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

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   *C. W. Gardiner*
October 1989

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   *C. W. Gardiner*
March 1985

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 approxima-tion techniques: small noise expansions and adiabatic elimination.
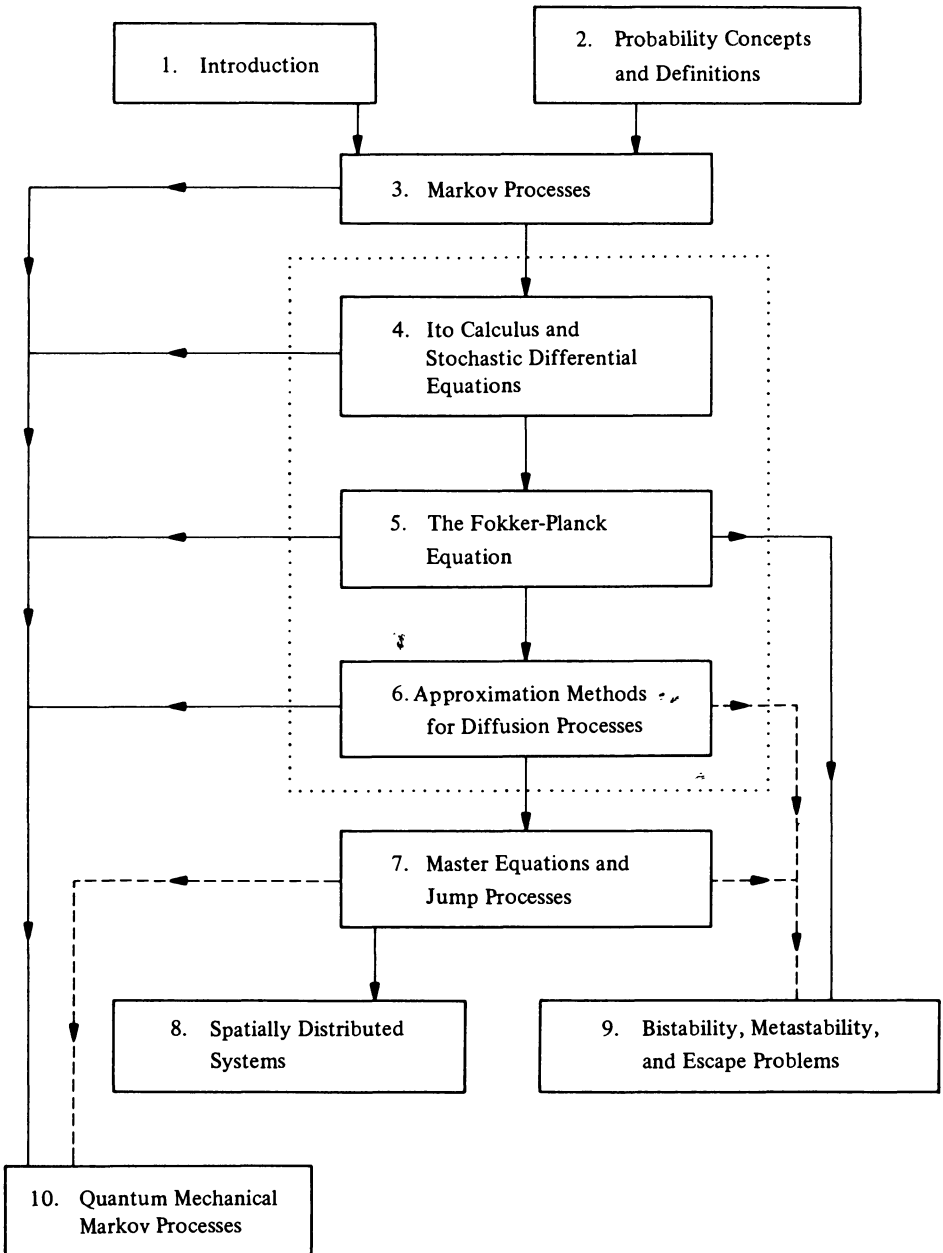
The remainder of the book is built around this core, since very many methods of treating the jump processes in Chapter 7 and the spatially distributed systems, themselves best treated as jump processes, depend on reducing the system to an approximating diffusion process. Thus, although *logically* the 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 proc-esses. It is with some disappointment that I have noted a tendency among mathe-maticians 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, renor-malisation 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

Contents

6. Approximation Methods for Diffusion Processes

7. Master Equations and Jump Processes

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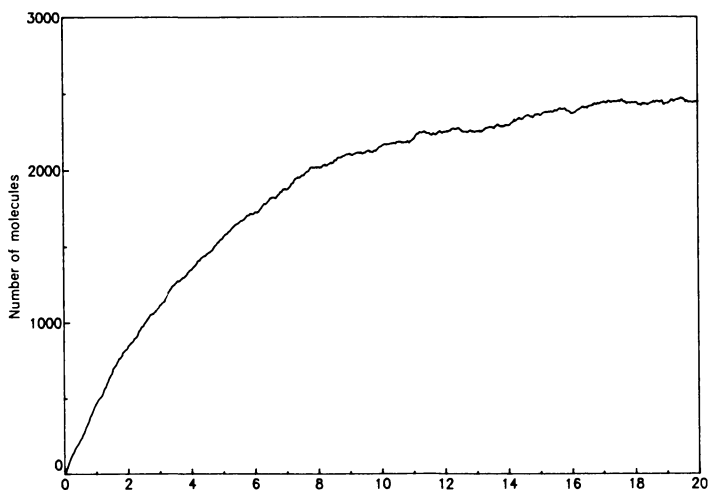
Fig. 1.1. Stochastic simulation of an isomerisation reaction $X \rightleftharpoons A$

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.
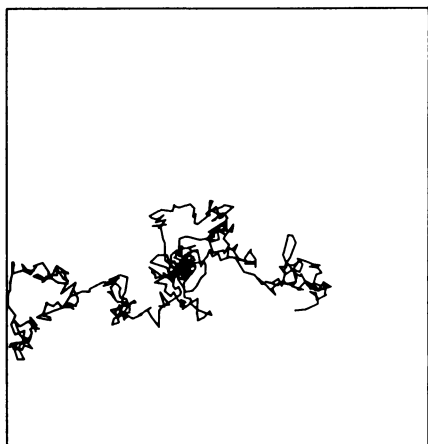
Fig. 1.2. Motion of a point undergoing Brownian motion

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 satisfies 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

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

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

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

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):

$$\int_{-\infty}^{\infty} \Delta\phi(\Delta)d\Delta, \quad \text{and} \quad \int_{0}^{\infty} \Delta^2\phi(\Delta)d\Delta . \tag{1.2.20}$$

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

$$X + A \rightarrow 2X \tag{1.3.1a}$$
$$X + Y \rightarrow 2Y \tag{1.3.1b}$$
$$Y \rightarrow B \tag{1.3.1c}$$

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

$$\frac{dx}{dt} = k_1ax - k_2xy \tag{1.3.2a}$$

$$\frac{dy}{dt} = k_2xy - k_3y . \tag{1.3.2b}$$

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.
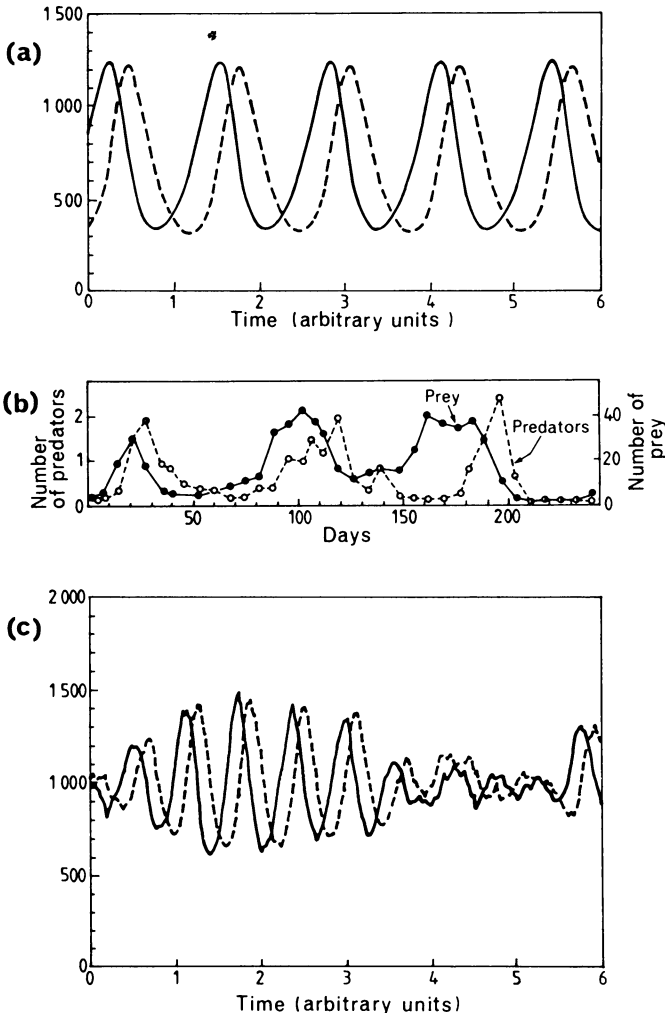
Fig. 1.3a-c. Time development in prey-predator systems. (a) Plot of solutions of the deterministic equations (1.3.2) (x = solid line, y = dashed line). (b) Data for a real prey-predator system. Here the predator is a mite (Eotetranychus sexmaculatus—dashed line) which feeds on oranges, and the prey is 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.

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

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

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

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

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

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

and letting $\Delta t \rightarrow 0$, $= \partial P(x, y, t)/\partial t$. 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 parti-cles 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

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

where $F(t-t_k)$ 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.
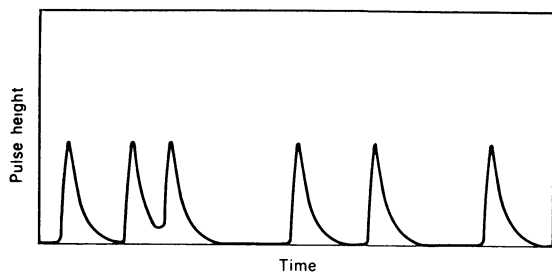
Fig. 1.4. Illustration of shot noise: identical electric pulses arrive at random times

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

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

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 $\Delta t \rightarrow 0$

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

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 $t \rightarrow \infty$, 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

$$F(t) = q\, e^{-\alpha t} \qquad (t > 0)$$
$$= 0 \qquad\qquad (t < 0) \tag{1.4.13}$$

so that (1.4.12) can be rewritten as

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

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

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

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)

$$\langle \mu(t)dt \rangle = \langle dN(t) \rangle = \lambda dt \qquad (1.4.17)$$

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

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

$$d\eta(t) = dN(t) - \lambda dt , \qquad (1.4.19)$$

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

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

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

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

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

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

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

$$\langle I^2(\infty) \rangle = (\lambda q/\alpha)^2 . \qquad (1.4.24)$$

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):

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

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 1st 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

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

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

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

with the corresponding inverse

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

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,

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

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

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

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

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

$$S(\omega) = R\,kT/[\pi(\omega^2\tau_C^2 + 1)] \tag{1.4.52}$$
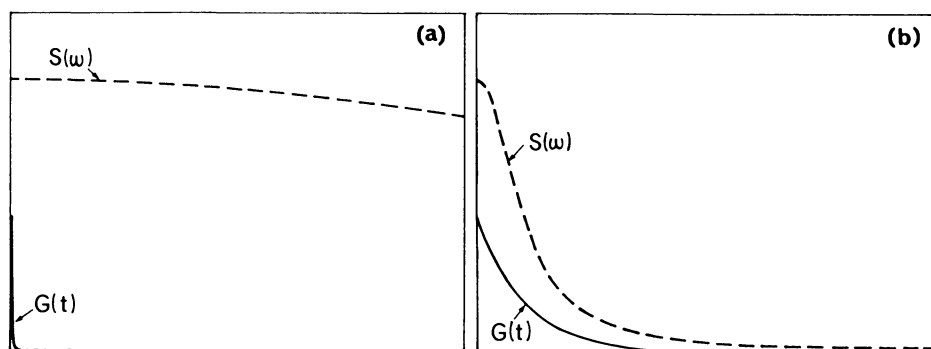
Fig. 1.5.a,b. Correlation Functions (————) and corresponding spectra (- - - - - -) for **(a)** short correlation time corresponding to an almost flat spectrum; **(b)** long correlation time, giving a quite rapidly decreasing spectrum

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

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

so that the autocorrelation function vanishes only for $\tau \gg \tau_{\mathrm{C}}$, 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 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

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

or a vector of real numbers

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

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 $A(r, \Delta V)$ that a molecule is within a volume element ΔV centred on a point 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 r, but to determine whether the particle is exactly at r is impossible. Thus, if we define the event $\omega(y)$ that the molecule is at point y, it makes sense to ask whether

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

and to assign a certain probability to the *set* $A(r, \Delta V)$, 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) $P(\varnothing) = 0.$ $\qquad$ (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 x_n at times t_n, $n = 0, 1, 2, ..., \infty$".

