8}$$

给出，这是千真万确的。

10.6.2 存在性定理

在本节我们将证明，广义P表示具有很强的存在特性。我们用若干定理做到这点。注意：为简明起见，我们将使用符号

$$\underset{\sim}{\alpha}=(\alpha,\ \beta)$$

定理 1 用有限个数态展开的算子存在复P表示

证明 设

$$\rho=\sum_{n}\sum_{m}C_{nm}(a^{+})^{m}|0\rangle\langle 0|a^{n} \tag{10.6.9}$$

因而，由柯西定理，

$$\rho=\oint_{C,C'}\oint\Lambda(\underset{\sim}{\alpha})P(\underset{\sim}{\alpha})\mathrm{d}\mu(\underset{\sim}{\alpha}) \tag{10.6.10}$$

而其中

$$P(\underset{\sim}{\alpha})=(-1/4\pi^{2})\mathrm{e}^{\alpha\beta}\sum_{n}\sum_{m}C_{nm}n!m!\alpha^{-m-1}\beta^{n-1} \tag{10.6.11}$$

式中C、C'是包围原点的积分路径。

定理 2 任何在相干态有限区域上展开的算子，存在复P表示，即

$$\rho=\int_{D,D'}\int\Lambda(\alpha,\ \beta)C(\alpha,\ \beta)\mathrm{d}^{2}\alpha\mathrm{d}^{2}\beta \tag{10.6.12}$$

式中D、D'是每个复平面的边界。

证明 柯西定理的应用表明，如果

$$P(\underset{\sim}{\alpha})=-\frac{1}{4\pi^{2}}\int_{D,D'}\int C(\alpha',\ \beta')/[(\alpha-\alpha')(\beta-\beta')]\mathrm{d}^{2}\alpha'\mathrm{d}^{2}\beta' \tag{10.6.13}$$

因此，

$$\rho=\int_{C,C'}\int\Lambda(\underset{\sim}{\alpha})P(\underset{\sim}{\alpha})\mathrm{d}\alpha\mathrm{d}\beta \tag{10.6.14}$$

式中C、C'分别包围D、D'。因而，在关于任何相干态投影算子有界展式的情况下，存在复P表示。

定理 3 任何量子密度算子 ρ，存在正 P 表示，它为

$$P(\alpha)=(1/4\pi^2)\exp(-|\alpha-\beta^*|^2/4)\left\langle\frac{1}{2}(\alpha+\beta^*)|\rho|\frac{1}{2}(\alpha+\beta^*)\right\rangle \tag{10.6.15}$$

证明 由于 ρ 是密度算子，并且它由对角矩阵元乘上正函数组成，所以 $P(\underset{\sim}{\alpha})$是正的。为了证明这代表一般情况下的一个量子密度算子，利用了特征函数

$$\chi(\lambda)\equiv\mathrm{Tr}\{\rho e^{\lambda a^+}e^{-\lambda^* a}\} \tag{10.6.16}$$

为了唯一地定义密度算子，在 10.2.5 节中已经证明了这一点。根据 ρ 的 R 表示，特征函数为

$$\chi(\lambda)=\frac{1}{\pi}\int R(\alpha^*,\ \lambda+\alpha)\exp(-\lambda^*\alpha-|\alpha|^2)\mathrm{d}^2\alpha \tag{10.6.17}$$

现在我们将 ρ 的 R 表示代入利用 ρ 的对角矩阵元定义 $P(\underset{\sim}{\alpha})$的(10.6.15)式中。所以我们定义 ρ_P，就象由前面的过程给出 ρ 那样，通过具有 $P(\underset{\sim}{\alpha})$的正 P 表示式 (10.6.4) 给出 ρ_P，我们利用(10.6.16)计算相应的特征函数 $\chi_P(\lambda)$，并证明这个特征函数与 ρ 的原特征函数相同。于是，

$$\begin{aligned}\chi_P(\lambda)&=\iint P(\underset{\sim}{\alpha})\exp(\lambda\beta-\lambda^*\alpha)\mathrm{d}^2\alpha\mathrm{d}^2\beta\\&=\frac{1}{4\pi^4}\iiiint R(\alpha'^*,\beta')\exp\Big[\lambda\beta-\lambda^*\alpha\\&\quad-\frac{1}{2}|\alpha|^2-\frac{1}{2}|\beta|^2-|\alpha'|^2-|\beta'|^2\\&\quad+\frac{1}{2}\beta'^*(\alpha+\beta^*)+\frac{1}{2}\alpha'^*(\alpha^*+\beta)\Big]\mathrm{d}^2\alpha\mathrm{d}^2\beta\,\mathrm{d}^2\alpha'\mathrm{d}^2\beta'\end{aligned}$$

现在由定义作变量代换

$$\begin{aligned}&\gamma=(\alpha+\beta^*)^2/2,\quad \delta=(\alpha-\beta^*)/2\\&\alpha=(\gamma+\delta),\quad \beta^*=(\gamma-\delta)\\&\mathrm{d}^2\alpha\mathrm{d}^2\beta=4\mathrm{d}^2\gamma\mathrm{d}^2\delta\end{aligned} \tag{10.6.18}$$

注意到 R 是解析函数，则下述恒等式

$$R(\alpha^*,\ \gamma)=\frac{1}{\pi}\int R(\alpha^*,\ \beta)\exp(\gamma\beta^*-|\beta|^2)\mathrm{d}^2\beta \tag{10.6.19}$$

是很有用的。因此，上面那个特征函数的表示可以简化，以给出

$$\chi_P(\lambda)=\frac{1}{\pi^3}\iiint R(\alpha'^*,\ \gamma)\exp[\lambda(\gamma-\delta)^*-\lambda^*(\gamma+\delta)-|\gamma|^2-|\delta|^2-|\alpha'|^2+\alpha'\gamma^*]\mathrm{d}^2\gamma\mathrm{d}^2\beta\mathrm{d}^2\alpha' \tag{10.6.20}$$

$$=\frac{1}{\pi^2}\iint R(\alpha'^*,\ \gamma)\exp(|\lambda|^2+\lambda\gamma^*-\lambda\gamma-|\gamma|^2-|\alpha'|^2+\alpha'\gamma^*)\mathrm{d}^2\gamma\mathrm{d}^2\alpha' \tag{10.6.21}$$

$$=\frac{1}{\pi}\int R(\alpha^*,\ \lambda+\alpha)\exp(-\lambda^*\alpha-|\alpha|^2)\mathrm{d}^2\alpha \tag{10.6.22}$$

所以

$$\chi_P(\lambda)=\mathrm{Tr}\{\rho\mathrm{e}^{\lambda a^\dagger}+\mathrm{e}^{-\lambda^* a}\}=\chi(\lambda) \tag{10.6.23}$$

上式最后一步是根据关系到 R 表示的 (10.6.17) 中定义的特征函数恒等式得到的。于是我们推导出 $\rho_P=\rho$ 的结论。

10.6.3 关于泊松表示

给出整数范围内的概率分布 $q(n)$ 后，我们总可以定义一个相应的正密度矩阵

$$\rho=\sum_n|n\rangle\langle n|q(n) \tag{10.6.24}$$

并且，ρ 的 P 表示给出相应于 $P(n)$ 的泊松表示，于是，

$$q(n)=\langle n|\rho|n\rangle$$

$$=\langle n|\int\mathrm{d}\mu(\alpha,\ \beta)P(\alpha,\ \beta)\frac{|\alpha\rangle\langle\beta^*|}{\langle\beta^*|\alpha\rangle}|n\rangle \tag{10.6.25}$$

$$=\int\mathrm{d}\mu(\alpha,\ \beta)\frac{\mathrm{e}^{-\alpha\beta}(\alpha\beta)^n}{n!}P(\alpha,\ \beta) \tag{10.6.26}$$

因此，我们可以写成

$$q(n)=\int\mathrm{d}\mu(\alpha_1)f(\alpha_1)\frac{\mathrm{e}^{-\alpha_1}\alpha_1^n}{n!} \tag{10.6.27}$$

其中

$$f(\alpha_1)\mathrm{d}\mu(\alpha_1)=\int\mathrm{d}\mu(\alpha,\ \beta)\delta_\mu(\alpha\beta-\alpha_1)P(\alpha,\ \beta) \tag{10.6.28}$$

而 $\delta_\mu(\alpha_1-\alpha_2)$ 是对于测度 $\mu(\alpha_1)$ 定义的狄拉克 δ 函数，即

$$\int d\mu(\alpha_1)\delta_\mu(\alpha_1-\alpha_2)\phi(\alpha_1)=\phi(\alpha_2) \tag{10.6.29}$$

于是我们可以从定理3得出结论，如同7.7节所断言的那样，正泊松表示恒存在。也可以利用第一个定理证明，如果$x>N$（N有限时）$q(x)=0$，则复P表示恒存在。然而，更普遍的结果已经在7.7.3节中得到证明。

10.6.4 算子恒等式

根据非对角相干态投影算子的定义(10.6.5)，可以得到下面的恒等式。再次请注意，$\underset{\sim}{\alpha}$用来表示$(\alpha,\ \beta)$：

$$\left.\begin{aligned}
&a\Lambda(\underset{\sim}{\alpha})=\alpha\Lambda(\underset{\sim}{\alpha})\\
&a^+\Lambda(\underset{\sim}{\alpha})=(\beta+\partial/\partial\alpha)\Lambda(\underset{\sim}{\alpha})\\
&\Lambda(\underset{\sim}{\alpha})a^+=\beta\Lambda(\underset{\sim}{\alpha})\\
&\Lambda(\underset{\sim}{\alpha})a=(\partial/\partial\beta+\alpha)\Lambda(\underset{\sim}{\alpha})
\end{aligned}\right\} \tag{10.6.30}$$

将上面的恒等式代入定义P表示的(10.6.4)式，并应用分部积分（假定边界项为零），这些恒等式就可以对依赖于这种表示的P函数产生作用。

a) 格劳伯-苏达香P表示

结果与(10.2.26)相同。

b) 复P表示

$$\left.\begin{aligned}
&a\rho\longleftrightarrow\alpha P(\underset{\sim}{\alpha})\\
&a^+\rho\longleftrightarrow(\beta-\partial/\partial\alpha)P(\underset{\sim}{\alpha})\\
&\rho a^+\longleftrightarrow\beta P(\underset{\sim}{\alpha})\\
&\rho a\longleftrightarrow(\alpha-\partial/\partial\beta)P(\underset{\sim}{\alpha})
\end{aligned}\right\} \tag{10.6.31}$$

c) 正P表示

现在利用$\Lambda(\alpha,\ \beta)$的解析性质，并注意，如果

$$\begin{aligned}
&\alpha=\alpha_x+i\alpha_y\\
&\beta=\beta_x+i\beta_y
\end{aligned} \tag{10.6.32}$$