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.,

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

(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, x, the position in space. The probability of a molecule being in the set whose only element in 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 $\{x\}$—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 $\{x\}$.

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 x. Then we can introduce the random variable $F(x)$ which is a function of x, which takes on certain values for each x. In particular, the identity function of x, written $X(x)$ is of interest; it is given by

$$X(x) = x. \qquad (2.2.10)$$

We shall normally use capitals in this book to denote random variables and small letters 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 $X(\omega)$ 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 x anyway, i.e., we identify 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., X^2, $\sin(a \cdot X)$, 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.

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

$P(A|B) = P(A \cap B)/P(B)$ (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

$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A).$ (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_i such that

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

$$\bigcup_i B_i = \Omega \tag{2.3.5}$$

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

$$\bigcup_i (A \cap B_i) = A \cap (\bigcup_i B_i) = A \cap \Omega = A \tag{2.3.6}$$

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

$$\sum_i P(A \cap B_i) = P[\bigcup_i (A \cup B_i)] \tag{2.3.7}$$

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

and thus

$$\sum_i P(A \mid B_i) P(B_i) = P(A) \tag{2.3.9}$$

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

$$\sum_i P(A_i \cap B_j \cap C_k \ldots) = P(B_j \cap C_k \cap \ldots). \tag{2.3.10}$$

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 $P(A|B)$ 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

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

This requires

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

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, ..., 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 $X_1 \in A_1$,
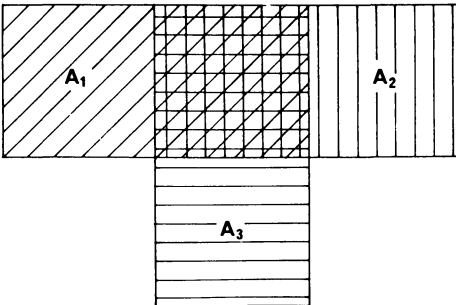
Fig. 2.1. Illustration of statistical independence in pairs, but not in threes. In the three sets $A_j \cap A_i$ is, in all cases, the central region. By appropriate choice of probabilities, we can arrange

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

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

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

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

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 $S(R_l)$, each containing only one point R_l, 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

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

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

for which one finds

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

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

and one finds approximately

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

In both these cases, 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 s is the vector $(s_1, s_2, ..., s_n)$, and X the vector of random variables $(X_1, X_2, ..., X_n)$, then the characteristic function (or moment generating function) is defined by

$$\phi(s) = \langle \exp(is \cdot X) \rangle = \int dx \, p(x) \exp(is \cdot x) . \tag{2.6.1}$$

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

i) $\phi(0) = 1$

ii) $\phi|(s)| \leqslant 1$

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

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

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

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

$$p(x) = (2\pi)^{-n} \int ds \, \phi(s) \exp(-ix \cdot s) \tag{2.6.3}$$

Because of this inversion formula, $\phi(s)$ determines $p(x)$ 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 ... are independent if and only if

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

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

$$Y = \sum_{i=1}^{n} X_i, \tag{2.6.6}$$

and the characteristic function of Y is

$$\phi_y(s) = \langle \exp{(isY)} \rangle, \tag{2.6.7}$$

then

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

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

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

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_i. 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_i 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:

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

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

$$\langle\!\langle X_i \rangle\!\rangle = \langle X_i \rangle \qquad (2.7.4)$$

$$\langle\!\langle X_i X_j \rangle\!\rangle = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle \qquad (2.7.5)$$

$$\langle\!\langle X_i X_j X_k \rangle\!\rangle = \langle X_i X_j X_k \rangle - \langle X_i X_j \rangle \langle X_k \rangle - \langle X_i \rangle \langle X_j X_k \rangle$$
$$- \langle X_i X_k \rangle \langle X_j \rangle + 2 \langle X_i \rangle \langle X_j \rangle \langle X_k \rangle \qquad (2.7.6)$$

(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 $\langle\!\langle X_1 X_2 X_3 \dots X_n \rangle\!\rangle$. The procedure is the following:

i) write a sequence of n dots $\dots\dots$;
ii) divide into $p + 1$ subsets by inserting angle brackets

$$\langle \dots \rangle \langle \dots \rangle \langle \dots\dots \rangle \dots \langle \dots \rangle ;$$

iii) distribute the symbols $X_1 \dots X_n$ 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 $C_p(X_1, X_2, \dots, X_n)$;

v) $\langle\!\langle X_1 X_2 \dots X_n \rangle\!\rangle = \sum_{p=0}^{n-1} (-1)^p p! C_p(X_1, X_2, \dots, X_n) .$ \qquad (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: $\langle\!\langle X_1 X_2 X_3 X_4 \rangle\!\rangle$

a) $p = 0$

only term is $\langle X_1 X_2 X_3 X_4 \rangle \equiv C_0(X_1 X_2 X_3 X_4)$.

b) $p = 1$

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

Term $\{\langle X_1 \rangle \langle X_2 X_3 X_4 \rangle + \langle X_2 \rangle \langle X_3 X_4 X_1 \rangle + \langle X_3 \rangle \langle X_4 X_1 X_2 \rangle$
$+ \langle X_4 \rangle \langle X_1 X_2 X_3 \rangle\} \equiv D_1$

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

Term $\langle X_1 X_2 \rangle \langle X_3 X_4 \rangle + \langle X_1 X_3 \rangle \langle X_2 X_4 \rangle + \langle X_1 X_4 \rangle \langle X_2 X_3 \rangle \equiv D_2$.

Hence,

$$D_1 + D_2 = C_1(X_1 X_2 X_3 X_4) .$$

c) $p = 2$

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

Term $\langle X_1 \rangle \langle X_2 \rangle \langle X_3 X_4 \rangle + \langle X_1 \rangle \langle X_3 \rangle \langle X_2 X_4 \rangle + \langle X_1 \rangle \langle X_4 \rangle \langle X_2 X_3 \rangle$
$+ \langle X_2 \rangle \langle X_3 \rangle \langle X_1 X_4 \rangle + \langle X_2 \rangle \langle X_4 \rangle \langle X_1 X_3 \rangle + \langle X_3 \rangle \langle X_4 \rangle \langle X_1 X_2 \rangle$
$= C_2(X_1 X_2 X_3 X_4) .$

d) $p = 3$

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

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

Hence,

$$\langle\!\langle X_1 X_2 X_3 X_4 \rangle\!\rangle = C_0 - C_1 + 2C_2 - 6C_3 \tag{2.7.8}$$

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

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

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

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

so that

$$\langle X \rangle = \int dx\, x\, p(x) = \bar{x} \qquad (2.8.2)$$

$$\langle XX^{\mathrm{T}} \rangle = \int dx\, xx^{\mathrm{T}}p(x) = \bar{x}\bar{x}^{\mathrm{T}} + \sigma \qquad (2.8.3)$$

and the characteristic function is given by

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

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 $\bar{x} = 0$, in which case the odd moments vanish and the even moments satisfy

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

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

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

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

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

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, ..., 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

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

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

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

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

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

and clearly, the *factorial moments*, defined by

$$\langle X^r \rangle_f = \langle x(x-1) \dots (x-r+1) \rangle , \qquad (2.8.14)$$

are given by

$$\langle X^r \rangle_f = \alpha^r , \qquad (2.8.15)$$

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

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

which is related to the characteristic function by

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

The generating function has the useful property that

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

For the Poisson distribution we have

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

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* $\langle\!\langle X^r \rangle\!\rangle_f$ by

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

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

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

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

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

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 $X(t)$ exists. We can measure values x_1, x_2, x_3, ..., etc., of $X(t)$ at times t_1, t_2, t_3, ... and we assume that a set of joint probability densities exists

$$p(x_1, t_1; x_2, t_2; x_3, t_3; ...) \tag{3.1.1}$$

which describe the system completely.

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

$$p(x_1, t_1,; x_2, t_2; ... | y_1, \tau_1; y_2, \tau_2; ...)$$
$$= p(x_1, t_1; x_2, t_2; ...; y_1, \tau_1; y_2, \tau_2; ...)/p(y_1, \tau_1; y_2, \tau_2; ...). \tag{3.1.2}$$

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.,

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

The concept of an evolution equation leads us to consider the conditional probabilities as predictions of the future values of $X(t)$ (i.e., x_1, x_2, ... at times t_1, t_2, ...), given the knowledge of the past (values y_1, y_2, ..., at times τ_1, τ_2, ...).

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:

$$p(x_1, t_1; x_2, t_2; x_3, t_3; ...) = \prod_i p(x_i, t_i) \tag{3.1 4}$$

which means that the value of X at time t is completely independent of its values in the past (or future). An even more special case occurs when the $p(x_i, t_i)$ 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

$$p(x_1, t_1; x_2, t_2; x_3, t_3; \ldots x_n, t_n)$$
$$= p(x_1, t_1 | x_2, t_2)p(x_2, t_2 | x_3, t_3)p(x_3, t_3 | x_4, t_4) \ldots \tag{3.2.}$$
$$\ldots p(x_{n-1}, t_{n-1} | x_n, t_n)p(x_n, t_n)$$

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.,

$$\sum_B P(A \cap B \cap C \ldots) = P(A \cap C \ldots); \tag{3.2.}$$

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

$$p(x_1, t_1) = \int dx_2 \, p(x_1, t_1; x_2, t_2)$$
$$= \int dx_2 \, p(x_1, t_1 | x_2, t_2)p(x_2, t_2). \tag{3.2.}$$

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

$$p(x_1, t_1 | x_3, t_3) = \int dx_2 \, p(x_1, t_1; x_2, t_2 | x_3, t_3)$$
$$= \int dx_2 \, p(x_1, t_1 | x_2, t_2; x_3, t_3) p(x_2, t_2 | x_3, t_3). \qquad (3.2.7)$$

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

$$p(x_1, t_1 | x_3, t_3) = \int dx_2 \, p(x_1, t_1 | x_2, t_2) p(x_2, t_2 | x_3, t_3) \qquad (3.2.8)$$

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 $p(x_i, t_i | x_j, t_j)$ 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 $p(x_1, t_1 | x_2, t_2)$.

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 $N = (N_1, N_2, N_3 \ldots)$, where the N_i are random variables which take on integral values. Clearly, we now replace

$$\int dx \leftrightarrow \sum_n \qquad (3.2.9)$$

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

$$P(n_1, t_1 | n_3, t_3) = \sum_{n_2} P(n_1, t_1 | n_2, t_2) \, P(n_2, t_2 | n_3, t_3) \, . \qquad (3.2.10)$$

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

3.2.3 More General Measures

A more general formulation would assume a measure $d\mu(x)$ instead of dx where a variety of choices can be made. For example, if $\mu(x)$ is a step function with steps at integral values of 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

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

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

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

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

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

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

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

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

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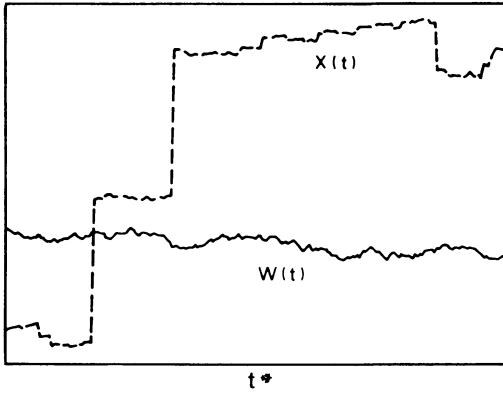
Fig. 3.1. Illustration of sample paths of the Cauchy process $X(t)$ (- - - - -) and Brownian motion $W(t)$ (———)

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* $\varepsilon > 0$:

i) $\lim\limits_{\Delta t \to 0} p(x, t + \Delta t \,|\, z, t)/\Delta t = W(x \,|\, z, t)$ \hfill (3.4.1)

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

ii) $\lim\limits_{\Delta t \to 0} \dfrac{1}{\Delta t} \int\limits_{|x-z|<\varepsilon} dx \,(x_i - z_i) p(x, t + \Delta t \,|\, z, t) = A_i(z, t) + O(\varepsilon)$; \hfill (3.4.2)

iii) $\lim\limits_{\Delta t \to 0} \dfrac{1}{\Delta t} \int\limits_{|x-z|<\varepsilon} dx \,(x_i - z_i)(x_j - z_j) p(x, t + \Delta t \,|\, z, t) = B_{ij}(z, t) + O(\varepsilon)$; \hfill (3.4.3)

the last two being uniform in 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

$$\lim\limits_{\Delta t \to 0} \frac{1}{\Delta t} \int\limits_{|x-z|<\varepsilon} dx \,(x_i - z_i)(x_j - z_j)(x_k - z_k)\, p(x, t + \Delta t \,|\, z, t)$$

$$\equiv C_{ijk}(z, t) + O(\varepsilon) . \hfill (3.4.4)$$

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,

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

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,

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

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

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

where we have (again by the twice continuous differentiability)

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

Now substitute in (3.4.9):

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

[notice that since $p(x, t + \Delta t | z, t)$ is a probability, the integral over 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]

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

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

$$\left| \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} dx \, |x - z|^2 R(x, z) p(x, t + \Delta t | z, t) \right|$$
$$\leqslant \left[\frac{1}{\Delta t} \int_{|x-z|<\varepsilon} dx \, |x - z|^2 p(x, t + \Delta t | z, t) \right] \underset{|x-z|<\varepsilon}{\text{Max}} |R(x, z)| \tag{3.4.14}$$
$$\to \left[\sum_i B_{ii}(z, t) + O(\varepsilon) \right] \left\{ \underset{|x-z|<\varepsilon}{\text{Max}} |R(x, z)| \right\}.$$

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

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

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

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

Notice, however, that we use the definition

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

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

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

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

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

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

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

It is clear that by definition we have

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

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_i and B_{ij} and by our reasoning above, one cannot assume that this is possible on the boundary of the region. Hence, let us choose $f(z)$ to be arbitrary but nonvanishing only in an arbitrary region R' entirely contained in R. We can then deduce that for all z in *the interior* of R,

$$
\begin{aligned}
\partial_t p(z, t \,|\, y, t') = &-\sum_i \frac{\partial}{\partial z_i} [A_i(z, t) p(z, t \,|\, y, t')] \\
&+ \sum_{i,j} \frac{1}{2} \frac{\partial^2}{\partial z_i \partial z_j} [B_{ij}(z, t) p(z, t \,|\, y, t')] \\
&+ \int dx \, [W(z \,|\, x, t) p(x, t \,|\, y, t') - W(x \,|\, z, t) p(z, t \,|\, y, t')] .
\end{aligned} \tag{3.4.22}
$$

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 $A(x, t)$, $B(x, t)$ (which must be positive semi-definite), and $W(x \,|\, y, t)$ (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,*

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

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

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

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

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

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

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

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

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

where $\langle \eta(t) \rangle = 0$ \hfill (3.5.11)

$$\langle \eta(t)\eta(t)^T \rangle = \underline{B}(y, t) . \tag{3.5.12}$$

It is easy to see that this picture gives

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

ii) sample paths which are nowhere differentiable, 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 $x(y, t') = y$, $\tag{3.5.15}$

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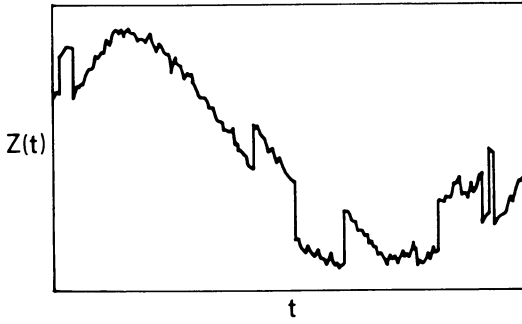
Fig. 3.2. Illustration of a sample path of a general Markov process, in which drift, diffusion and jumps exist

Z(t)

t

Z(t)

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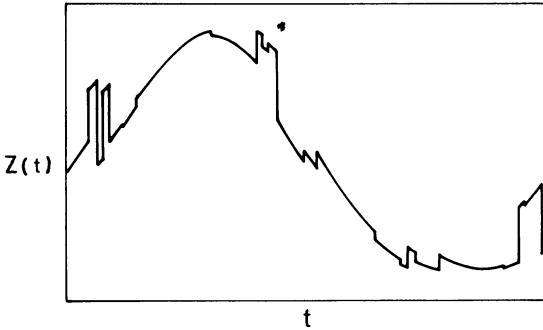
Fig. 3.3. Sample path of a Markov process with only drift and jumps

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 $p(x, t \mid y, t')$.

We consider

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

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

$$- p(x, t \mid z, t' + \Delta t')] \tag{3.6.2}$$

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

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

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

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

$$\frac{\partial p(x, t | y, t')}{\partial t'} = - \sum_i A_i(y, t') \frac{\partial p(x, t | y, t')}{\partial y_i} - \frac{1}{2} \sum_{ij} B_{ij}(y, t') \frac{\partial^2 p(x, t | y, t')}{\partial y_i \partial y_j}$$

$$+ \int dz\, W(z | y, t') \left[p(x, t | y, t') - p(x, t | z, t') \right] \qquad (3.6.4)$$

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

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

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.,

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

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

$$p_s(x) \qquad (3.7.2)$$

and the two-time joint probability as

$$p_s(x_1, t_1 - t_2; x_2, 0). \qquad (3.7.3)$$

Finally, the conditional probability can also be written as

$$p_s(x_1, t_1 - t_2 | x_2, 0). \qquad (3.7.4)$$

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

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

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

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

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

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

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

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

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 auto-correlation 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

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

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

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

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

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

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

$$p(x, t) = p_s(x, t \,|\, x_0, t_0) \qquad \text{and} \tag{3.7.24}$$

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

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 $t_0 \to -\infty$,

$$p(x, t) \to p_s(x)$$

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 x is by some external agency fixed to have a value 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

$$0 = -\sum_i \frac{\partial}{\partial z_i} [A_i(z)p(z, t \,|\, y, t')] + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial z_i \partial z_j} [B_{ij}(z)p(z, t \,|\, y, t')]$$

$$+ \int dx \, [W(z \,|\, x)p(x, t \,|\, y, t') - W(x \,|\, z)p(z, t \,|\, y, t')], \tag{3.7.26}$$

where we have used the fact that the process is homogeneous to note that A, 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, 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

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

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

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

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)

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

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

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

which can be rearranged to give

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

Similarly, we may calculate

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

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.,

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

The constant must equal one since $p_1(x, t)$ and $p_s(x)$ 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,

$p_s(x) \neq 0$ with probability 1

$W(x|x') \neq 0$ with probability 1, $\qquad (3.7.40)$

$B_{ij}(x)$ positive definite with probability 1,

that any solution of the differential Chapman-Kolmogorov equation approaches the stationary solution $p_s(x)$ 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, $W(x|x') = 0$ if x and x' are not both in R_1 or R_2. Then it is possible to have $dK/dt = 0$ if

$$p_1(x, t) = \lambda_1 p_s(x) \qquad x \in R_1 \qquad (3.7.41)$$
$$= \lambda_2 p_s(x) \qquad x \in R_2$$

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) $p_s(x)$ Vanishes in Some Definite Region

If we have

$$p_s(x) = 0 \qquad x \in R_1 \qquad (3.7.42)$$
$$\neq 0 \qquad x \in R_2$$

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

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

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

$$\langle X(t) | [x_0, t_0] \rangle = \int dx \, x \, p(x, t | x_0, t_0) \,, \tag{3.7.44}$$

then the autocorrelation matrix

$$\langle X(t) X(t_0)^\mathrm{T} \rangle = \int dx \, dx_0 \, xx_0^\mathrm{T} p(x, t; x_0, t_0) \tag{3.7.45}$$

$$= \int dx_0 \, \langle X(t) | [x_0, t_0] \rangle x_0^\mathrm{T} p(x_0, t_0) \,. \tag{3.7.46}$$

Thus we see that (3.7.44) defines the mean of $X(t)$ under the condition that X had the value 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 x_0^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 $\langle X(t) | [x_0, t_0] \rangle$ is a uniquely defined quantity, since the knowledge of 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

$$p_s(x_1, t_1; x_2, t_2; \ldots x_n, t_n), \tag{3.7.47}$$

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

$$p_a(x_1, t_1; x_2, t_2; \ldots x_n, t_n) = p_s(x_1, t_1, x_2, t_2; \ldots x_n, \; t_n | a, 0). \tag{3.7.48}$$

Then for this process we note that

$$\langle X(t) | [x_0, t_0] \rangle_a = \int dx \, x \, p_s(x, t | x_0, t_0; a, 0) \tag{3.7.49}$$

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

$$\lim_{\tau \to \infty} p_s(x, t + \tau | x_0, t_0 + \tau; a, 0) = p_s(x, t - t_0 | x_0, 0) \tag{3.7.50}$$

so that we will also have a stationary conditional average of x

$$\langle X(t) | [x_0, t_0] \rangle_s = \lim_{\tau \to \infty} \langle X(t + \tau) | [x_0, t_0 + \tau] \rangle_a \tag{3.7.51}$$

and the stationary autocorrelation matrix is given by

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

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

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

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

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

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

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

$$\frac{\partial}{\partial t} p(w, t \,|\, w_0, t_0) = \frac{1}{2} \frac{\partial^2}{\partial w^2} p(w, t \,|\, w_0, t_0) \,. \tag{3.8.1}$$

Utilising the initial condition

$$p(w, t_0 \,|\, w_0, t_0) = \delta(w - w_0) \tag{3.8.2}$$

on the conditional probability, we solve (3.8.1) by use of the characteristic function

$$\phi(s, t) = \int dw \, p(w, t \,|\, w_0, t_0) \exp(isw) \tag{3.8.3}$$

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

$$\phi(s, t_0) = \exp(isw_0)$$

so that

$$\phi(s, t) = \exp\left[isw_0 - \tfrac{1}{2} s^2(t - t_0)\right]. \tag{3.8.6}$$

Performing the Fourier inversion, we have the solution to (3.8.1):

$$p(w, t \,|\, w_0, t_0) = [2\pi(t - t_0)]^{-1/2} \exp\left[-(w - w_0)^2/2(t - t_0)\right].$$ (3.8.7)