则

$$(\partial/\partial\alpha)\Lambda(\underset{\sim}{\alpha})=(\partial/\partial\alpha_x)\Lambda(\underset{\sim}{\alpha})=(-i\partial/\partial\alpha_y)\Lambda(\underset{\sim}{\alpha}) \tag{10.6.33}$$

及

$$(\partial/\partial\beta)\Lambda(\underset{\sim}{\alpha}) = (\partial/\partial\beta_x)\Lambda(\underset{\sim}{\alpha}) = (-i\partial/\partial\beta_y)\Lambda(\underset{\sim}{\alpha}) \tag{10.6.34}$$

以致在这种情况下和(10.6.31)都是正确的那样，我们也得到

$$a^+\rho \longleftrightarrow (\beta - \partial/\partial\alpha_x)P(\underset{\sim}{\alpha}) \longleftrightarrow (\beta + i\partial/\partial\alpha_y)P(\underset{\sim}{\alpha}) \tag{10.6.35}$$

$$\rho a \longleftrightarrow (\alpha - \partial/\partial\beta_x)P(\underset{\sim}{\alpha}) \longleftrightarrow (\alpha + i\partial/\partial\beta_y)P(\underset{\sim}{\alpha})$$

所有这些对应关系现在都可以在适当的时候用来导出福克-普朗克方程。

10.7 广义 P 表示对时间演化方程的应用

首先我们想证明格劳伯-苏达香 P 表示不会总产生一个合适的福克-普朗克方程，而实际的情况就是这样。

考虑一个相干地驱动的信号模式，它与非线性吸收器相干。为此目的，我们设[10.9]

$$H_1 = \sum_i \hbar\omega_i\left(\Gamma_i^+\Gamma_i + \frac{1}{2}\right) \tag{10.7.1}$$

$$H_2 = \sum_i [(a^+)^2 g_i\Gamma_i + a^2 g\Gamma_i^+] \tag{10.7.2}$$

$$H_3 = \hbar\omega\left(a^+a + \frac{1}{2}\right) + i\hbar(\varepsilon e^{-i\omega t}a^+ - \varepsilon^* e^{i\omega t}a) \tag{10.7.3}$$

这与10.4.1节所处理的例子非常相似。在相互作用图景中，(假定热槽的温度为零)利用(10.3.38)，我们得到主方程

$$\boxed{\frac{\partial\rho}{\partial t} = [\varepsilon a^+ - \varepsilon^* a,\ \rho] + \frac{1}{2}K[2a^2\rho(a^+)^2 - (a^+)^2a^2\rho - \rho(a^+)^2a^2]} \tag{10.7.4}$$

利用普通的格劳伯-苏达香算子对应关系，我们应得到

$$\frac{\partial}{\partial t}P(\alpha, \alpha^*) = \left[-\frac{\partial}{\partial\alpha}(\varepsilon - K\alpha^2\alpha^*)\right.$$

$$-\frac{1}{2}\frac{\partial^2}{\partial\alpha^2}(K\alpha^2)+\text{复共轭}\Big]P(\alpha,\alpha^*) \tag{10.7.5}$$

利用实变量：

$$x=(\alpha+\alpha^*)/\sqrt{2}$$

$$y=(\alpha-\alpha^*)/i\sqrt{2}$$

扩散矩阵为

$$\begin{bmatrix}-K/2 & -K\\ -K & K/2\end{bmatrix} \tag{10.7.6}$$

它并非半正定的。因此，式(10.7.4)那种时间演化方程将演化为奇性方程，所以我们被导致以各种P表示的另外的方程。

如果存在一个半正定的扩散矩阵，则下面将应用到一些规则，而且扩散矩阵将导致随机微分方程

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\alpha\\ \alpha^*\end{bmatrix}=\begin{bmatrix}\varepsilon-K\alpha^2\alpha^*\\ \varepsilon-K(\alpha^*)^2\alpha\end{bmatrix}+i\sqrt{K}\begin{bmatrix}\alpha\xi_1(t)\\ \alpha^*\xi(t)\end{bmatrix} \tag{10.7.7}$$

式中$\xi_1(t)$和$\xi_2(t)$是独立的白噪声。然而，矛盾出现了，因为ξ_1和ξ_2是独立的，所以α和α^*就不再保持复共轭关系。我们得到与泊松表示的场合相类似的情形，出现负的扩散矩阵。假定我们作代替$\alpha^*\to\beta$，以及把它们看成是正P表示的变量，则我们将证明，(10.7.7)式实际上是正确的。

10.7.1 复P表示

这里的方法产生与格劳伯-苏达香的情况非常类似的方程。我们假定，通过将不同算子的次序进行适当的重新排列，我们能够将量子力学主方程简化成[式中$(\alpha,\beta)=\underset{\sim}{\alpha}\equiv(\alpha^{(1)},\alpha^{(2)};\mu=1,2]$，

$$\frac{\partial\rho}{\partial t}=\int_{\sigma,\sigma'}\int A(\underset{\sim}{\alpha})\frac{\partial P(\underset{\sim}{\alpha})}{\partial t}\,\mathrm{d}\alpha\,\mathrm{d}\beta$$

$$=\int_{\sigma,\sigma'}\int\left\{\left[A^{\mu}(\underset{\sim}{\alpha})\frac{\partial}{\partial\alpha^{\mu}}+\frac{1}{2}D^{\mu\nu}(\underset{\sim}{\alpha})\frac{\partial}{\partial\alpha^{\mu}}\frac{\partial}{\partial\alpha^{\nu}}\right]A(\underset{\sim}{\alpha})\right\}P(\underset{\sim}{\alpha})\mathrm{d}\alpha\,\mathrm{d}\beta \tag{10.7.8}$$

的形式。现在我们作分部积分，如果我们能够忽略边界项，这可以

通过选取环路 C、C' 而做到这点，则至少由 $\Lambda(\underset{\sim}{\alpha})$ 的系数相等而得到一个解：

$$\frac{\partial P(\underset{\sim}{\alpha})}{\partial t}=\Big[-\frac{\partial}{\partial\alpha^{\mu}}A^{\mu}(\underset{\sim}{\alpha})+\frac{1}{2}\frac{\partial}{\partial\alpha^{\mu}}\frac{\partial}{\partial\alpha^{\nu}}D^{\mu\nu}(\underset{\sim}{\alpha})\Big]P(\underset{\sim}{\alpha}) \tag{10.7.9}$$

虽然这个主方程对导出 (10.7.8) 是充分的，但是由于 $\Lambda(\underset{\sim}{\alpha})$ 不是线性独立的，所以上述方程不是唯一的方程。应该注意到，对于这个复 P 表示来说，$A^{\mu}(\underset{\sim}{\alpha})$ 和 $D^{\mu\nu}(\underset{\sim}{\alpha})$ 对 $\underset{\sim}{\alpha}$ 总是解析的。因此，如果 $P(\alpha)$ 最初解析，则 (10.7.9) 在时间演化过程中保留了这种解析性。

10.7.2 正 P 表示

我们假定在考虑同样的方程 (10.7.8)，但用的是正 P 表示。对称矩阵总可分折成因子

$$\underset{\sim}{\boldsymbol{D}}(\underset{\sim}{\alpha})=\underset{\sim}{\boldsymbol{B}}(\underset{\sim}{\alpha})\underset{\sim}{\boldsymbol{B}}^{T}(\underset{\sim}{\alpha}) \tag{10.7.10}$$

现在写出

$$\boldsymbol{A}(\underset{\sim}{\alpha})=\boldsymbol{A}_{x}(\underset{\sim}{\alpha})+i\boldsymbol{A}_{y}(\underset{\sim}{\alpha}) \tag{10.7.11}$$

$$\underset{\sim}{\boldsymbol{B}}(\underset{\sim}{\alpha})=\underset{\sim}{\boldsymbol{B}}_{x}(\underset{\sim}{\alpha})+i\underset{\sim}{\boldsymbol{B}}_{y}(\underset{\sim}{\alpha}) \tag{10.7.12}$$

式中 $\boldsymbol{A}_{x}$、$\boldsymbol{A}_{y}$、$\underset{\sim}{\boldsymbol{B}}_{x}$ 及 $\underset{\sim}{\boldsymbol{B}}_{y}$ 都是实数。因此我们发现主方程产生

$$\begin{aligned}\frac{\partial\rho}{\partial t}&=\iint \mathrm{d}^{2}\alpha\mathrm{d}^{2}\beta\Lambda(\underset{\sim}{\alpha})(\partial P(\underset{\sim}{\alpha})/\partial t)\\&=\iint P(\underset{\sim}{\alpha})\Big[A_{x}^{\mu}(\underset{\sim}{\alpha})\partial_{\mu}^{x}+A_{y}^{\mu}(\underset{\sim}{\alpha})\partial_{\mu}^{y}\\&\quad+\frac{1}{2}(B_{x}^{\mu\sigma}B_{x}^{\nu\sigma}\partial_{\mu}^{x}\partial_{\nu}^{x}+B_{y}^{\mu\sigma}B_{y}^{\nu\sigma}\partial_{\mu}^{y}\partial_{\nu}^{y}\\&\quad+2B_{x}^{\mu\sigma}B_{y}^{\nu\sigma}\partial_{\mu}^{x}\partial_{\nu}^{y}\Big]\Lambda(\underset{\sim}{\alpha})\mathrm{d}^{2}\alpha\mathrm{d}^{2}\beta\end{aligned} \tag{10.7.13}$$

这里，为了简化符号，我们写成 $\partial/\partial\alpha_{x}^{\mu}=\partial_{\mu}^{x}$，等等，并利用了 $\Lambda(\underset{\sim}{\alpha})$ 的解析性及关于产生 (10.7.8) 的方法，作出下列任一个代换

$$\partial/\partial\alpha^{\mu}\longleftrightarrow\partial_{\mu}^{x}\longleftrightarrow-i\partial_{\mu}^{y} \tag{10.7.14}$$

现在假定分部积分是允许的，则我们把福克-普朗克方程简化成

$$\partial P(\underset{\sim}{\alpha})/\partial t = \Big\{ -\partial_\mu^x A_x^\mu(\underset{\sim}{\alpha}) - \partial_\mu^y A_y^\mu(\underset{\sim}{\alpha}) + \frac{1}{2}[\partial_\mu^x \partial_\nu^x B_x^{\mu\sigma}(\underset{\sim}{\alpha}) B_x^{\nu\sigma}(\underset{\sim}{\alpha}) + 2\partial_\mu^x \partial_\nu^y B_x^{\mu\sigma}(\underset{\sim}{\alpha}) B_y^{\nu\sigma}(\underset{\sim}{\alpha}) + \partial_\mu^y \partial_\nu^y B_y^{\mu\sigma}(\underset{\sim}{\alpha}) B_y^{\nu\sigma}(\underset{\sim}{\alpha})] \Big\} P(\underset{\sim}{\alpha}) \tag{10.7.15}$$