This represents a Gaussian, with

$$\langle W(t) \rangle = w_0$$ (3.8.8)

$$\langle [W(t) - w_0]^2 \rangle = t - t_0 ,$$ (3.8.9)

so that an initially sharp distribution spreads in time, as graphed in Fig.3.4.
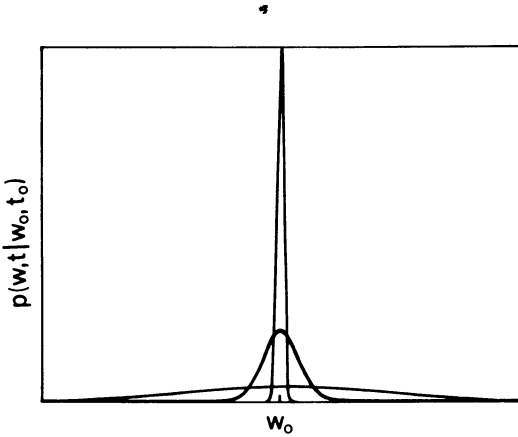
Fig. 3.4. Wiener process: spreading of an initially sharp distribution $p(w, t \,|\, w_0, t_0)$ with increasing time $t - t_0$

A multivariate Wiener process can be defined as

$$W(t) = [W_1(t), W_2(t), \dots , W_n(t)]$$ (3.8.10)

which satisfies the multivariable Fokker-Planck equation

$$\frac{\partial}{\partial t} p(\boldsymbol{w}, t \,|\, \boldsymbol{w}_0, t_0) = \frac{1}{2} \sum_i \frac{\partial^2}{\partial w_i^2} p(\boldsymbol{w}, t \,|\, \boldsymbol{w}_0, t_0)$$ (3.8.11)

whose solution is

$$p(\boldsymbol{w}, t \,|\, \boldsymbol{w}_0, t_0) = [2\pi(t - t_0)]^{-n/2} \exp\left[-(\boldsymbol{w} - \boldsymbol{w}_0)^2/2(t - t_0)\right],$$ (3.8.12)

a multivariate Gaussian with

$$\langle W(t) \rangle = \boldsymbol{w}_0$$ (3.8.13)

and

$$\langle [W_i(t) - w_{0i}] [W_j(t) - w_{0j}] \rangle = (t - t_0)\delta_{ij} .$$ (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
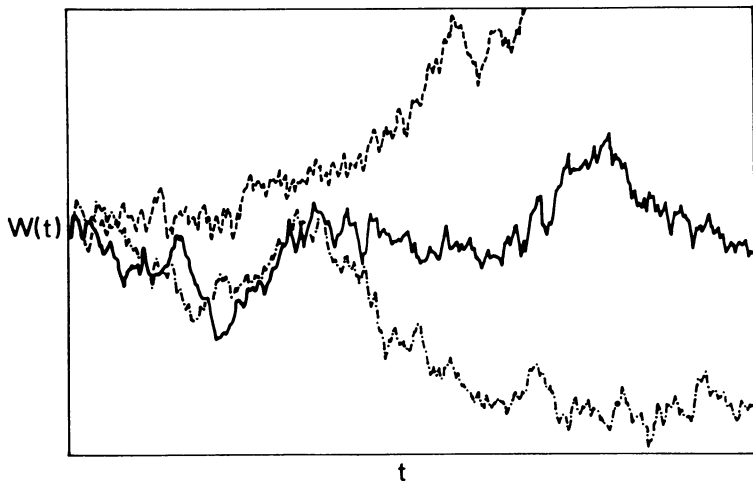
Fig. 3.5. Three simulated sample paths of the Wiener process, illustrating their great variability

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

$$p(w_n, t_n; w_{n-1}, t_{n-1}; w_{n-2}, t_{n-2}; \ldots ; w_0, t_0)$$
$$= \prod_{i=0}^{n-1} p(w_{i+1}, t_{i+1} | w_i, t_i) p(w_0, t_0) , \qquad (3.8.17)$$

and using the explicit form of the conditional probabilities (3.8.7), we see that

$$p(w_n, t_n; w_{n-1}, t_{n-1}; w_{n-2}, t_{n-2}; \ldots; w_0, t_0)$$
$$= \prod_{i=0}^{n-1} \{[2\pi(t_{i+1} - t_i)]^{-1/2} \exp[-(w_{i+1} - w_i)^2/2(t_{i+1} - t_i)]\} p(w_0, t_0) . \qquad (3.8.18)$$

If we define the variables

$$\Delta W_i \equiv W(t_i) - W(t_{i-1}) \qquad (3.8.19)$$

$$\Delta t_i \equiv t_i - t_{i-1} , \qquad (3.8.20)$$

then the joint probability density for the ΔW_i is

$$p(\Delta w_n; \Delta w_{n-1}; \Delta w_{n-2}; \ldots \Delta w_1; w_0)$$
$$= \prod_{i=1}^{n} \{(2\pi\Delta t_i)^{-1/2} \exp(-\Delta w_i^2/2\Delta t_i)\} p(w_0, t_0) \qquad (3.8.21)$$

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

$$\langle W(t)W(s) | [w_0, t_0] \rangle = \int dw_1 dw_2 \, w_1 w_2 p(w_1, t; w_2, s | w_0, t_0), \qquad (3.8.22)$$

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

$$\partial_t P(n, t|n', t') = d[P(n + 1, t|n', t') + P(n - 1, t|n', t') \\ - 2P(n, t|n't')]. \tag{3.8.26}$$

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

$$P(n, (N + 1)\tau|n', N'\tau) = \tfrac{1}{2}[P(n + 1, N\tau|n', N'\tau) \\ + P(n - 1, N\tau|n', N'\tau)] . \tag{3.8.27}$$

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

$$G(s, t) = \langle e^{ins} \rangle = \sum_n P(n, t|n', t')e^{ins} \tag{3.8.29}$$

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

$$G(s, (N+1)\tau) = \frac{1}{2}(e^{is} + e^{-is})G(s, N\tau) . \tag{3.8.31}$$

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

$$G_1(s, t) = \exp\left[(e^{is} + e^{-is} - 2)td\right] , \tag{3.8.33}$$

and to (3.8.31)

$$G_2(s, N\tau) = [\frac{1}{2}(e^{is} + e^{-is})]^N \tag{3.8.34}$$

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

$$P_1(n, t|0, 0) = e^{-2td}I_n(4td) \tag{3.8.38}$$

$$P_2(n, N\tau \,|\, 0, 0) = (\tfrac{1}{2})^N N! \left[\left(\frac{N-n}{2}\right)! \left(\frac{N+n}{2}\right)! \right]^{-1} . \tag{3.8.39}$$

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 $\quad D = \lim_{l \to 0} (l^2 d).$ \hfill (3.8.43)

This is the characteristic function of a Gaussian (Sect.2.8.1) of the form

$$p(x, t \,|\, 0, 0) = (4\pi Dt)^{-1/2} \exp(-x^2/4Dt) \tag{3.8.44}$$

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

$$\partial_t p(x, t \,|\, 0, 0) = (l^2 d) \partial_x^2 p(x, t \,|\, 0, 0) . \tag{3.8.46}$$

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):

$$\partial_t G(s, t) = d[\exp(is) - 1]G(s, t) \tag{3.8.50}$$

with the solution

$$G(s, t) = \exp\{td[\exp(is) - 1] \tag{3.8.51}$$

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 $l \rightarrow 0$, 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

$$\phi(s, t) = G(ls, t) \simeq \exp[t(ivs - s^2 D/2)] \tag{3.8.57}$$

where, as in the previous section,

$$D = l^2 d .$$

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

$$P(n - 1, t | 0, 0) = d\, p(x - l, t | 0, 0)$$
$$\simeq d\, p(x, t | 0, 0) - l d\, \partial_x p(x, t | 0, 0) + \tfrac{1}{2} l^2 d\, \partial_x^2 p(x, t | 0, 0). \tag{3.8.60}$$

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 $p(x, t | x_0, 0)$. This is the *Ornstein-Uhlenbeck process* [3.5]. The equation for the characteristic function

$$\phi(s) = \int_{-\infty}^{\infty} e^{isx} p(x, t | x_0, 0) dx \qquad \text{is} \tag{3.8.62}$$

$$\partial_t \phi + ks \partial_s \phi = -\tfrac{1}{2} Ds^2 \phi . \tag{3.8.63}$$

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)

$$\frac{dt}{1} = \frac{ds}{ks} = -\frac{d\phi}{\tfrac{1}{2} Ds^2 \phi}, \tag{3.8.65}$$

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

$$p(x, 0 | x_0, 0) = \delta(x - x_0) \tag{3.8.69}$$

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

$$\frac{1}{p}\partial_x p = -\frac{2kx}{D} \tag{3.8.76}$$

so that $p_s(x) = (\pi D/k)^{-1/2} \exp(-kx^2/D)$ \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

$$\lim_{t_0 \to -\infty} p(x_2, s \,|\, x_0, t_0) = p_s(x_2) = (\pi D/k)^{-1/2} \exp\left(- kx_2^2/D\right) . \tag{3.8.81}$$

and by straightforward substitution and integration and noting that the stationary mean is zero, we get

$$\langle X(t)X(s)\rangle_s = \langle X(t), X(s)\rangle_s = \frac{D}{2k} \exp\left(-k \,|\, t - s \,|\right) . \tag{3.8.82}$$

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

$$\left.\begin{array}{l} \partial_t P(a, t \,|\, x, t_0) = -\lambda P(a, t \,|\, x, t_0) + \mu P(b, t \,|\, x, t_0) \\[1mm] \partial_t P(b, t \,|\, x, t_0) = \lambda P(a, t \,|\, x, t_0) - \mu P(b, t \,|\, x, t_0) \end{array}\right\} \tag{3.8.85}$$

for which the solution can simply be found by noting that

$$P(a, t \,|\, x, t_0) + P(b, t \,|\, x, t_0) = 1$$

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

$$P(x', t_0 \,|\, x, t_0) = \delta_{x, x'} \,, \tag{3.8.86}$$

$$\lambda P(a, t \,|\, x, t_0) - \mu P(b, t \,|\, x, t_0) = \exp\left[-(\lambda + \mu)(t - t_0)\right](\lambda \delta_{a, x} - \mu \delta_{b, x}) \tag{3.8.87}$$

so that

$$\left.\begin{aligned}
P(a, t \,|\, x, t_0) &= \frac{\mu}{\lambda + \mu} + \exp\left[-(\lambda + \mu)(t - t_0)\right]\left(\frac{\lambda}{\lambda + \mu} \delta_{a, x} - \frac{\mu}{\lambda + \mu} \delta_{b, x}\right) \\
P(b, t \,|\, x, t_0) &= \frac{\lambda}{\lambda + \mu} - \exp\left[-(\lambda + \mu)(t - t_0)\right]\left(\frac{\lambda}{\lambda + \mu} \delta_{a, x} - \frac{\mu}{\lambda + \mu} \delta_{b, x}\right)
\end{aligned}\right\} \tag{3.8.88}$$

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:

$$\langle X(t) \,|\, [x_0, t_0]\rangle = \sum xP(x, t \,|\, x_0, t_0)$$

$$= \frac{a\mu + b\lambda}{\mu + \lambda} + \exp\left[-(\lambda + \mu)(t - t_0)\right]\left(x_0 - \frac{a\mu + b\lambda}{\mu + \lambda}\right) \tag{3.8.90}$$

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

$$\langle X(t)X(s)\rangle_s = \langle X\rangle_s^2 + \exp\left[-(\lambda + \mu)(t - s)\right](\langle X^2\rangle_s - \langle X\rangle_s^2) \qquad (3.8.95)$$

$$= \left(\frac{a\mu + b\lambda}{\mu + \lambda}\right)^2 + \exp\left[-(\lambda + \mu)(t - s)\right]\frac{(a - b)^2\mu\lambda}{(\lambda + \mu)^2}. \qquad (3.8.96)$$

Hence,

$$\langle X(t), X(s)\rangle_s = \langle X(t)X(s)\rangle_s - \langle X\rangle_s^2 = \exp\left[-(\lambda + \mu)|t - s|\right]\frac{(a-b)^2\mu\lambda}{(\lambda + \mu)^2}. \quad (3.8.97)$$

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)
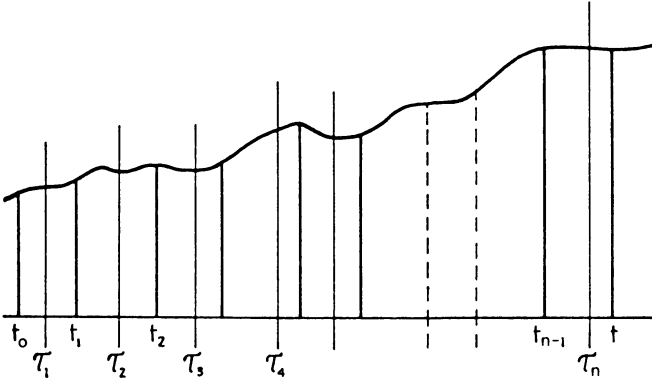
Fig. 4.1. Partitioning of the time interval used in the definition of stochastic integration

$$t_0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_{n-1} \leqslant t , \qquad (4.2.1)$$

and define intermediate points τ_i such that

$$t_{i-1} \leqslant \tau_i \leqslant t_i . \qquad (4.2.2)$$

The stochastic integral $\int_{t_0}^{t} G(t')dW(t')$ is defined as a limit of the partial sums.

$$S_n = \sum_{i=1}^{n} G(\tau_i)[W(t_i) - W(t_{i-1})] . \qquad (4.2.3)$$

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

$$\tau_i = \alpha t_i + (1 - \alpha)t_{i-1} \quad (0 < \alpha < 1), \tag{4.2.7}$$

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

$$\int_{t_0}^{t} G(t')dW(t') = \text{ms -lim}_{n \to \infty} \left\{ \sum_{i=1}^{n} G(t_{i-1})[W(t_i) - W(t_{i-1})] \right\}. \tag{4.2.10}$$

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,

$$\langle[\sum_i (W_i - W_{i-1})^2 - (t - t_0]^2\rangle$$

$$= \langle\sum_i (W_i - W_{i-1})^4 + 2 \sum_{i>j} (W_i - W_{i-1})^2 (W_j - W_{j-1})^2$$

$$- 2(t - t_0) \sum_i (W_i - W_{i-1})^2 + (t - t_0)^2\rangle. \qquad (4.2.15)$$

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,

$$\langle(W_i - W_{i-1})^2 (W_j - W_{j-1})^2\rangle = (t_i - t_{i-1}) (t_j - t_{j-1}) \qquad (4.2.16)$$

and also [using (2.8.6)]

$$\langle(W_i - W_{i-1})^4\rangle = 3 \langle(W_i - W_{i-1})^2\rangle^2 = 3(t_i - t_{i-1})^2 \qquad (4.2.17)$$

which combined with (4.2.16) gives

$$\langle[\sum_i (W_i - W_{i-1}) - (t - t_0)]^2\rangle = 2 \sum_i (t_i - t_{i-i})^2 \qquad (4.2.18)$$

$$+ \sum_{i,j} [(t_i - t_{i-1}) - (t - t_0)] [(t_j - t_{j-1}) - (t - t_0)]$$

$$= 2 \sum_i (t_i - t_{i-1})^2$$

$$\rightarrow 0 \quad \text{as} \quad n \rightarrow \infty.$$

Thus,

$$\text{ms-lim}_{n\rightarrow\infty} \sum_i (W_i - W_{i-1})^2 = t - t_0 \qquad (4.2.19)$$

by definition of the mean square limit, so

$$\boxed{\int_{t_0}^{t} W(t')dW(t') = \tfrac{1}{2}[W(t)^2 - W(t_0)^2 - (t - t_0)].}$$

$$(4.2.20)$$

Comments

i) $\langle\int_{t_0}^{t} W(t)dW(t)\rangle = \tfrac{1}{2} [\langle W(t)^2\rangle - \langle W(t_0)^2\rangle - (t - t_0)] = 0 \qquad (4.2.21)$

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

$$S \int_{t_0}^{t} W(t') \, dW(t')$$

$$= \underset{n \to \infty}{\text{ms-lim}} \sum_{i=1}^{n} \frac{W(t_i) + W(t_{i-1})}{2} [W(t_i) - W(t_{i-1})] \tag{4.2.22}$$

$$= \tfrac{1}{2} [W(t)^2 - W(t_0)^2] \,. \tag{4.2.23}$$

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) $\int\limits_{t_0}^{t} dt' G(t')$

$\left.\vphantom{\int\int}\right\}$ when $G(t)$ is itself a nonanticipating function

v) $\int\limits_{t_0}^{t} dW(t') G(t')$

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

$$\int\limits_{t_0}^{t} [dW(t')]^{2+N} G(t') \equiv \underset{n\to\infty}{\text{ms-lim}} \sum_i G_{i-1} \Delta W_i^{2+N}$$

$$= \int\limits_{t_0}^{t} dt' G(t') \text{ for } \qquad N = 0 \qquad\qquad (4.2.25)$$

$$= 0 \qquad \text{for } N > 0$$

for an arbitrary *nonanticipating function* $G(t)$.

The proof is quite straightforward. For $N = 0$, let us define

$$I = \lim_{n\to\infty} \langle [\sum_i G_{i-1}(\Delta W_i^2 - \Delta t_i)]^2 \rangle \qquad\qquad (4.2.26)$$

$$= \lim_{n\to\infty} \{\langle \underbrace{\sum_i (G_{i-1})^2 (\Delta W_i^2 - \Delta t_i)^2}_{\text{Independent}} + \sum_{i>j} \underbrace{2 G_{i-1} G_{j-1}(\Delta W_j^2 - \Delta t_j)(\Delta W_i^2 - \Delta t_i)}_{\text{Independent}} \rangle \}. \qquad (4.2.27)$$

The horizontal braces indicate factors which are statistically independent of each other because of the properties of the Wiener process, and because the G_i are values of a nonanticipating function which are independent of all ΔW_i for $j > i$.

Using this independence, we can factorise the means, and also using

i) $\langle \Delta W_i^2 \rangle = \Delta t_i$

ii) $\langle (\Delta W_i^2 - \Delta t_i)^2 \rangle = 2 \Delta t_i^2$ (from Gaussian nature of ΔW_i),

we find

$$I = 2 \lim_{n\to\infty} [\sum_i \Delta t_i^2 \langle (G_{i-1})^2 \rangle] . \qquad\qquad (4.2.28)$$

Under reasonably mild conditions on $G(t)$ (e.g., boundedness), this means that

$$\text{ms-lim}_{n\to\infty} \left(\sum_i G_{i-1}\Delta W_i^2 - \sum_i G_{i-1}\Delta t_i\right) = 0 \tag{4.2.29}$$

and since

$$\text{ms-lim}_{n\to\infty} \sum_i G_{i-1}\Delta t_i = \int_{t_0}^t dt' G(t') , \tag{4.2.30}$$

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_i is independent of G_{i-1}. In the Stratonovich integral,

$$\Delta W_i = W(t_i) - W(t_{i-1}) \tag{4.2.33}$$

$$G_{i-1} = G[\tfrac{1}{2}(t_i + t_{i-1})] \tag{4.2.34}$$

and although $G(t)$ is nonanticipating, this is not sufficient to guarantee the independence of ΔW_i and G_{i-1} as thus defined.

iv) By similar methods one can prove that

$$\int_{t_0}^t G(t')dt'dW(t') \equiv \text{ms-lim}_{n\to\infty} \sum G_{i-1}\Delta W_i\Delta t_i = 0 \tag{4.2.35}$$

and similarly for higher powers. The simplest way of characterising these results is to say that $dW(t)$ is an infinitesimal order of $\tfrac{1}{2}$ 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

$$\int_{t_0}^{t} W(t')^n dW(t') = \frac{1}{n+1} [W(t)^{n+1} - W(t_0)^{n+1}] - \frac{n}{2} \int_{t_0}^{t} W(t)^{n-1} dt. \tag{4.2.37}$$

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,

$$d\{\exp[W(t)]\} = \exp[W(t) + dW(t)] - \exp[W(t)]$$
$$= \exp[W(t)][dW(t) + \tfrac{1}{2} dW(t)^2]$$
$$= \exp[W(t)][dW(t) + \tfrac{1}{2} dt] \tag{4.2.38}$$

or more generally,

$$df[W(t), t] = \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} (dt)^2 + \frac{\partial f}{\partial W} dW(t) + \frac{1}{2} \frac{\partial^2 f}{\partial W} [dW(t)]^2$$
$$+ \frac{\partial^2 f}{\partial W \partial t} dt dW(t) + \dots$$

and we use $(dt)^2 \to 0$

$dt dW(t) \to 0$ [Sect. 4.2.5, comment (iv)]

$[dW(t)]^2 = dt$

and all higher powers vanish, to arrive at

$$df[W(t), t] = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W^2}\right) dt + \frac{\partial f}{\partial W} dW(t) . \qquad (4.2.39)$$

e) Mean Value Formula

For nonanticipating $G(t)$,

$$\left\langle \int_{t_0}^{t} G(t')dW(t')\right\rangle = 0 . \qquad (4.2.40)$$

For, since $G(t)$ is nonanticipating, in the definition of the stochastic integral,

$$\left\langle \sum_i G_{i-1}\Delta W_i \right\rangle = \sum_i \langle G_{i-1}\rangle\langle \Delta W_i\rangle = 0 \qquad (4.2.41)$$

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_{i-1} is chosen in the middle of the interval, and can be correlated with ΔW_i.

f) Correlation Formula

If $G(t)$ and $H(t)$ are arbitrary continuous nonanticipating functions,

$$\left\langle \int_{t_0}^{t} G(t')dW(t') \int_{t_0}^{t} H(t')dW(t')\right\rangle = \int_{t_0}^{t}dt \,\langle G(t')H(t')\rangle . \qquad (4.2.42)$$

Proof. Notice that

$$\left\langle \sum_i G_{i-1}\Delta W_i \sum_j H_{j-1}\Delta W_j \right\rangle = \left\langle \sum_i G_{i-1}H_{i-1}(\Delta W_i)^2\right\rangle$$
$$+ \left\langle \sum_{i>j}(G_{i-1}H_{j-1} + G_{j-1}H_{i-1})\Delta W_j\Delta W_i\right\rangle . \qquad (4.2.43)$$

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

$$\langle \Delta W_i^2\rangle = \Delta t_i \qquad (4.2.44)$$

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

$$dW(t) \rightarrow \xi(t)dt , \qquad (4.2.45)$$

it is clear that if $F(t)$ and $G(t)$ are nonanticipating, $\xi(t)$ is independent of them, and we get

$$\int_{t_0}^{t} dt' \int_{t_0}^{t} ds' \, \langle G(t')H(s')\xi(t')\xi(s')\rangle = \int_{t_0}^{t}\int_{t_0}^{t} dt'ds' \langle G(t')H(s')\rangle \langle \xi(t')\xi(s')\rangle$$

$$= \int_{t_0}^{t} dt' \langle G(t')H(t')\rangle \qquad (4.2.46)$$

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

$$I_1 = \int_{t_1}^{t_2} dt \, f(t)\delta(t - t_1) \qquad (4.2.47)$$

or

$$I_2 = \int_{t_1}^{t_2} dt \, f(t)\delta(t - t_2) \qquad (4.2\,48)$$

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

$$I_1 = f(t_1) \qquad (4.2.49)$$

$$I_2 = 0 \qquad (4.2.50)$$

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

$$\langle \int_{t_0}^{t} G(t')dW(t') \, [\int_{t_0}^{t'} H(s')dW(s')]\rangle = 0 \qquad (4.2.51)$$