这也不是唯一的时间演化方程，但是(10.7.13)是(10.7.15)的结果。

然而，福克-普朗克方程(10.7.15)在向量为

$$(\alpha_x^{(1)}, \alpha_x^{(2)}, \alpha_y^{(3)}, \alpha_y^{(4)}) \equiv (\alpha_x, \beta_x, \alpha_y, \beta_y) \tag{10.7.16}$$

的四维空间中具有半正定的扩散矩阵。

我们发现漂移向量为

$$\mathscr{A}(\underset{\sim}{\alpha}) \equiv (A_x^{(1)}(\underset{\sim}{\alpha}), A_x^{(2)}(\underset{\sim}{\alpha}), A_y^{(3)}(\underset{\sim}{\alpha}), A_y^{(4)}(\underset{\sim}{\alpha})) \tag{10.7.17}$$

以及扩散矩阵为

$$\mathscr{D}(\underset{\sim}{\alpha}) = \begin{bmatrix} \underset{\sim}{B_x}\underset{\sim}{B_x^T} & \underset{\sim}{B_x}\underset{\sim}{B_y^T} \\ \underset{\sim}{B_y}\underset{\sim}{B_x^T} & \underset{\sim}{B_y}\underset{\sim}{B_y^T} \end{bmatrix}(\underset{\sim}{\alpha}) \equiv \mathscr{B}(\underset{\sim}{\alpha})\mathscr{B}^T(\underset{\sim}{\alpha}) \tag{10.7.18}$$

式中

$$\mathscr{B}(\underset{\sim}{\alpha}) = \begin{bmatrix} \underset{\sim}{B_x} & 0 \\ \underset{\sim}{B_y} & 0 \end{bmatrix}(\underset{\sim}{\alpha}) \tag{10.7.19}$$

于是$\mathscr{D}$显然是半正定(及非正定)的。相应的伊藤随机微分方程可写成

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \underset{\sim}{\alpha_x} \\ \underset{\sim}{\alpha_y} \end{bmatrix} = \begin{bmatrix} A_x(\underset{\sim}{\alpha}) \\ A_y(\underset{\sim}{\alpha}) \end{bmatrix} + \begin{bmatrix} \underset{\sim}{B_x}(\underset{\sim}{\alpha}) & \xi(t) \\ \underset{\sim}{B_y}(\underset{\sim}{\alpha}) & \xi(t) \end{bmatrix} \tag{10.7.20}$$

或重新组合成实数和虚数部分

$$\mathrm{d}\underset{\sim}{\alpha}/\mathrm{d}t = A(\underset{\sim}{\alpha}) + \underset{\sim}{B}(\underset{\sim}{\alpha})\xi(t) \tag{10.7.21}$$

除了代换$\alpha^* \to \beta$以外，(10.7.21)正好就是这种随机微分方程，它是通过利用格劳伯-苏达香表示以及自然而然地把具有非正定扩散矩阵的福克-普朗克方程变换成伊藤随机微分方程而得到的。

在我们的推导中，两种形式的变量(α, α^*)已被复平面内允许独立涨落的变量(α, β)所代替。正如这里所定义的，正P表示于是作为这个步骤的数学证明而出现。这个步骤密切地平行于7.7.4节中对正泊松表示所用的方法。

10.7.3 例子

我们讨论10.7节的例子。利用适当的算子对应关系，复P表示的福克-普朗克方程为

$$\partial_t P(\alpha, \beta) = \Big[-\frac{\partial}{\partial\alpha}(\varepsilon - K\alpha^2\beta) - \frac{1}{2}\frac{\partial^2}{\partial\alpha^2}(K\alpha^2) - \frac{\partial}{\partial\beta}(\varepsilon - K\alpha\beta^2) - \frac{1}{2}\frac{\partial^2}{\partial\beta^2}(K\beta^2)\Big] P(\alpha, \beta) \tag{10.7.22}$$

相当奇怪，我们看到这个福克-普朗克方程满足5.3.3节的位势条件。用这些符号，因为

$$\boldsymbol{A} = \begin{bmatrix} \varepsilon - K\alpha^2\beta \\ \varepsilon - K\beta^2\alpha \end{bmatrix}, \quad \underset{\sim}{B} = \begin{bmatrix} -K\alpha^2 & 0 \\ 0 & -K\beta^2 \end{bmatrix} \tag{10.7.23}$$

所以，利用(5.3.22, 23)

$$\boldsymbol{Z} = -\frac{2}{K}\begin{bmatrix} \varepsilon/\alpha^2 - K\beta + K/\alpha \\ \varepsilon/\beta^2 - K\alpha + K/\beta \end{bmatrix}$$

$$\frac{\partial Z_\alpha}{\partial\beta} = \frac{\partial Z_\beta}{\partial\alpha} = 2 \tag{10.7.24}$$

及

$$\phi(\alpha, \beta) = -\int (Z_\alpha d\alpha + Z_\beta d\beta)$$

$$= -\frac{2\varepsilon}{K}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right) + 2\log(\alpha\beta) - 2\alpha\beta \tag{10.7.25}$$

以致

$$\boxed{P_s(\alpha, \beta) = (\alpha\beta)^{-2}\exp\left[2\alpha\beta + \frac{2\varepsilon}{K}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)\right]} \tag{10.7.26}$$

关于这平稳分布，唯一可采用的环路是C、C'，它们是α和β

平面上的独立环路，并包围 $\alpha=0$ 及 $\beta=0$ 这两个本性奇点。

这种类型的可能解是非常有用的，而且不可能利用格劳伯-苏达香 P 表示而得到。矩可以从

$$\iint \mathrm{d}\alpha\,\mathrm{d}\beta\,\beta^m\alpha^n(\alpha\beta)^{-2}\exp\left[2\alpha\beta+\frac{2\varepsilon}{K}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\right] \tag{10.7.27}$$

得到，而且我们可以把 $\exp(2\alpha\beta)$ 展开成幂级数以及逐项进行围道积分，以得到

$$\langle(a^+)^m a^n\rangle=\sum_{r=0}^{\infty}\frac{\left(\frac{2\varepsilon}{K}\right)^{n+m-2r-2}}{r!\,(n-r-1)!\,(m-r-1)!} \tag{10.7.28}$$

上式是一个容易计算的级数。

利用正 P 表示，我们得到随机微分方程

$$\begin{bmatrix}\mathrm{d}\alpha\\ \mathrm{d}\beta\end{bmatrix}=\begin{bmatrix}\varepsilon-K\alpha^2\beta\\ \varepsilon-K\alpha\beta^2\end{bmatrix}\mathrm{d}t+i\sqrt{K}\begin{bmatrix}\alpha\mathrm{d}W_1(t)\\ \beta\mathrm{d}W_2(t)\end{bmatrix} \tag{10.7.29}$$

应该注意，这个方程并不包含任何非常明显的小噪声参量。然而，一个大的驱动场极限可以由设

$$\begin{aligned}K&=\tilde{K}/\varepsilon^2\\ \alpha&=\tilde{\alpha}\varepsilon\\ \beta&=\tilde{\beta}\varepsilon\end{aligned} \tag{10.7.30}$$

以致

$$\begin{bmatrix}\mathrm{d}\tilde{\alpha}\\ \mathrm{d}\tilde{\beta}\end{bmatrix}=\begin{bmatrix}1-\tilde{K}\tilde{\alpha}^2\tilde{\beta}\\ 1-\tilde{K}\tilde{\alpha}\tilde{\beta}^2\end{bmatrix}\mathrm{d}t+i\frac{\sqrt{\tilde{K}}}{\varepsilon}\begin{bmatrix}\tilde{\alpha}\mathrm{d}W_1(t)\\ \tilde{\beta}\mathrm{d}W_2(t)\end{bmatrix} \tag{10.7.31}$$

来得到。小噪声线性化过程可以在大驱动场及小非线性的这种限止下得到。实际上，这种性质是实际中在使用的。

参 考 文 献

第 1 章

[1.1] Lord Rayleigh: 1891 [Phil. Mag. 32, Scientific Papers III, 473, (1902)]

[1.2] A. Einstein: Ann. Phys. (Leipzig) 17, 549(1905)

[1.3] M. von Smoluchowski: Ann. Phys. (Leipzig) 21, 756(1906)

[1.4] P. Langevin: Comptes. Rendues 146, 530(1908)

[1.5] M. von Smoluchowski: Phys. Z. 16, 321 (1915); 17, 557, 585 (1916)

[1.6] B. J. Berne, R. Pecora: *Dynamic Light Scattering*(Wiley, New York 1976)

[1.7] A. J. Lotka: *Elements of Physical Biology* (Williams and Wilkins, Baltimore 1925) [Reissued as"Elements of Mathematical Biology", Dover, 1956]

[1.8] V. Volterra: Mem. Acad. Lincei 2, 31-113(1926)

[1.9] S. O. Rice: Bell Syst. Tech. J. 23, 282(1944); 24, 46(1945)

[1.10] W. Schottky: Ann. Phys. (Leipzig) 57, 541(1918)

[1.11] G. I. Taylor: Proc. London Math. Soc. 22, 196(1920)

[1.12] N. Wiener: Acta Math. 55, 117(1930)

[1.13] A. Khinchin: Math. Annalen 109, 604(1934)

[1.14] J. B. Johnson: Phys. Rev. 32, 97(1928)

[1.15] H. Nyquist: Phys. Rev. 32, 110(1928)

[1.16] C. G. Huffaker: Hilgardia 27, 343-383(1958)

[1.17] M.P. Hassell: *The Dynamics of Arthropod Predator-Prey Systems*, Monographs in Population Biology, No. 13 (Princeton University Press, Princeton 1978)

第 2 章

[2.1] W. Feller: *An Introduction to Probability Theory and Its Applications*, 2nd ed., Vol. II(Wiley, New York 1974)

[2.2] A. Papoulis: *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York 1965)

[2.3] A. N. Kolmogorov: *Foundations of the Theory of Probability* (Chelsea, New York 1950) [The German original appeared in 1933]

[2.4] R. L. Stratonovich: *Introduction to the Theory of Random Noise*(Gordon and Breach, New York, 1963)

[2.5] B. V. Gnedenko: *The Theory of Probability*, 2nd ed (Chelsea, New York 1962)

[2.6] E.Meeron: J. Chem. Phys. 27. 1238(1957)

[2.7] N. G. van Kampen: Physica 74, 215(1973); 74, 239(1973)

[2.8] J. Marcienkiewicz: Math. Z. 44, 612(1939)

[2.9] A. K. Rajagopal, E. C. G. Sudarshan: Phys. Rev. A. 10, 1852 (1974)

第 3 章

[3.1] F. Haake: Springer Tracts in Modern Physics, Vol. 66(Springer, Berlin, Heidelberg, New York 1973)

[3.2] I. F. Gihman, A. V. Skorohod: *The Theory of Stochastic Processes* Vols. I, II, III. (Springer, Berlin, Heidelberg, New York 1975)

[3.3] [Ref. 3.2, Vol. II]

[3.4] W. Feller: Ann. Math. 55, 227-246(1952)

[3.5] G. E. Uhlenbeck, L. S. Ornstein: Phys. Rev. 36, 823(1930)

[3.6] A. Papoulis: *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York 1965)

第 4 章

[4.1] I. I. Gihman, A. V. Skorohod: *Stochastic Differential Equations* (Springer, Berlin, Heidelberg, New York 1972)

[4.2] R. L. Stratonovich: *Introduction to the Theory of Random Noise* (Gordon and Breach, New York, London 1963)
[4.3] L. Arnold: *Stochastic Differential Equations* (Wiley-Interscience, New York 1974)
[4.4] R. Kubo: In *Stochastic Processes in Chemical Physics*, ed. by K. E. Shuler (Wiley-Interscience, New York 1969)
[4.5] N. G. van Kampen: Phys. Rept. 24C, 171—228(1976)

第 5 章

[5.1] A. D. Fokker: Ann. Phys. (Leipzig) 43. 310(1915)
[5.2] M. Planck: Sitzungsber. Preuss. Akad. Wiss. Phys.Math. Kl., 325(1917)
[5.3] A. N. Kolmogorov. Math. Ann. 104, 415—418(1931)
[5.4] W. Feller: Ann. Math. 55, 227—246(1952)
[5.5] A. T. Bharucha-Reid: *Elements of the Theory of Markov Processes and Their Application* (McGraw-Hill, New York 1960)
[5.6] A. Erdelyi: *Bateman Manuscript* Project, Vols. 1—3 (McGraw-Hill, New York 1953)
[5.7] N. G. van Kampen: Physica 23, 707(1957); 23, 816(1957)
[5.8] U. Uhlhorn: Ark. Fys. 17, 361(1960)
[5.9] R. Graham, H. Haken: Z. Phys. 243.289-302(1971)
[5.10] H. A. Kramers: Physica 7.284(1940)
[5.11] L. Onsager: Phys. Rev. 37.405(1931)
[5.12] H. B. G. Casimir: Kev. Mod. Phys. 17, 343(1945)

第 6 章

[6.1] H. Haken: *Synergetics* An Introduction, 2nd ed. (Springer, Berlin, Heidelberg, New York 1978)
[6.2] H. C. Brinkman: Physica 22, 29(1956)
[6.3] R. L. Stratonovich: *Introduction to the Theory of Random Noise* (Gordon and Breach, New York, London 1963)
[6.4] G. Wilemski: J. Stat. Phys. 14, 153—170(1976)
[6.5] U. M. Titulaer: Physica 91A, 321—344(1978)

[6.6] G. C. Papanicolaou: In *Modern Modelling of Continuum Phenomena*, Lecture Notes in Applied Maths, Vol. 16 (Am Math. Soc., Providence RI 1977), p. 109
[6.7] G. C. Papanicolaou, W. Kohler: Comm. Pure Appl. Math. 27, 641—668(1974)
[6.8] R. Graham, T. Tel: J. Stat. Phys. 35, 729(1984)
[6.9] R. Graham, T. Tel: Phys. Rev. A31, 1109(1985)
[6.10] R. Beals, V. Protopescu: J. Stat. Phys. 32, 565(1983)
[6.11] J. V. Selinger, U. M. Titulaer: J. Stat. Phys. 36, 293(1984)
[6.12] U. M. Titulaer: J. Stat. Phys. 37, 589(1985)

第 7 章

[7.1] I. Matheson, D. F. Walls, C. W. Gardiner: J. Stat. Phys. 12, 21(1975)
H. K. Janssen: Z. Phys. 260, 67(1974)
[7.2] N. G. van Kampen: Can. J. Phys. 39 551—567(1961)
[7.3] H. A. Kramers: Physica 7, 284(1940)
[7.4] J. E. Moyal: J. R. Stat. Soc. 11, 151—210(1919)
[7.5] A. Einstein: Ann. Phys. (Leipzig)549(1905)
[7.6] T. G. Kurtz: J. Chem. Phys. 50, 460(1969)
[7.7] S. Karlin, H. M. Taylor: *A First Course in Stochastic Processes*, 2nd ed. (Academic Press, New York 1975)
[7.8] H. Haken: *Synergetics* An Introduction, 2nd ed. (Springer, Berlin, Heidelberg, New York 1978)
[7.9] J. Schnakenberg: Rev Mod. Phys. 48, 571—586 (1976); also *Thermodynamic Network Analysis of Biological Systems*, 2nd ed. (Springer, Berlin, Heidelberg, New York 1981)
[7.10] C. W. Gardiner, S. Chaturvedi: J. Stat. Phys. 17, 429—468 (1977); 18, 501—522(1978)
[7.11] W. Bernard, H. B. Callen: Rev. Mod. Phys. 31, 1017 (1959), Phys. Rev. 118, 1466(1960)
[7.12] K. Hochberg: Ann. Probab. 3, 433—458(1978)
[7.13] N. G. van Kampen: Adv. Chem. Phys. 34, 245—309(1976)

第 8 章

[8.1] R. F. Curtain: "Stochastic Partial Differential Equations", in *Stochastic Nonlinear Systems*, ed. by L. Arnold, R. Lefever (Springer, Berlin, Heidelberg, New York 1981)

[8.2] J. Keizer: J. Chem. Phys. 64, 1679—1687(1976)

[8.3] G. Nicolis: J. Stat. Phys. 6, 195(1972)

[8.4] F. Schlogl: Z. Phys. 253, 147(1972)

[8.5] L. Arnold: "Consistency of Models of Chemical Reactions" in *Dynamics of Synergetic Systems*, ed. by H. Haken (Springer, Berlin, Heidelberg, New York 1980)

[8.6] N. G. van Kampen: "Fluctuations in Continuous Systems", in *Topics in Statistical Mechanics and Biophysics*, ed. by R. A. Piccirelli(Am. Inst. of Physics, 1976)

[8.7] C. van den Broek, W. Horsthemke, M. Malek- Mansour: Physica 89A, 339—352(1977);
L. Brenig, C. van den Broek: Phys. Rev. 21A, 1039(1980)

[8.8] H. Grad: "Principles of the Kinetic Theory of Gases", in *Handbuch der Physik*, Vol. 12, ed. by S. Flügge (Springer, Berlin, Göttingen, Heidelberg 1958)

[8.9] C. W. Gardiner, M. L. Steyn-Ross: Phys. Rev. 29A, 2823 (1984)

第 9 章

[9.1] M. Abramowitz, I. Stegun: *Handbook of Mathematical Functions* (Dover, New York 1964, 1965)

[9.2] H. A. Kramers: Physica 7, 284(1940)

[9.3] A summary of the situation is given in the appendix to M. Buttiker, R. Landauer: "Transport and Fluctuations in Linear Arrays of Multistable Systems" in *Nonlinear Phenomena at Phase Transitions and Instabilities*, ed. by T. Riste (Plenum Press, New York, London 1982)

[9.4] H. Eyring: J. Chem. Phys. 3, 1935(1935)

[9.5] S. A. Hawley: Biochem. 10, 2436(1971);

S. A. Hawley, R. M.Mitchell: Biochem. 14, 3257(1975)
[9.6] J. S. Langer: Ann. Phys. N. Y. 54, 258(1969);
R. Landauer, J. A. Swanson: Phys. Rev. 121, 1668(1961)
[9.7] Z. Schuss: S. I. A. M. Rev. 22, 119(1980);
B. J. Matkowsky, Z. Schuss: S. I. A. M. J. Appl. Math. 33, 365(1977)
[9.8] A. D. Ventsel, M. I. Freidlin: Russian. Math. Surveys 25, 1 (1970)

第 10 章

[10.1] R. J. Glauber: Phys. Rev. 130, 2529(1963); 131, 2766(1963)
[10.2] E. C. G. Sudarshan: Phys. Rev. Lett. 10, 277(1963)
[10.3] J. R. Klauder, E. C. G. Sudarshan: *Fundamentals of Quantum Optics* (Benjamin, New York 1970)
[10.4] R. J. Glauber: In *Quantum Optics*, ed. by S. Kay, A. Maitland (Academic Press, New York 1970)
[10.5] W. Louisell: In *Quantum Optics*, ed. by S. Kay, A. Maitland (Academic Press, New York 1970); also *Quantum Statistical Properties of Radiation*(Wiley, New York 1974)
[10.6] H. J. Carmichael, D. F. Walls: J. Phys. B9, 1199(1976)
[10.7] A. Einstein: Phys. Z. 18, 121(1917)
[10.8] M. Lax: Phys. Rev. 172, 350(1968)
[10.9] P. D. Drummond, C. W. Gardiner: J. Phys. A 13, 2353—2368 (1980)

参 考 书 目

Agarwal, G. S.: "Quantum Statistical Theories of Spontaneous Emission and their Relation to Other Approaches", Springer Tracts in Modern Physics, Vol. 70 (Springer, Berlin, Heidelberg, New York 1974)

书中列出了实际应用量子马尔可夫过程的一些好例子。

Arnold, L.: *Stochastic Differential Equations: Theory and Applications*, (Wiley, New York 1974)

关于随机微分方程方面非常值得一读的数学教科书。读者如对随机微分方程严格而可以理解的论述有兴趣，本书是值得推荐的。

Arnold, L., Lefever, R. (eds.): *Stochastic Nonlinear Systems* (Springer, Berlin, Heidelberg, New York 1981)

本书包括了应用随机方法方面的最新领域，并附有大多数有关应用领域的参考文献。

Bharucha-Reid, A. T.: *Elements of the Theory of Markov Processes and their Applications*, (McGraw-Hill, New York 1960)

一本应用方面很有用的书。包括了群体生物学、核过程、天文学、化学和排队。对于一些近似于真实系统的模型，本书给出了许多严格解。

Cox, D. R., Miller, H. D.: *The Theory of Stochastic Processes* (Methuen, London 1965)

按英国传统写成的书，它更适用于统计学而不是自然科学，非常容易阅读。

Feller, W.: *An Introduction to Probability Theory and its Applications*, Vol. II, 2nd ed. (Wiley, New York 1971)

一本很值得一读、甚至还很有趣的书，主要是关于随机过程基础的。作者是本科目的创始人之一。本书对实用来说并不特别有用，但是更确切地说，对应用工作者而言本书是该科目内在优美的一个启示。

Gihman, I.I., Skorohod, A. V.: *Stochastic Differential Equations* (Springer, Berlin, Heidelberg, New York 1972)

详尽的俄国著作。不容易阅读，但值得努力一读。本书发展了几乎所有象随机微分方程那样的马尔可夫过程理论，包括某些非马尔可夫过程。

Gihman, I.I., Skorohod, A. V.: *The Theory of Stochastic Processes*, Vol. I (1974), Vol. II, Vol. III (Springer, Berlin, Heidelberg, New York

详尽而严格的俄国传统式的大部头著作，不很实用，但是可以认为是最现代的严格工作。经过努力，一个非数学家也可以了解。

Goel, N. S., Richter-Dyn, N.: *Stochastic Models in Biology* (Academic, New York 1974)