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

$$\int_{t_0}^{t} dt' \int_{t_0}^{t'} ds' \langle G(t')H(s')\rangle \delta(t' - s') = 0 \qquad (4.2.52)$$

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

$$dW(t) = W(t + dt) - W(t) .$$ (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

$$\left. \begin{array}{l} I_1 = \tfrac{1}{2}f(t_1) \\ I_2 = \tfrac{1}{2}f(t_2). \end{array} \right\} \quad \text{(Stratonovich)}$$ (4.2.54)

This means that in both cases, the delta function occuring at the limit of an integral has $\tfrac{1}{2}$ 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

$$\frac{dx}{dt} = a(x, t) + b(x, t)\xi(t)$$ (4.3.1)

is a stochastic integral equation

$$x(t) - x(0) = \int_0^t dt' a[x(t'), t'] + \int_0^t dW(t') b[x(t'), t'] .$$ (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

$$x_{i+1} = x_i + a(x_i, t_i)\Delta t_i + b(x_i, t_i)\Delta W_i . \tag{4.3.6}$$
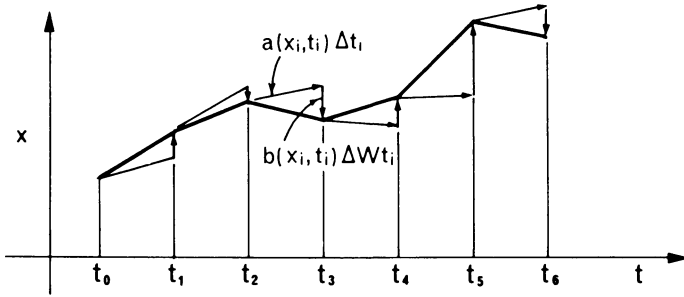
Fig. 4.2. Illustration of the Cauchy-Euler procedure for constructing an approximate solution of the stochastic differential equation $dx(t) = a[x(t), t]dt + b[x(t), t]dW(t)$

Here,

$$\left.\begin{array}{l} x_i = x(t_i) \\ \Delta t_i = t_{i+1} - t_i \\ \Delta W_i = W(t_{i+1}) - W(t_i). \end{array}\right\} \tag{4.3.7}$$

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

$$b(x_i, t_i)\Delta W_i . \tag{4.3.9}$$

The stochastic term contains an element ΔW_i, which is the increment of the Wiener process, but is statistically independent of x_i 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_i 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

$$
\boxed{
\begin{aligned}
df[x(t)] &= \{a[x(t), t]f'[x(t)] + \tfrac{1}{2}b[x(t), t]^2 f''[x(t)]\}\, dt \\
&\quad + b[x(t), t]f'[x(t)]dW(t) .
\end{aligned}}
\tag{4.3.14}
$$

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 $\tfrac{1}{2}$. By similar methods to those used in Sect.4.2.5, we can show that for an n dimensional Wiener process $W(t)$,

$$dW_i(t)dW_j(t) = \delta_{ij}dt \qquad (4.3.15a)$$

$$[dW_i(t)]^{N+2} = 0 \quad (N > 0) \qquad (4.3.15b)$$

$$dW_i(t)dt \quad = 0 \qquad (4.3.15c)$$

$$dt^{1+N} \quad = 0 \quad (N > 0) . \qquad (4.3.15d)$$

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 $x(t)$ satisfying the stochastic differential equation

$$dx = A(x, t)dt + \underset{\sim}{B}(x, t)dW(t) , \qquad (4.3.16)$$

we simply follow this procedure. The result is

$$df(x) = \{\sum_i A_i(x, t)\partial_i f(x) + \tfrac{1}{2} \sum_{i,j} [\underset{\sim}{B}(x, t)\underset{\sim}{B}^T(x, t)]_{ij}\partial_i\partial_j f(x)\} dt$$
$$+ \sum_{i,j} B_{ij}(x, t)\partial_i f(x)dW_j(t) . \qquad (4.3.17)$$

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

$$\langle df[x(t)]\rangle/dt = \left\langle \frac{df[x(t)]}{dt} \right\rangle = \frac{d}{dt} \langle f[x(t)]\rangle$$
$$= \langle a[x(t), t]\partial_x f + \tfrac{1}{2}b[x(t), t]^2\partial_x^2 f\rangle . \qquad (4.3.18)$$

However, $x(t)$ has a conditional probability density $p(x, t | x_0, t_0)$ and

$$\frac{d}{dt} \langle f[x(t)]\rangle = \int dx f(x)\partial_t p(x, t | x_0, t_0)$$
$$= \int dx[a(x, t)\partial_x f + \tfrac{1}{2}b(x, t)^2\partial_x^2 f]p(x, t | x_0, t_0) . \qquad (4.3.19)$$

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,

$$\partial_t p(x, t | x_0, t_0) = -\partial_x[a(x, t)p(x, t | x_0, t_0)] + \tfrac{1}{2}\partial_x^2[b(x, t)^2 p(x, t | x_0, t_0)] .$$

$$(4.3.20)$$

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

$$dx = A(x, t)dt + \underset{\sim}{B}(x, t)dW(t), \tag{4.3.21}$$

where $dW(t)$ 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 $p(x, t | x_0, t_0) \equiv p$ is

$$\partial_t p = -\sum_i \partial_i [A_i(x, t)p] + \tfrac{1}{2} \sum_{i,j} \partial_i \partial_j \{ [\underset{\sim}{B}(x, t) \underset{\sim}{B}^T(x, t)]_{ij} p \} . \tag{4.3.22}$$

Notice that the same Fokker-Planck equation arises from all matrices $\underset{\sim}{B}$ such that $\underset{\sim}{B}\underset{\sim}{B}^T$ is the same. This means that we can obtain the same Fokker-Planck equation by replacing $\underset{\sim}{B}$ by $\underset{\sim}{B}\underset{\sim}{S}$ where $\underset{\sim}{S}$ is orthogonal, i.e., $\underset{\sim}{S}\underset{\sim}{S}^T = 1$. Notice that $\underset{\sim}{S}$ may depend on $x(t)$. This can be seen more directly. Suppose $\underset{\sim}{S}(t)$ is an orthogonal matrix with an arbitrary *nonanticipating* dependence on t. Then define

$$dV(t) = \underset{\sim}{S}(t)dW(t). \tag{4.3.23}$$

Now the vector $dV(t)$ is a linear combination of Gaussian variables $dW(t)$ with coefficients $\underset{\sim}{S}(t)$ which are independent of $dW(t)$, since $\underset{\sim}{S}(t)$ is nonanticipating. For any fixed value of $\underset{\sim}{S}(t)$, the $dV(t)$ 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 $\underset{\sim}{S}(t)$ is orthogonal. Hence, all the moments are independent of $\underset{\sim}{S}(t)$ and are the same as those of $dW(t)$, so $dV(t)$ is itself Gaussian with the same correlation matrix as $dW(t)$. 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 $\underset{\sim}{S}(t)$ 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

$$dx = A(x, t)dt + \underline{B}(x, t)\underline{S}^T(t)\underline{S}(t)dW(t) \tag{4.3.27}$$
$$= A(x, t)dt + \underline{B}(x, t)\underline{S}^T(t)dV(t), \tag{4.3.28}$$

and since $V(t)$ is itself simply a Wiener process, this equation is equivalent to

$$dx = A(x, t)dt + \underline{B}(x, t)\underline{S}^T(t)dW(t) \tag{4.3.29}$$

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

$$S \int_{t_0}^{t} G[x(t'), t'] \, dW(t') = \text{ms-}\lim_{n \to \infty} \sum_{i=1}^{n} G\left\{ \frac{x(t_i) + x(t_{i-1})}{2}, t_{i-1} \right\} [W(t_i) - W(t_{i-1})] \tag{4.3.30}$$

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

$$dx(t_i) = a[x(t_{i-1}), t_{i-1}](t_i - t_{i-1}) + b[x(t_{i-1}), t_{i-1}][W(t_i) - W(t_{i-1})]. \qquad (4.3.34)$$

Then, applying Ito's formula, we can write

$$\beta\left[\frac{x(t_i) + x(t_{i-1})}{2}, t_{i-1}\right] = \beta[x(t_{i-1}) + \tfrac{1}{2}dx(t_{i-1}), t_{i-1}]$$

$$= \beta(t_{i-1}) + [a(t_{i-1})\,\partial_x\beta(t_{i-1}) + \tfrac{1}{4}b^2(t_{i-1})][\tfrac{1}{2}(t_i - t_{i-1})]$$

$$+ \tfrac{1}{2}b(t_{i-1})\,\partial_x^2\beta(t_{i-1})[W(t_i) - W(t_{i-1})] \qquad (4.3.35)$$

(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

$$S\!\int = \sum_i \beta(t_{i-1})\{W(t_i) - W(t_{i-1})\}$$

$$+ \tfrac{1}{2}\sum_i b(t_{i-1})\,\partial_x\beta(t_{i-1})(t_i - t_{i-1}).$$

Hence we derive

$$\boxed{S\int_{t_0}^{t}\beta[x(t'), t']dW(t') = \int_{t_0}^{t}\beta[x(t'), t']dW(t') + \tfrac{1}{2}\int_{t_0}^{t}b[x(t'), t']\partial_x\beta[x(t'), t']dt'.} \quad (4.3.37)$$

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

$$\alpha(x, t) = a(x, t) - \tfrac{1}{2}b(x, t)\partial_x b(x, t)$$

$$\beta(x, t) = b(x, t) \qquad (4.3.38)$$

$$\boxed{\begin{array}{ll} \text{We see that the Ito SDE} & dx = a\,dt + b\,dW(t), \\ \text{is the same as the Stratonovich SDE } dx = [a - \tfrac{1}{2}b\partial_x b]dt + b\,dW(t), \end{array}} \quad \begin{array}{l}(4.3.39a)\\(4.3.39b)\end{array}$$

or conversely,

$$\boxed{\begin{array}{ll} \text{the Stratonovich SDE} & dx = \alpha\,dt + \beta\,dW(t) \\ \text{is the same as the Ito SDE} & dx = [\alpha + \tfrac{1}{2}\beta\partial_x\beta]dt + \beta\,dW(t). \end{array}} \quad \begin{array}{l}(4.3.40a)\\(4.3.40b)\end{array}$$

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 $\bar{\alpha}(y) = \alpha[g(y)]$

$\qquad \bar{\beta}(y) = \beta[g(y)]$.