一本主要关于人口生灭过程方面的实用书籍。书中用到了某些逼近方法。

Graham, R.: "Statistical Theory of Instabilities in Stationary Nonequilibrium Systems with Applications to Lasers and Non linear Optics", Springer Tracts in Modern Physics, Vol. 66 (Springer, Berlin, Heidelberg, New York 1973)

随机方法在量子光学和其他非平衡系统中的应用。

Haake, F.: "Statistical Treatment of Open Systems by Generalised Master Equations", Springer Tracts in Modern Physics, Vol. 66 (Springer, Berlin, Heidelberg, New York 1973)

从物理学家的观点来看，本书给出了应用投影算子方法的很好概述。

Haken, H.: "Laser Theory" (Band XXV/2c of Handbuch der Physik, ed. by L. Genzel) (Springer, Berlin, Heidelberg, New York 1970)

一本非常全面地处理量子光学的书。本书包含了光学方面的专业工作者感兴趣的大量细节，同时还从跟本书所介绍的完全不同的观点讨论了量子马尔可夫过程，给出了P表示和别的量子-经典对应关系的全面处理方法。

Haken, H.: *Synergetics, An Introduction*, 2nd ed. (Springer, Berlin, Heidelberg, New York 1978)

关于把随机过程应用于各种合作现象的主要参考书。本书既提供了许多初等的例子，又提供了高深的工作。

Hänggi, P., Thomas, H.: "Stochastic Processes: Time Evolution, Symmetries and Linear Response" (Physics Reports 88, 207—319,

1982)

本书是从随机微分方程对物理的应用观点来写的，它也包含了某些非马尔可夫的过程。

Karlin, S., Taylor, H. M.: *A First Course in Stochastic Processes*, 2nd ed. (Academic, New York 1975)

一本包含许多加工过的例子的、非常值得一读的数学书。本书着重讨论马尔可夫链和跳过程。

Kay, S. M., Maitland, A. (eds): "Quantum Optics" — Proceedings of Scottish Universities Summer School (Academic, New York 1970)

一本关于应用量子马尔可夫系统的经典著作，这种系统出现于量子光学作为一门科学刚刚获得发展的时代。本书特别与格劳伯、哈肯和刘维的论文有关。

Lax, M.: "Classical Noise" I Rev. Mod. Phys. 32, 25 (1960); II Phys. Chem. Solids 14, 248 (1960) (with P. Mengert); III Rev. Mod. Phys. 38, 359 (1966); IV Rev. Mod, Phys. 38, 541(1966)

这是一本随机方法对物理系统、特别是对固态物理应用的开创性的、但处理相当涣散的书。

Louisell, W.: *Radiation and Noise in Quantum Electronics* (McGraw-Hill, New York 1964)

第一本系统的量子电子学课本。主要面向电子工程、而不包含任何量子主方程的内容。

Louisell, W.: *Quantum Statistical Properties of Radiation* (Wiley, New York 1974)

本书包含了马尔可夫方法对量子光学应用方面的全部资料，实质上是作者早期著作的第二版。

Nicolis, G, Prigogine, I.: *Self Organisation in Nonequilibrium Systems* (Wiley, New York 1977)

主要论述有关自组织的书。本书的确是一本关于布鲁塞尔学派随机方法的值得一读和有用的书。作为关于随机方法在反应扩散系统及非平衡统计力学中的应用方面最新工作的一个原始出处，非常值得推荐。

Oppenheim, I., Shuler, K. E., Weiss, G. H.: *Stochastic Processes in Chemical Physics - the Master Equation* (M.I.T. Press, Camb-

ridge 1977)

本书是随机方法的简明介绍。包括了关于此科目基本论文的非常重要和有用的提纲，其中有（茨温采）Zwanzig 首先引进的投影技术的基本论文。

Papoulis, A.: *Probability. Random Variables, and Stochastic Processes* (McGraw-Hill, New York 1965)

一本学习上实用的、但又是合理地严格的概率方法的优秀参考书。主要适用于电子工程应用方面。

Risken, H.: "The Fokker Planck Equation" (Springer, Berlin, Heidelberg, New York 1984)

这是非常全面地处理求解和应用福克-普朗克方程的方法的一本书。包含了连分数方法的处理、在周期性势场中的运动的处理以及包含了一些不太专门的书中不包括的大量材料。

Stratonovich, R. L.: *Topics in the Theory of Random Noise*, Vols. I, II (Gordon and Breach, New York 1963)

一本无线电工程师对随机方法看法的纲要，包含许多优美的结果，它们以一种颇有特色的、常常是非常令人启发的风格出现。随机微分方程是以真实过程的极限的方式引入的，本书是应用随机微分方程方面的基础书籍之一。

van Kampen, N. G.: *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, New York, Oxford 1981)

一本介绍随机过程在物理及化学现象中应用的书。以清晰的推理和严谨的风格为其特征。本书特别注意解决物理和化学应用中所用近似的合理性条件。

Wax N.: *Selected Papers on Noise and Stochastic Processes* (Dover, New York 1954)

一本十分重要的、有关历史的参考书，它既包含了乌伦贝克（Uhlenbeck）和奥恩斯坦（Ornstein）在布朗运动方面的工作及Chandrasekhar 在物理和天文学中的随机问题的经典论文，也包含了其他的经典论文。

符 号 索 引

检索页码以符号在正文中首次述及为准. 次序以字母表方式排列,小写在先,不计字体. 索引有三部分:拉丁符号、希腊符号和数学符号.

拉丁符号

a^+, a 产生和湮灭算子

ac-lim 几乎必然收敛

$A(z, t)$ 漂移向量

$A_a[J(\alpha)]$ 泊松表示下福克-普朗克方程的漂移向量

$\hat{A}_a(\eta)$ 泊松表示的漂移向量

$B(z, t)$ 扩散矩阵

$\tilde{B}_{ab}[J(\alpha)]$ 泊松表示下福克-普朗克方程的扩散矩阵

$\hat{B}_{ab}(\eta)$ 泊松表示中去掉因子 $e^2=V^{-1}$ 后的扩散矩阵

d_{if} 扩散跳跃概率

$d\tilde{W}(q, t)$ 局部维纳过程的傅里叶变换

$dW(r, t)$ 空间相关的局部维纳过程

$\mathscr{D}$ 泊松表示中的积分区域

$\tilde{D}(r)$ 非对角空间相关的扩散张量

$\tilde{(\mathscr{D}, \eta)}(r, t)$ 非局部扩散

$\mathrm{erf}(x)=\sqrt{\frac{\pi}{2}}\int_0^x \mathrm{d}t e^{-t^2}$ 误差函数

$f(\alpha, t)$ 泊松表示的拟概率

$f(\alpha, t|\alpha', t')$ 泊松表示的条件拟概率

$G(s, t)$ 母函数

$G(\tau)$ 自相关函数

H	哈密顿算子
$H_n(x)$	厄密多项式
$j(r,\ t)$	扩散流
$J(x)$	跳过程的概率流
$J(x,\ t)$	概率流
$J_A(\alpha)$	泊松表示的概率流
$\hat{J}_A[\eta]$	泊松表示的概率流
$k_A^{\pm}$	化学主方程的速率常数
$\mathrm{L}_n(x)$	拉盖尔多项式
ms-lim	均方极限
M^A	化学主方程中的终态参数
$\mathcal{N}$	归一化常数
N^A	化学主方程中的初态参数
$p(x_1,\ t_1;\ x_2,\ t_2;\ x_3,\ t_3;\ \cdots)$	联合概率密度
$p(x_1,\ t_1 \mid x_2,\ t_2)$	条件概率密度
$P(A \mid B)$	条件概率
P	绝热消去中的投影算子
$P(x,\ t \mid x',\ t')$	跳过程的条件概率
$P(\alpha,\ \alpha^*)$	格劳伯—苏达香 P 函数
$P(\alpha,\ \beta)$	广义 P 函数
$P_\lambda(x)$	前向福克—普朗克算子的本征函数
q	傅里叶变换变量
$Q_\lambda(x)$	后向福克—普朗克算子的本征函数
$r^A = M^A - N^A$	
$R\ (\alpha^*, \beta)$	格劳伯 R 函数
st-lim	随机极限
$S(w)$	谱
$S^+,\ S^-,\ S_z$	泡利矩阵
$t^+(x),\ t^-(x)$	每单位时间跳变 $x \longrightarrow x+1$，$x \longrightarrow x-1$ 的转移概率
$t_A^{\pm}(x)$	化学主方程中的每单位时间的转移概率

$\bar{T}(x)$	平均首次通过时间
$\mathrm{var}\{X\}$	X的方差
V, Ω	系统体积
$W(t)$	维纳过程
$\boldsymbol{W}(t)$	多变量维纳过程
$W(x\|z,\ t)$	跳概率
$x[V,\ t]=\int_V \mathrm{d}^3r\rho(r,\ t)$	体积V内的分子数

希腊符号

$\boldsymbol{\alpha}$, α, $\alpha(t)$, $\boldsymbol{\alpha}(t)$	泊松表示的变量
$\underset{\sim}{\alpha}=(\alpha,\ \beta)$	
$\alpha_n(x)$	n阶导出矩
Γ_i	热槽算子
$\delta(x)$	狄拉克δ函数
$\delta(m,\ n)\equiv\delta m,\ n$	克罗内克δ
$\frac{\delta F(\rho)}{\delta\rho(r)}$	泛函导数
η	泊松表示的浓度变量
$\kappa_\lambda^{\pm}$	以浓度表示的化学速率常数
$\Lambda(\alpha,\ \beta)=\|\alpha\rangle\langle\beta^*\|/\langle\beta^*\|\alpha\rangle$	
$\mu(\alpha)$	用于泊松表示的测度
$\xi(t)$	涨落力
$\pi_a(x)$	经由a出口的概率
ρ	密度矩阵或密度算子
$\rho(T)$	正则系综的密度算子
$\rho(r,\ t)$	浓度变量
$\sigma[X]$	X的标准差
τ_c	相关时间
$\Phi(S)$	累积母函数
$\phi(S)$	矩母函数
$\chi(\lambda,\ \lambda^*)$, $\chi(\lambda)$	量子特征函数
Ω, V	系统大小参数

数学符号

$[A, B] = AB - BA$	算子 A, B 的交换子
$\langle X_i, X_j\rangle \equiv \langle X_iX_j\rangle - \langle X_i\rangle\langle X_j\rangle$	X_i 和 X_j 的协方差矩阵
$\langle R\rangle$	R 的均值
$\langle X^r\rangle_t$	X 的阶乘矩
$\langle\alpha(t)\mid[\alpha', s]\rangle$	在初始条件 $\alpha(s)=\alpha'$ 下 $\alpha(t)$ 的均值
$\ll X^r\gg_t$	X 的阶乘累积量
$\ll X_1^{m_1}X_2^{m_2}\cdots X_n^{m_n}\gg$	变量 X_1、X_2、…、X_n 的累积量
$\mid n\rangle$	(粒子)数状态
$\mid\alpha\rangle$	相干态
$\parallel\alpha\rangle$	伯格曼态
$\int_{t_0}^{t} G(t')\mathrm{d}W(t')$	伊藤随机积分
$S\int_{t_0}^{t} G(t')\mathrm{d}W(t')$	斯特拉托诺维奇随机积分
$\tilde{f}(s)=\int_0^{\infty}\mathrm{d}te^{-st}f(t)$	$f(t)$ 的拉普拉斯变换

作者索引

Beals, R.	比尔斯
Bernard, W.	伯纳德
Boltzmann, L.	玻耳兹曼
Brinkman, H. C.	布林克曼
Brown, R.	布朗
Callen, H. B.	卡伦
Carmichael, H. J.	卡迈克尔
Casimir, H. B. G.	卡西米尔
Drummond, P. D.	德拉蒙德
Einstein, A.	爱因斯坦
Feller, W.	费勒
Fokker, A. D.	福克
Freidlin, M. I.	弗雷德林
Gardiner, C. W.	加德纳
Glauber, R. J.	格劳伯
Graham, R.	格雷厄姆
Haken, H.	哈肯
Hochberg, K.	霍克伯格
Johnson, J. B.	约翰逊
Keizer, J.	凯泽
Khinchin, A.	辛钦
Klauder, J. R.	克劳德
Kohler, W.	科勒
Kolmogorov, A. N.	柯尔莫哥洛夫
Kramers, A. H.	雷默斯
Kurtz, T. G.	库尔茨
Landauer, R.	兰道

Langer, J. S.	兰格
Langevin, P.	郎之万
Lax, M.	拉克斯
Lotka, A. J.	洛特卡
Louisell, W.	路易塞尔
Marcienkiewicz, J.	马辛凯维奇
Matkowsky, B. J.	马特可夫斯基
Meeron, E.	米隆
Moyal, J. E.	莫耶尔
Nyquist, H.	尼奎斯
Onsager, L.	翁萨格
Ornstein, L. S.	奥恩斯坦
Papanicolaou, G. C.	帕普尼科劳
Papoulis, A.	帕普里斯
Planck, M.	普朗克
Protopescu, V.	普罗托斯库
Rayleigh, Lord	瑞利
Rice, S. O.	赖斯
Schnakenberg, J.	施纳肯伯格
Schottky, W.	肖特基
Schuss, M.	舒斯
Selinger, J. V.	塞林格
Smoluchowski, M.	斯莫路乔夫斯基
Stratonovich, R. L.	斯特拉托诺维奇
Sudarshan, E.C.G.	苏达香
Taylor, G. I.	泰勒
Tel, T.	特尔
Titulaer, U. M.	泰特莱尔
Uhlenbeck, G. E.	乌伦贝克
Uhlhorn, U.	乌尔霍恩
Van den Broek, C.	范登布罗克
Van Kampen, N. G.	范坎彭
Ventsel, A. D.	文切尔
Volterra, V.	伏尔特拉

Walls, D. F.	沃尔斯
Wiener, N.	维纳
Wilemski, G.	威伦斯基

论 题 索 引

Absorbing boundary condition for backward FPE 129 后向福克-普朗克方程的吸收边界条件
Additive invariants 加法不变量
Adiabatic elimination 绝热消去
  in quantum systems 量子系统中的绝热消去
  of fast variables 快变变量的绝热消去
  of inhomogeneous modes in reaction diffusion systems 反应扩散系统中非均匀模式的绝热消去
Almost certain 几乎必然
Arrhenius formula 阿伦尼厄斯公式
Autocorrelation function 自相关函数
  and Poisson rep. 自相关函数与泊松表示
  for Markov process 马尔可夫过程的自相关函数
  in terms of eigenfunctions 用本征函数表示的自相关函数
  with detailed balance 细致平衡时的自相关函数
Autocorrelation matrix 自相关矩阵
  and eigenfunctions 自相关矩阵与本征函数
  stationary 平稳自相关矩阵

Backward differential Chapman-Kolmogorov equation 后向查普曼-柯尔莫哥洛夫方程
Backward FPE, boundary conditions for 后向福克-普朗克方程及其边界条件
Backward master equation and first passage times 后向主方程及首次通过时间
Baker-Hausdorff formula 贝克-豪斯道夫公式
Bargmann states, defined 伯格曼态及其定义
Bath correlation functions 热槽相关函数

Bernoulli distribution	伯努里分布
Bernoulli trials	伯努里试验
Birth-death master equation	生灭主方程
and bistability	生灭主方程及双稳定性
in quantum harmonic oscillator	量子谐振子中的生灭主方程
one-variable	单变量生灭主方程
Birth-death process	生灭过程
for flow	流的生灭过程
Birth-death processes, boundary conditions for	生灭过程的边界条件
Birth-death systems, many-variable	多变量生灭系统
Bistability	双稳定性
in multivariable systems	多变量系统中的双稳定性
with birth-death master equation	生灭主方程的双稳定性
Bistable system, chemical	化学双稳系统
Boltzmann equation from Boltzmann master equation	由玻耳兹曼主方程得到的玻耳兹曼方程
Boltzmann master equation	玻耳兹曼主方程
in Poisson representation	泊松表示下的玻耳兹曼主方程
Boundary conditions for	
backward FPE	后向福克-普朗克方程的边界条件
birth-death processes	生灭过程的边界条件
Kramers' equation	克喇默斯方程的边界条件
many-variable FPE	多变量福克-普朗克方程的边界条件
Smoluchowski equation	斯莫路乔夫斯基方程的边界条件
Boundary conditions for FPE	福克-普朗克方程的边界条件
absorbing	福克-普朗克方程的吸收边界条件
at a discontinuity	福克-普朗克方程在间断点的边界条件
at infinity	福克-普朗克方程在无穷远处的边界条件
reflecting	福克-普朗克方程的反射边界条件
Brownian motion	布朗运动
and adiabatic elimination	布朗运动与绝热消去

continuous sample path	连续样本路径的布朗运动
corrections to Smoluchowski equation	布朗运动的斯莫路乔夫斯基方程修正
in double potential	双井势中的布朗运动
Kramers' equation for	布朗运动的克雷默斯方程
Brownian particle	布朗粒子
Cauchy process	柯西过程
Central limit theorem	中心极限定理
Chaos	混沌
Chapman-Kolmogorov equation	查普曼-柯尔莫哥洛夫方程
differential	微分查普曼-柯尔莫哥洛夫方程
Characteristic function	特征函数
of a Gaussian	高斯变量的特征函数
quantum	量子特征函数
Chemical FPE	化学福克-普朗克方程
Chemical intermediates, elimination of	化学中间体及其消去
Chemical reaction	
$A+2X\rightleftarrows 3X$, $A\rightleftarrows X$	化学反应 $A+2X\rightleftarrows 3X$, $A\rightleftarrows X$
$A+2X\rightleftarrows 3X$, $A\rightleftarrows X$, master equation	化学反应 $A+2X\rightleftarrows 3X$, $A\rightleftarrows X$ 的主方程
$A+2X\rightleftarrows 3X$, in Poisson rep.	化学反应 $A+2X\rightleftarrows 3X$ 的泊松表示
$A+X\rightarrow 2X+D$, $B+X\rightleftarrows C$, solution of ME	化学反应 $A+X\rightleftarrows 2X+D$, $B+X\rightleftarrows C$, 及其主方程之解
$A+X\rightleftarrows 2X$, $B+X\rightleftarrows C$, positive Poisson rep.	化学反应 $A+X\rightleftarrows 2X$, $B+X\rightleftarrows C$ 及其正泊松表示
$A+X\rightleftarrows 2X$, $B+X\rightleftarrows C$, Poisson rep. solution	化学反应 $A+X\rightleftarrows 2X$, $B+X\rightleftarrows C$ 及其泊松表示解
$A+X\rightleftarrows 2X$, $B+X\rightleftarrows C$, complex Poisson rep.	化学反应 $A+X\rightleftarrows 2X$, $B+X\rightleftarrows C$ 及其复泊松表示
$A+X\rightleftarrows 2X$, $B+X\rightleftarrows C$, spatially distributed	化学反应 $A+X\rightleftarrows 2X$, $B+X\rightleftarrows C$ 及其空间分布
$A+Y\rightleftarrows X+Y$, $Y\rightleftarrows 2X$, elimination of intermediate	化学反应 $A+Y\rightleftarrows X+Y$, $Y\rightleftarrows 2X$ 及中间体的消去

$B+X\rightleftarrows C, A+X\to 2X$, spatially distributed	化学反应 $B+X\rightleftarrows C, A+X\to 2X$ 及其空间分布
$B\to X$, $2X\to A$, and positive Poisson rep.	化学反应 $B\to X$, $2X\to A$ 以及正泊松表示
$B\to X$, $2X\to A$, complex Poisson rep.	化学反应 $B\to X$, $2X\to A$ 的复泊松表示
$B\rightleftarrows X$, $A+X\to 2X$, analogy to harmonic oscillator	化学反应 $B\rightleftarrows X$, $A+X\to 2X$ 与谐振子相似
$X+A\rightleftarrows 2X$, FPE	化学反应 $X+A\rightleftarrows 2X$ 及其福克-普朗克方程
$X+A\rightleftarrows 2X$, chemical FPE	化学反应 $X+A\rightleftarrows 2X$ 的化学福克-普朗克方程
$X\rightleftarrows A$, and system size expansion	化学反应 $X\rightleftarrows A$ 及系统大小展开
$X\rightleftarrows A$, master equation	化学反应 $X\rightleftarrows A$ 的主方程
$X\rightleftarrows Y$, as reaction diffusion system	作为反应扩散系统的化学反应 $X\rightleftarrows Y$
$X\rightleftarrows Y\rightleftarrows A$	化学反应 $X\rightleftarrows Y\rightleftarrows A$
$X\rightleftarrows Y\rightleftarrows A$, chemical FPE	化学反应 $X\rightleftarrows Y\rightleftarrows A$ 的化学福克-普朗克方程
$X_1\rightleftarrows 2X_2$, reaction diffusion equations	化学反应 $X_1\rightleftarrows 2X_2$ 的反应扩散方程
$X_1\rightleftarrows X_2$, solution of master equation	化学反应 $X_1\rightleftarrows X_2$ 及其主方程的解
nonlocal	非局部化学反应
prey-predator system, FPE	被捕食-捕食系统的化学反应及其福克-普朗克方程
unimolecular	单分子化学反应
Chemical reactions, transition state theory	化学反应及过渡态理论
Chymotrypsinogen, reversible denaturation of	糜蛋白酶及其可逆质变
Coherent states	相干态
defined	相干态的定义
expansion of an operator in	相干态中算子的展开

expansion of arbitrary states in terms of Poissonian number distribution of	任意态按相干态的展开
properties of	相干态的性质
Collisions	碰撞
and flow	碰撞和流
Boltzmann master equation for	关于碰撞的玻耳兹曼主方程
Combinatorial kinetics; defined	组合动力学的定义
Commutation relations	交换(对易)关系
Complex P-representation	复P表示
defined	复P表示的定义
FPE from	由复P表示得到的福克-普朗克方程
operator identities for	复P表示的算子恒等式
Conditional probability	条件概率
and eigenfunctions	条件概率与本征函数
Continuity in stochastic processes	随机过程连续性
Corrected Smoluchowski equation and escape problem	修正的斯莫路乔夫斯基方程及逃逸问题
Correlation	相关
Correlation functions	相关函数
defined	相关函数的定义
of heat bath	热槽的相关函数
Correlation length	相关长度
Correlation time	相关时间
Correlations	相关
behaviour at instability point	在不稳定点的相关行为
of smooth functions of bath operators	热槽算子平滑函数的相关
space-time	空间-时间相关
Covariance	协方差
Covariance matrix	协方差矩阵
Creation operator	产生算子
Critical fluctuations	临界涨落
Critical slowing down	临界慢化