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

$$dx = A(x, t)dt + B(x, t)dW(t) , \tag{4.3.42}$$

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,

$$(S) \ dx = A^s(x, t)dt + B^s(x, t)dW(t) \tag{4.3.44}$$

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

$$\partial_t p = -\sum_i \partial_i \{A_i^s p\} + \tfrac{1}{2}\sum_{i,j,k} \partial_i \{B_{ik}^s \partial_j [B_{jk}^s p]\} \tag{4.3.45}$$

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

$$dx = a(\lambda, x, t)dt + b(\lambda, x, t)dW(t)$$

with initial condition

$$x(t_0) = c(\lambda)$$

$$(4.3.46)$$

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) $\underset{\lambda \to \lambda_0}{\lim}\{\sup_{\substack{t \in [t_0, T] \\ |x| \leqslant N}}[|a(\lambda, x, t) - a(\lambda_0, x, t)| + |b(\lambda, x, t) - b(\lambda_0, x, t)|]\} = 0$;

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,

$$\underset{\lambda \to \lambda_0}{\text{st-lim}}\{\sup_{t \in [t_0, T]} |x(\lambda, t) - x(\lambda_0, t)|\} = 0 .$$

$$(4.3.47)$$

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,

$$G(t') = b(t') \left. \right\} \quad t' < t$$
$$= 0 \quad \left. \right\} \quad t' \geqslant t$$

$$H(t') = b(t') \left. \right\} \quad t' < s$$
$$= 0 \quad \left. \right\} \quad t' \geqslant s.$$

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 Stratono-vich 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

$$\frac{dz}{dt} = i[\omega + \sqrt{2\gamma} \, \xi(t)]z \tag{4.4.14}$$

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

$$(S) \quad dz = i[\omega dt + \sqrt{2\gamma} \, dW(t)]z \tag{4.4.15}$$

which is equivalent to the Ito equation (from Sect. 4.3.6)

$$dz = [(i\omega - \gamma)dt + i\sqrt{2\gamma} \, dW(t)]z \, . \tag{4.4.16}$$

Taking the mean value, we see immediately that

$$\frac{d\langle z \rangle}{dt} = (i\omega - \gamma)\langle z \rangle \tag{4.4.17}$$

with the *damped* oscillatory solution

$$\langle z(t) \rangle = \exp\left[(i\omega - \gamma)t\right]\langle z(0) \rangle \, . \tag{4.4.18}$$

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

$$\langle z(t)z(s) \rangle = \langle z(0)^2 \rangle \exp\left[(i\omega - \gamma)(t + s) - 2\gamma \min(t, s)\right] \, . \tag{4.4.19}$$

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

$$\begin{aligned}
\langle z(t)z^*(s) \rangle &= \langle |z(0)|^2 \rangle \langle \exp\{i\omega(t - s) + i\sqrt{2\gamma}[W(t) - W(s)]\} \rangle \\
&= \langle |z(0)|^2 \rangle \exp\{i\omega(t - s) - \gamma[t + s - 2\min(t, s)]\} \\
&= \langle |z(0)|^2 \rangle \exp\left[i\omega(t - s) - \gamma|t - s|\right] \, .
\end{aligned} \tag{4.4.21}$$

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

$$dy = (dx)d(e^{kt}) + (dx)e^{kt} + xd(e^{kt})$$

$$= [-kx \, dt + \sqrt{D} \, dW(t)]k \, e^{kt}dt$$

$$+ [-kx \, dt + \sqrt{D} \, dW(t)]e^{kt} + kx \, e^{kt}dt. \tag{4.4.24}$$

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

$$\langle x(t) \rangle = \langle x(0) \rangle e^{-kt} \tag{4.4.27}$$

$$\text{var}\{x(t)\} = \langle \{[x(0) - \langle x(0) \rangle]e^{-kt} + \sqrt{D} \int_0^t e^{-k(t-t')}dW(t')\}^2 \rangle. \tag{4.4.28}$$

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

$$\text{var}\{x(t)\} = \text{var}\{x(0)\} e^{-2kt} + D \int_0^t e^{-2k(t-t')}dt'$$

$$= \{[\mathrm{var}\,\{x(0)\} - D/2k]\,e^{-2kt} + D/2k\,. \tag{4.4.29}$$

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,

$$\langle x(t), x(s)\rangle = \mathrm{var}\,\{x(0)\}\,e^{-k\,(t+s)} + D\langle \int\limits_0^t e^{-k\,(t-t')} dW(t') \int\limits_0^s e^{-k\,(s-s')} dW(s')\rangle$$

$$= \mathrm{var}\,\{x(0)\}\,e^{-k\,(t+s)} + D\int\limits_0^{\min(t,\,s)} e^{-k\,(t+s-2t')} dt'$$

$$= \left[\mathrm{var}\,\{x(0)\} - \frac{D}{2k}\right] e^{-k\,(t+s)} + \frac{D}{2k}\,e^{-k|t-s|}\,. \tag{4.4.30}$$

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

$$x(t) = \sqrt{D}\,\int\limits_{-\infty}^t e^{-k\,(t-t')} dW(t')\,. \tag{4.4.31}$$

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.,

$$dE_1(t) = -\gamma E_1(t)\,dt + \varepsilon\,dW_1(t)$$

$$dE_2(t) = -\gamma E_2(t)\,dt + \varepsilon\,dW_2(t)\,. \tag{4.4.32}$$

It is of interest to convert to polar coordinates. We set

$$E_1(t) = a(t)\cos\phi(t)$$

$$E_2(t) = a(t)\sin\phi(t) \tag{4.4.33}$$

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

$$
\begin{aligned}
d(\mu + i\phi) &= \frac{d(E_1 + iE_2)}{E_1 + iE_2} - \frac{1}{2}\frac{[d(E_1 + iE_2)]^2}{(E_1 + iE_2)^2} \\
&= -\frac{\gamma(E_1 + iE_2)}{E_1 + iE_2}\, dt + \frac{\varepsilon[dW_1(t) + idW_2(t)]}{(E_1 + iE_2)} \\
&\quad - \frac{1}{2}\frac{\varepsilon^2[dW_1(t) + i\, dW_2(t)]^2}{(E_1 + iE_2)^2}
\end{aligned}
\tag{4.4.36}
$$

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 $dW_\phi(t)$ 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 $[\tfrac{1}{2}\varepsilon^2/a(t)]dt$ 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

$$dx(t) = -Ax(t)dt + B\,dW(t), \tag{4.4.42}$$

(A and B are constant matrices) for which the solution is easily obtained (as in Sect. 4.4.4):

$$x(t) = \exp(-At)x(0) + \int_0^t \exp[-A(t-t')]B\,dW(t'). \tag{4.4.43}$$

The mean is

$$\langle x(t) \rangle = \exp(-At)\langle x(0) \rangle. \tag{4.4.44}$$

The correlation function follows similarly

$$
\begin{aligned}
\langle x(t), x^{\mathrm{T}}(s) \rangle &\equiv \langle [x(t) - \langle x(t) \rangle][x(s) - \langle x(s) \rangle]^{\mathrm{T}} \rangle \\
&= \exp(-At)\langle x(0), x^{\mathrm{T}}(0) \rangle \exp(-As) \\
&\quad + \int_0^{\min(t,s)} \exp[-A(t-t')]BB^{\mathrm{T}}\exp[-A^{\mathrm{T}}(s-t')]dt'.
\end{aligned} \tag{4.4.45}
$$

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 $\langle x(0), x^{\mathrm{T}}(0) \rangle = 0$, 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

$$SS^+ = 1$$
$$SAS^+ = SA^{\mathrm{T}}S^+ = \mathrm{diag}(\lambda_1, \lambda_2, ...\lambda_n). \tag{4.4.46}$$

For simplicity, assume $t \geqslant s$. Then

$$\langle x(t), x^{\mathrm{T}}(s) \rangle = S^+ G(t, s)S,$$

where

$$[G(t, s)]_{ij} = \frac{(BB^T)_{ij}}{\lambda_i + \lambda_j} [\exp(-\lambda_i|t - s|) - \exp(-\lambda_i t - \lambda_j s)] . \tag{4.4.47}$$

b) Variance in Stationary Solution

If A has only eigenvalues with positive real part, a stationary solution exists of the form

$$x_s(t) = \int_{-\infty}^{t} \exp[-A(t - t')]B \, dW(t) . \tag{4.4.48}$$

We have of course

$$\langle x_s(t) \rangle = 0$$

and

$$\langle x_s(t), x_s^T(s) \rangle = \int_{-\infty}^{\min(t,s)} \exp[-A(t - t')]BB^T \exp[-A^T(s - t')]dt' . \tag{4.4.49}$$

Let us define the stationary covariance matrix σ by

$$\sigma = \langle x_s(t), x_s^T(t) \rangle . \tag{4.4.50}$$

Then the evaluation of this quantity can be achieved algebraically for

$$A\sigma + \sigma A^T = \int_{-\infty}^{t} A \exp[-A(t - t')]BB^T \exp[-A^T(t - t')]dt'$$

$$+ \int_{-\infty}^{t} \exp[-A(t - t')]BB^T \exp[-A^T(t - t')]A^T dt'$$

$$= \int_{-\infty}^{t} \frac{d}{dt'} \{\exp[-A(t - t')]BB^T \exp[-A(t - t')]\} \, dt' .$$

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

$$\boxed{A\sigma + \sigma A^T = BB^T} \tag{4.4.51}$$

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

$$A^2 - (\text{Tr } A)A + (\text{Det } A) = 0 \tag{4.4.52}$$

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

$$\sigma = \alpha BB^T + \beta(ABB^T + BB^T A^T) + \gamma ABB^T A^T .$$

Using (4.4.52), we find (4.4.51) is satisfied if

$$\alpha + (\text{Tr } A)\beta - (\text{Det } A)\gamma = 0$$

$$2\beta(\text{Det } A) + 1 = 0$$

$$\beta + (\text{Tr } A)\gamma = 0 .$$

From which,

$$\sigma = \frac{(\text{Det } A)BB^{\text{T}} + [A - (\text{Tr } A)1]BB^{\text{T}}[A - (\text{Tr } A)1]^{\text{T}}}{2(\text{Tr } A)(\text{Det } A)} . \qquad (4.4.53)$$

d) Time Correlation Matrix in the Stationary Sate
From the solution of ‏(4.4.49), we see that if $t > s$,

$$\langle x_s(t), x_s^{\text{T}}(s) \rangle = \exp[-A(t-s)] \int_{-\infty}^{s} \exp[-A(s-t')]BB^{\text{T}}\exp[-A^{\text{T}}(s-t')]dt'$$

$$= \exp[-A(t-s)]\sigma \qquad t > s \qquad (4.4.54a)$$

and similarly,

$$= \sigma \exp[-A^{\text{T}}(s-t)] \qquad t < s \qquad (4.4.54b)$$

which gives the dependence on $s - t$ as expected of a stationary solution. Defining then

$$G_s(t-s) = \langle x_s(t), x_s^{\text{T}}(s) \rangle , \qquad (4.4.55)$$

we see (remembering $\sigma = \sigma^{\text{T}}$) that

$$G_s(t-s) = [G_s(s-t)]^{\text{T}} . \qquad (4.4.56)$$

e) Spectrum Matrix in Stationary State
The spectrum matrix turns out to be rather simple. We define similarly to Sect. 1.4.2:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} G_s(\tau)d\tau \qquad (4.4.57)$$

$$= \frac{1}{2\pi} \left\{ \int_{0}^{\infty} \exp[-(i\omega + A)\tau]\sigma \, d\tau + \int_{-\infty}^{0} \sigma \exp[(-i\omega + A^{\text{T}})\tau]d\tau \right\}$$

$$= \frac{1}{2\pi}[(A + i\omega)^{-1}\sigma + \sigma(A^{\text{T}} - i\omega)^{-1}] .$$

Hence, $(A + i\omega)S(\omega)(A^{\text{T}} - i\omega) = \frac{1}{2\pi}(\sigma A^{\text{T}} + A\sigma)$ and using (4.4.51), we get

$$S(\omega) = \frac{1}{2\pi}(A + i\omega)^{-1}BB^{\text{T}}(A^{\text{T}} - i\omega)^{-1} . \qquad (4.4.58)$$

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

$$\frac{d}{d\tau}[G_s(\tau)]d\tau = \frac{d}{d\tau}\langle x_s(\tau), x_s^T(0)\rangle d\tau$$

$$= \langle[-Ax_s(\tau)d\tau + B\,dW(\tau)], x_s^T(0)\rangle \qquad (4.4.59)$$

and since $\tau > 0$, $dW(\tau)$ is uncorrelated with $x_s^T(0)$, so

$$\boxed{\frac{d}{d\tau}[G_s(\tau)] = -A\,G_s(\tau)\,.} \qquad (4.4.59)$$

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

$$dx = [b(t)dt + g(t)dW(t)]x \qquad (4.4.60)$$

and using the usual Ito rules, write

$$y = \log x \qquad (4.4.61)$$

so that

$$dy = dx/x - \tfrac{1}{2}(dx)^2/x^2$$
$$= [b(t)dt + g(t)dW(t)] - \tfrac{1}{2}g(t)^2 dt, \qquad (4.4.62)$$

and integrating and inverting (4.4.61), we get

$$x(t) = x(0)\exp\left\{\int_0^t [b(t') - \tfrac{1}{2}g(t')^2]dt' + \int_0^t g(t')dW(t')\right\} \qquad (4.4.63)$$

$$\equiv x(0)\phi(t) \qquad (4.4.64)$$

which serves to define $\phi(t)$.