Cubic process	立方过程
Cumulant	累积量
Cumulant, factorial	阶乘累积量
Cumulant generating function	累积母函数
factorial	阶乘累积母函数
Density matrix	密度矩阵
properties of	密度矩阵的性质
Destruction operator	湮灭算子
Detailed balance	细致平衡
conditions for	细致平衡的条件
defined	细致平衡的定义
in birth-death master equation	生灭主方程中的细致平衡
in diffusion	扩散中的细致平衡
Deterministic process	确定性过程
Differential Chapman-Kolmogorov equation	微分查普曼-柯尔莫哥洛夫方程
backward	后向微分查普曼-柯尔莫哥洛夫方程
Diffusion	扩散
and detailed balance	扩散及细致平衡
coefficient	扩散系数
current, defined	扩散流及其定义
equation	扩散方程
fluctuating partial differential equation	扩散涨落偏微分方程
in a gravitational field	重力场中的扩散
in multivariate master equation	多变量主方程中的扩散
inhomogeneous anisotropic	非均匀各向异性扩散
Diffusion master equation	扩散主方程
continuum form of	扩散主方程的连续统形式
system size expansion of	扩散主方程的系统大小展开
Diffusion matrix, defined	扩散矩阵的定义
Diffusion process	扩散过程
approximation by jump process	用跳过程逼近扩散过程
Divergence problems in nonlinear	非线性反应扩散系统中的分歧问题

reaction diffusion systems
Double well potential 双井势
diffusion in 双井势中的扩散
Drift vector; defined 漂移向量的定义

Eigenfunctions 本征函数
and autocorrelation matrix 本征函数与自相关矩阵
and conditional probability 本征函数与条件概率
and exit time 本征函数与出口时间
and spectrum matrix 本征函数与谱矩阵
for many-variable FPE 多变量福克-普朗克方程的本征函数
for one-variable FPE 单变量福克-普朗克方程的本征函数
Eigenfunctions for FPE, variational principle for 福克-普朗克方程的本征函数及其变分原理
Einstein equations and two-level atom 爱因斯坦方程与双能级原子
Ensemble average 系综平均
Entrance boundary for FPE 福克-普朗克方程的入口边
Equilibration of populations in bistable system 双稳系统中粒子数的平衡
Ergodic 各态历经
Ergodic properties and stationary processes 各态历经性与平稳过程
Escape 逃逸
Escape probability of particle in a double well 双井势中粒子的逃逸概率
Escape time 逃逸时间
Event 事件
Existence theorems for *P*-representations *P* 表示的存在定理
Exit boundary for FPE 福克-普朗克方程的出口边界
Exit points, distribution of by asymptotic method 出口点及由渐近方法得到的分布
Exit time 出口时间
see also first passage time 亦见首次通过时间

and eigenfunctions	出口时间与本征函数
asymptotic analysis of	出口时间的渐近分析
from a region, many variable FPE	多变量福克-普朗克方程的区域出口时间
in a double well potential	双井势中的出口时间
through a particular end of an interval	经由区间特定端点的出口时间
Factorial correlation in reaction diffusion system	反应扩散系统中的阶乘相关
Factorial cumulant	阶乘累积量
Factorial moment	阶乘矩
Fast variable	快变变量
Fick's law	斐克定律
with fluctuations	具有涨落的斐克定律
First passage time	首次通过时间
for one dimensional FPE	一维福克-普朗克方程的首次通过时间
of particle in a double well	双井势中粒子的首次通过时间
Flow	流
and collisions	流和碰撞
as a birth death process	作为生灭过程的流
in position space	位置空间中的流
in velocity space	速度空间中的流
Fluctuating force	涨落力
Fluctuation dissipation theorem	涨落耗散定理
Flux, phenomenological	通量及唯象学的通量
Fokker-Planck equation (FPE)	福克-普朗克方程
backward, boundary conditions	后向福克-普朗克方程的边界条件
boundary conditions	福克-普朗克方程的边界条件
chemical	化学福克-普朗克方程
connection with SDE	福克-普朗克方程与随机微分方程的关系
defined	福克-普朗克方程的定义
from complex P-representation	由复P表示得到的福克-普朗克方程
from Poisson representation	由泊松表示得到的福克-普朗克方程

from positive P-representation	由正P表示得到的福克-普朗克方程
from P-representation	由P表示得到的福克-普朗克方程
functional	泛函福克-普朗克方程
many dimensional	多维福克-普朗克方程
many variable, boundary conditions	多变量福克-普朗克方程的边界条件
many variables, eigenfunctions	多变量福克-普朗克方程的本征函数
one dimensional	一维福克-普朗克方程
one variable, eigenfunctions	单变量福克-普朗克方程的本征函数
small noise expansion for	福克-普朗克方程的小噪声展开
stationary solution for one variable	单变量福克-普朗克方程的平稳解
stratonovich form	福克-普朗克方程的斯特拉托诺维奇形式
Force, phenomenological	唯象力
Fourier analysis of fluctuating functions	涨落函数的傅里叶分析
FPE	福克-普朗克方程
Functional derivative, defined	泛函导数及其定义
Functional FPE, for reaction diffusion systems	反应扩散系统的泛函福克-普朗克方程
Gaussian distribution	高斯分布
Gaussian random variable	高斯随机变量
Generalised P-representations	广义P表示
and time development equations	广义P表示及时间演化方程
defined	广义P表示的定义
relation to Poisson representation	广义P表示与泊松表示的关系
Generating function	母函数
for birth-death master equation	生灭主方程的母函数
for Poisson distribution	泊松分布的母函数
Glauber-Sudarshan P-representation	格劳伯-苏达香P表示

Haken's model	哈肯模型
Harmonic oscillator	谐振子
defined	谐振子的定义
driven by fluctuating field	由涨落场驱动的谐振子
driven damped	受驱动的阻尼谐振子
driven; in P-representation	P 表示中的驱动谐振子
interaction with external field	谐振子与外场的相互作用
quantum master equation	谐振子的量子主方程
quantum mechanics of	谐振子的量子力学
Heat bath; defined	热槽及其定义
Hermite polynomials	厄密多项式
in adiabatic elimination	绝热消去中的厄密多项式
Homogeneous Markov process	齐次马尔可夫过程
defined	齐次马尔可夫过程的定义
Independence	独立性
Independent random variables	独立随机变量
and characteristic function	独立随机变量与特征函数
Interaction picture	相互作用图景
and driven atom	相互作用图景与被驱动的原子
Ito stochastic integral; defined	伊藤随机积分的定义
properties of	伊藤随机积分的性质
Ito's formula; derivation of	伊藤公式的推导
Johnson noise	约翰逊噪声
Joint probability	联合概率
Jump process	跳过程
Kramers' equation	克雷默斯方程
boundary conditions	克雷默斯方程的边界条件
Kramers' method	克雷默斯方法
for escape problems	逃逸问题的克雷默斯方法
in several dimensions	多维克雷默斯方法
Kramers-Moyal expansion	克雷默斯-莫耶尔展开

and system size expansion	克雷默斯-莫耶尔展开与系统大小展开
defined	克雷默斯-莫耶尔展开的定义
in birth-death master equation	生灭主方程中的克雷默斯-莫耶尔展开
in Boltzmann master equation	玻耳兹曼主方程中的克雷默斯-莫耶尔展开
Kurtz's theorem	库尔茨定理
Laguerre polynomial	拉盖尔多项式
Langevin equation	郎之万方程
Laplace transform and adiabatic elimination	拉普拉斯变换与绝热消去
Laser light scattering	激光散射
Law of large numbers	大数法则
Limit of sequence of random variables	随机变量序列的极限
Limit	极限
almost certain	几乎必然极限
in distribution	依分布收敛的极限
in probability	依概率收敛的极限
mean square	均方极限
stochastic	随机极限
Lindeberg condition	林德贝格条件
Linear SDE	线性随机微分方程
multivariable	多变量线性随机微分方程
single-variable	单变量线性随机微分方程
Liouville equation	刘维方程
quantum	量子刘维方程
Liouville operators; defined	刘维算子及其定义
Local and global descriptions of chemical reaction, connection between	化学反应的局部和整体描述以及它们之间的关系
Local and global fluctuations	局部和整体涨落
Markov assumption	马尔可夫假设
Markov postulate	马尔可夫公设

Markov process	马尔可夫过程
autocorrelation for	马尔可夫过程的自相关
continuous; defined	连续马尔可夫过程的定义
defined	马尔可夫过程的定义
homogeneous; defined	齐次马尔可夫过程的定义
quantum mechanical	量子力学马尔可夫过程
stationary; defined	平稳马尔可夫过程的定义
Master equation	主方程
approximation by Fokker-Planck equation	用福克-普朗克方程逼近主方程
many-variable	多变量主方程
many-variable; Kramers-Moyal expansion for	多变量主方程的克雷默斯-莫耶尔展开
mean first passage times	主方程的平均首次通过时间
one-variable; stationary solution of	单变量主方程的平稳解
phase space	相空间主方程
quantum; defined	量子主方程及其定义
quantum; derivation	量子主方程及其导出
stationary solutions without detailed balance	不具有细致平衡时的主方程平稳解
ME	主方程
Mean first passage time	平均首次通过时间
in one-variable FPE	单变量福克-普朗克方程中的平均首次通过时间
for master equations	主方程的平均首次通过时间
Mean square limit in definition of stochastic integral	依随机积分定义的均方极限
Mean value	均值
Metastability	亚稳定性
Microscopic reversibility	微观可逆性
Moment	矩
Moment generating function	矩母函数
Moments and characteristic function	矩与特征函数

Multitime averages in quantum systems	量子系统中的多次平均
Multivariate master equation	多变量主方程
for diffusing system	扩散系统的多变量主方程
Natural boundary for FPE	福克-普朗克方程的自然边界
Noise	噪声
Nonanticipating function; defined	非可料函数的定义
Nonlinear absorber	非线性吸收器
Nyquist's formula	尼奎斯公式
Nyquist's theorem	尼奎斯定理
Onsager relations	翁萨格关系
Operator correspondences in Glauber-Sudarshan P-representation	格劳伯-苏达香P表示中的算子对应
Operator identities	算子恒等式
for complex P-representation	复P表示的算子恒等式
for positive P-representation	正P表示的算子恒等式
Ornstein-Uhlenbeck process	奥恩斯坦-乌伦贝克过程
and Onsager relations	奥恩斯坦-乌伦贝克过程与翁萨格关系
and quantum harmonic oscillator	奥恩斯坦-乌伦贝克过程与量子谐振子
eigenfunctions	奥恩斯坦-乌伦贝克过程与本征函数
exit time from a circle	奥恩斯坦-乌伦贝克过程的圆形区域的出口时间
in small noise expansion of SDE	随机微分方程的小噪声展开中的奥恩斯坦-乌伦贝克过程
spectrum matrix	奥恩斯坦-乌伦贝克过程的谱矩阵
stationary distribution	奥恩斯坦-乌伦贝克过程的平稳分布
stationary variance	奥恩斯坦-乌伦贝克过程的平稳方差
stationary variance in two dimensions	二维奥恩斯坦-乌伦贝克过程的平稳方差
time correlation function	奥恩斯坦-乌伦贝克过程的时间相关函数
time correlation matrix in	奥恩斯坦-乌伦贝克过程的稳态时间相

stationary state	关矩阵
time dependent	时间独立奥恩斯坦-乌伦贝克过程
Ornstein-Uhlenbeck process(SDE)	奥恩斯坦-乌伦贝克过程的随机微分方程
multivariate	多变量奥恩斯坦-乌伦贝克过程的随机微分方程
Oscillator, noisy	噪声振子
P-representation	P表示
complex; defined	复P表示及其定义
existence theorems for	复P表示的存在定理
FPE from	由复P表示得到的福克-普朗克方程
generalised	广义P表示
generalised; defined	广义P表示的定义
Glauber-Sudarshan	格劳伯-苏达香P表示
positive; defined	正P表示及其定义
Pauli matrices, defined	泡利矩阵及其定义
Phase space master equation	相空间主方程
Phase transition, second order;	相变、二级相变及反应扩散系统中的
in reaction diffusion system	相变和二级相变
Phenomenological flux	唯象学通量
Phenomenological force	唯象力
Poisson distribution	泊松分布
and coherent states	泊松分布和相干态
Poisson distribution relaxation in reaction diffusion system	反应扩散系统中泊松分布的衰减
Poisson process	泊松过程
Poisson representation	泊松表示
and Boltzmann master equation	泊松表示和玻耳兹曼主方程
and time correlation functions	泊松表示与时间相关函数
and unimolecular reactions	泊松表示与单分子反应
complex; defined	复泊松表示的定义
in reaction diffusion system	反应扩散系统中的泊松表示
positive; defined	正泊松表示及其定义
real; defined	实泊松表示及其定义

relation to generalised P-representation	泊松表示和广义P表示的关系
Poissonian time dependent solutions of certain master equations	一个主方程的时间相依泊松分布解
Positive P-representation	正P表示
defined	正P表示的定义
FPE from	由正P表示得到的福克-普朗克方程
operator identities for	正P表示的算子恒等式
Potential barrier, escape over	势垒及从势垒上逸出
Potential conditions	位势条件
for FPE in complex P-representation	复P表示中福克-普朗克方程的位势条件
for many-variable FPE	多变量福克-普朗克方程的位势条件
Potential solution of one-variable FPE	单变量福克-普朗克方程的位势解
Prey-predator system	被捕食者-捕食者系统
chemical FPE for	被捕食-捕食者系统的化学福克-普朗克方程
Probability	概率
a priori	先验概率
conditional	条件概率
joint	联合概率
Probability axioms	概率公理
Probability current; defined	概率流及其定义
Probability density	概率密度
Projection operator in adiabatic elimination	绝热消去中的投影算子
Q-representation	Q表示
Quantum characteristic function	量子特征函数
Quantum Liouville equation	量子刘维方程
Quantum Markov process, derivation of	量子马尔可夫过程及其导出

Quantum Markov process, time correlation functions in	量子马尔可夫过程中的时间相关函数
Quantum master equation	量子主方程
defined	量子主方程的定义
derivation of	量子主方程的导出
for harmonic oscillator	谐振子的量子主方程
Quantum mechanics of harmonic oscillator	谐振子的量子力学
Quantum regression theorem	量子回归定理
R-representation	R表示
Random telegraph process	随机电报过程
and driven atom	随机电报过程及驱动原子
Random variable	随机变量
Gaussian	高斯随机变量
Random variables	随机变量
independent	独立随机变量
independent(and central limit theorem)	独立随机变量与中心极限定理
Random walk	随机游动
approximation by FPE	用福克-普朗克方程逼近随机游动
Rayleigh process	瑞利过程
eigenfunctions	瑞利过程的本征函数
Reaction diffusion chemical equation	反应扩散化学方程
Reaction diffusion master equation; defined	反应扩散主方程及其定义
Reaction diffusion system	反应扩散系统
fluctuating partial differential equations	反应扩散系统的涨落偏微分方程
in Poisson representation	泊松表示中的反应扩散系统
Reaction diffusion systems, divergence problems	反应扩散系统中的发散问题
Reflecting boundary condition for backward FPE	后向福克-普朗克方程的反射边界条件

Regression theorem	回归定理
quantum	量子回归定理
Relaxation time, dependence on peak populations	弛豫时间以及它与顶值布居数的关系
Sample path, continuous	连续样本路径
Scaling assumption in approximation of master equation	逼近主方程时的标度假设
Schrödinger's equation	薛定谔方程
SDE	随机微分方程
Separatrix	分隔号
Set of events	事件集
Sets of probability zero	零概率集
Shot noise	散粒噪声
Sigma algebra	σ 代数
Save	
noisy	有噪声支配
silent	无噪声支配
Slaving principle	支配原理
Small noise expansion	小噪声展开
for SDE	随机微分方程的小噪声展开
of FPE	福克-普朗克方程的小噪声展开
Smoluchowski equation	斯莫路乔夫斯基方程
and escape problem	斯莫路乔夫斯基方程和逃逸问题
boundary conditions	斯莫路乔夫斯基方程的边界条件
corrected	修正过的斯莫路乔夫斯基方程
corrections to	对斯莫路乔夫斯基方程的修正
defined	斯莫路乔夫斯基方程的定义
for diffusion in a double well	在双井势中扩散的斯莫路乔夫斯基方程
Spatial correlations	空间相关
Spectrum	谱
flat	平直的谱
of Ornstein-Uhlenbeck process	奥恩斯坦-乌伦贝克过程的谱
Spectrum matrix	谱矩阵

and eigenfunctions 谱矩阵与本征函数
with detailed balance 具有细致平衡时的谱矩阵
Splitting probability in double well potential 在双井势中的分裂概率
Standard deviation 标准差
Stationary 平稳
Stationary distribution of FPE, asymptotic method for 福克-普朗克方程的平稳分布及其渐近方法
Stationary Markov process; defined 平稳马尔可夫过程的定义
Stationary process, approach to 趋向于平稳过程
Stationary solutions, many-variable FPE 多变量福克-普朗克方程的平稳解
Stationary systems 平稳系统
Statistical mechanics 统计力学
Stochastic differential equation 随机微分方程
and positive Poisson rep. 随机微分方程与正泊松表示
as limit of nonwhite process 随机微分方程作为非白噪声过程的极限
connection with FPE 随机微分方程与福克-普朗克方程的关系
definition and properties of 随机微分方程的定义与性质
dependence of solutions on initial conditions 随机微分方程的解对初始条件的依赖性
dependence of solutions on parameters partial 随机微分方程的解对参数的依赖性
small noise expansion for 随机微分方程的小噪声展开
Stochastic integral 随机积分
defined 随机积分的定义
Ito 伊藤随机积分
Stratonovich 斯特拉托诺维奇随机积分
Stochastic partial differential equation 随机偏微分方程
Stochastic process 随机过程
defined 随机过程的定义

separable	可分随机过程
Stratonovich SDE	斯特拉托诺维奇型随机微分方程
as limit of nonwhite process	斯特拉托诺维奇型随机微分方程作为非白噪声过程的极限
definition and connection with Ito SDE	斯特拉托诺维奇型随机微分方程的定义以及跟伊藤型随机微分方程的关系
Stratonovich stochastic integral	斯特拉托诺维奇随机积分
System size	系统大小
System size expansion	系统大小展开
and critical fluctuations	系统大小展开与临界涨落
and Kramers-Moyal expansion	系统大小展开与克雷默斯-莫耶尔展开
and Kurtz's theorem	系统大小展开与库尔茨定理
in birth-death master equation	生灭主方程中的系统大小展开
in Boltzmann master equation	玻耳兹曼主方程中的系统大小展开
in reaction diffusion master equation	反应扩散主方程中的系统大小展开
of master equation	主方程的系统大小展开
Third-order noise	三阶噪声
defined	三阶噪声的定义
Three-state interpretation of bistable system	双稳系统的三态解释
Time correlation function	时间相关函数，见自相关函数
Time correlation functions	
and P-representation	时间相关函数与 P 表示
and quantum harmonic oscillator	时间相关函数与量子谐振子
for two-level atom	双能级原子的时间相关函数
in quantum Markov processes	量子马尔可夫过程中的时间相关函数
Time reversal invariance	时间反演不变性
Transition probability	转移概率
Transition state theory of chemical reactions	化学反应中的过渡态理论
Trimolecular reaction as a limit of bimolecular reaction	三分子反应作为双分子反应的极限
Two-level atom	双能级原子

driven	受驱动的双能级原子
time correlations	双能级原子的时间相关
Unimolecular reactions in Poisson representation	泊松表示中的单分子反应
Unstable state, decay from	非稳态及来自非稳态的衰变
Van Kampen's system size expansion	范坎彭系统大小展开
Variance	方差
Variational principle for eigenfunctions of FPE	福克-普朗克方程的本征函数的变分原理
Viscosity	粘滞性
Volterra-Lotka system	伏尔特拉-洛特卡系统，见被捕食者-捕食者系统
Von Neumann's equation	冯·诺伊曼方程
White noise	白噪声
as limit of nonwhite process	白噪声作为非白噪声过程的极限
Wiener process	维纳过程
as limit of random walk	维纳过程作为随机游动的极限
autocorrelation function	维纳过程的自相关函数
eigenfunctions	维纳过程的本征函数
independence of increments	维纳过程的增量独立性
irregularity of sample paths	维纳过程样本路径的不规则性
nondifferentiability of sample paths	维纳过程样本路径的不可微性
Wiener-Khinchin theorem	维纳-辛钦定理
Wigner representation	维格纳表示