We note that [using (4.4.8)]

$$\langle[x(t)]^n\rangle = \langle[x(0)]^n\rangle\left\langle\exp\left\{n\int_0^t [b(t') - \tfrac{1}{2}g(t')^2]dt' + n\int_0^t g(t')dW(t')\right\}\right\rangle$$

$$= \langle[x(0)]^n\rangle\exp\left\{n\int b(t')dt' + \tfrac{1}{2}n(n-1)\int g(t')^2 dt'\right\}. \qquad (4.4.65)$$

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

$$d[x(t)^n] = nx(t)^{n-1}dx(t) + \frac{n(n-1)}{2} x(t)^{n-2}[dx(t)]^2$$

$$= nx(t)^{n-1}dx(t) + \frac{n(n-1)}{2} x(t)^{n-2}[f(t) + g(t)x(t)]^2 dt . \tag{4.4.70}$$

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

$$dx(t) = [B(t)dt + \sum_i G_i(t)dW_i(t)]x(t) , \tag{4.4.72}$$

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.

$$G_i(t)G_j(t') = G_j(t')G_i(t)$$
$$B(t)G_i(t') = G_i(t')B(t) \tag{4.4.73}$$
$$B(t)B(t') = B(t')B(t)$$

In this case, the solution is completely analogous to the one variable case and we have

$$x(t) = \Phi(t)x(0)$$

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

$$dx(t) = [A(t) + B(t)x]dt + \sum_i [F_i(t) + G_i(t)x]dW(t) \tag{4.4.75}$$

and write

$$y(t) = \psi(t)^{-1}x(t) , \tag{4.4.76}$$

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

$$dy(t) = \psi(t)^{-1}\{[A(t) - \sum_i G_i(t)F_i(t)]dt + \sum_i F_i(t)dW_i(t)\}.$$

Hence,

$$x(t) = \psi(t)\{x(0) + \int_0^t \psi(t')^{-1}\{[A(t') - \sum_i G_i(t')F_i(t')]dt' + \sum_i F_i(t')dW_i(t')\}\}.$$

$$\text{(4.4.78)}$$

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,

$$dx(t) = -A(t)x(t)dt + B(t)dW(t).$$

$$\text{(4.4.79)}$$

This is clearly of the same form as (4.4.75) with the replacements

$$A(t) \rightarrow 0$$
$$B(t) \rightarrow -A(t)$$
$$\sum_i F_i(t)dW_i(t) \rightarrow B(t)dW(t) \qquad\qquad\qquad \text{(4.4.80)}$$
$$G_i(t) \rightarrow 0.$$

The corresponding homogeneous equation is simply the deterministic equation

$$dx(t) = -A(t)x(t)\,dt \qquad\qquad\qquad\qquad\qquad \text{(4.4.81)}$$

which is soluble provided $A(t)A(t') = A(t')A(t)$ and has the solution

$$x(t) = \psi(t)x(0)$$

with

$$\psi(t) = \exp[-\int_0^t A(t')dt']. \qquad\qquad\qquad\qquad \text{(4.4.82)}$$

Thus, applying (4.4.78),

$$x(t) = \exp[-\int_0^t A(t')dt']x(0) + \int_0^t \{\exp[-\int_{t'}^t A(s)ds]\}B(t')dW(t'). \qquad \text{(4.4.83)}$$

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

$$\partial_t p(\mathbf{z}, t) = -\sum_i \frac{\partial}{\partial z_i} A_i(\mathbf{z}, t) p(\mathbf{z}, t) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial z_i \partial z_j} B_{ij}(\mathbf{z}, t) p(\mathbf{z}, t) \qquad (5.2.6)$$

We note that this can also be written

$$\frac{\partial p(\mathbf{z}, t)}{\partial t} + \sum_i \frac{\partial}{\partial z_i} J_i(\mathbf{z}, t) = 0 \qquad (5.2.7)$$

where we define the *probability current*

$$J_i(\mathbf{z}, t) = A_i(\mathbf{z}, t) p(\mathbf{z}, t) - \frac{1}{2} \sum_j \frac{\partial}{\partial z_j} B_{ij}(\mathbf{z}, t) p(\mathbf{z}, t) \qquad (5.2.8)$$

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 S and define

$$P(R, t) = \int_R d\mathbf{z}\, p(\mathbf{z}, t)$$

then (5.2.7) is equivalent to

$$\frac{\partial P(R, t)}{\partial t} = -\int_S dS\, \mathbf{n} \cdot \mathbf{J}(\mathbf{z}, t) \qquad (5.2.9)$$

where $\mathbf{n}$ is the outward pointing normal to S. Thus (5.2.9) indicates that the total loss of probability is given by the surface integral of $\mathbf{J}$ over the boundary of R. We can show as well that the current $\mathbf{J}$ does have the somewhat stronger property, that a surface integral over any surface S gives the net flow of probability across that surface.
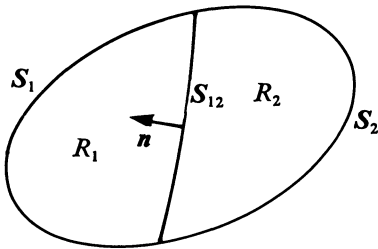
Fig. 5.1. Regions used to demonstrate that the probability current is the flow of probability

For consider two adjacent regions R_1 and R_2, separated by a surface S_{12}. Let S_1 and S_2 be the surfaces which, together with 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.

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{R_1} dx \int_{R_2} dy\, [p(x, t + \Delta t; y, t) - p(y, t + \Delta t; x, t)]. \qquad (5.2.10)$$

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)

$$= -\int_{R_1} dx \sum_i \frac{\partial}{\partial x_i}\, J_i(x, t; R_2, t) + \int_{R_2} dy \sum_i \frac{\partial}{\partial y_i}\, J_i(y, t; R_1, t) \qquad (5.2.11)$$

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 \qquad (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

$$\boldsymbol{n} \cdot \boldsymbol{J}(z, t) = 0 \quad \text{for } z \in S, \quad \boldsymbol{n} = \text{normal to } S \tag{5.2.13}$$

where $\boldsymbol{J}(z, t)$ is given by (5.2.8).

Since the particle cannot cross S, it must be reflected there, and hence the name *reflecting barrier* 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.

$$p(z, t) = 0 \quad \text{for } z \in S \tag{5.2.14}$$

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,

$$\boldsymbol{n} \cdot \boldsymbol{J}(z) |_{S_+} = \boldsymbol{n} \cdot \boldsymbol{J}(z) |_{S_-} \tag{5.2.15}$$

$$p(z) |_{S_+} = p(z) |_{S_-} \tag{5.2.16}$$

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 $p(z)$ 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

$$A(b, t) = A(a, t) \tag{5.2.19}$$

$$B(b, t) = B(a, t)$$

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

$$\mathscr{S}(x_1, x_2) \tag{5.2.29}$$

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 $f(x) \in \mathscr{S}(a, x_0)$, and $g(x) \in \mathscr{S}(a, x_0)$

II: Exit: if $g(x) \notin \mathscr{S}(a, x_0)$, and $h_1(x) \in \mathscr{S}(a, x_0)$

III: Entrance: if $g(x) \in \mathscr{S}(a, x_0)$, and $h_2(x) \in \mathscr{S}(a, x_0)$

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

$$p_s(x) = p_s(a)\psi(x)/\psi(a)$$
$$= p_s(a)\exp\left[-\frac{k}{D}(x^2 - a^2)\right]$$

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

$$B(0, t) \equiv (ax + x^2)_{x=0} = 0$$
$$A(0, t) \equiv (ax - x^2)_{x=0} = 0 \tag{5.2.60}$$
$$\partial_x B(0, t) \equiv (a + 2x)_{x=0} > 0$$

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 $x = 0$.
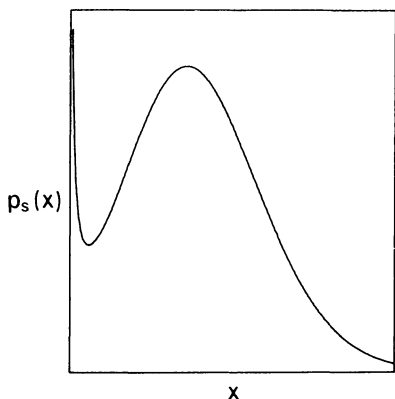

Fig. 5.2. Non-normalisable "stationary" $p_s(x)$ for the reaction $X + A \rightleftharpoons 2X$

5.2.4 Boundary Conditions for the Backward Fokker-Planck Equation

We suppose that $p(x, t | x', t')$ obeys the forward Fokker-Planck equation for a set of x, t and x', t', and that the process is confined to a region R with boundary S. Then, if s is a time between t and t',

$$0 = \frac{\partial}{\partial s} p(x, t | x', t') = \frac{\partial}{\partial s} \int dy\, p(x, t | y, s) p(y, s | x', t') , \qquad (5.2.61)$$

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

$$p(y, s) = p(y, s | x', t') \qquad (5.2.62)$$
$$\bar{p}(y, s) = p(x, t | y, s) .$$

Then,

$$0 = \int_R dy \left[-\sum_i \frac{\partial}{\partial y_i} (A_i p) + \sum_{i,j} \frac{\partial^2}{\partial y_i \partial y_j} (B_{ij} p) \right] \bar{p} \qquad (5.2.63)$$
$$+ \int_R dy \left[-\sum_i A_i \frac{\partial \bar{p}}{\partial y_i} - \sum_{i,j} B_{ij} \frac{\partial^2 \bar{p}}{\partial y_i \partial y_j} \right] p$$

and after some manipulation

$$= \int_R dy \sum_i \frac{\partial}{\partial y_i} \left\{ -A_i p \bar{p} + \frac{1}{2} \sum_j \left[\bar{p} \frac{\partial}{\partial y_j} (B_{ij} p) - p B_{ij} \frac{\partial \bar{p}}{\partial y_j} \right] \right\} \qquad (5.2.64)$$

$$= \int_S \sum_i dS_i \left\{ \bar{p} \left[-A_i p + \frac{1}{2} \sum_j \frac{\partial}{\partial y_i} (B_{ij} p) \right] \right\}$$

$$- \frac{1}{2} \int_S \sum_i dS_i p \left(\sum_j B_{ij} \frac{\partial \bar{p}}{\partial y_j} \right). \tag{5.2.65}$$

We now treat the various cases individually.

a) Absorbing Boundaries
This requires $p = 0$ on the boundary. That it also requires $\bar{p}(y, t) = 0$ 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

$$p(x, t \,|\, y, s) = 0, \qquad \text{for } y \in \text{boundary} \tag{5.2.67}$$

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

$$\sum_{i,j} n_i B_{ij}(y) \frac{\partial}{\partial y_j} [p(x, t \,|\, y, s)] = 0 . \tag{5.2.68}$$

In one dimension this reduces to

$$\frac{\partial}{\partial y} p(x, t \,|\, y, s) = 0 \tag{5.2.69}$$

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

$$\partial_t p(x, t) = -\partial_x[A(x)p(x, t)] + \tfrac{1}{2} \partial_x^2 [B(x)p(x, t)] . \qquad (5.2.70)$$

We define a function $q(x, t)$ by

$$p(x, t) = p_s(x)q(x, t) \qquad (5.2.71)$$

and, by direct substitution, find that $q(x, t)$ satisfies the *backward equation*

$$\partial_t q(x, t) = A(x)\partial_x q(x, t) + \tfrac{1}{2} B(x)\partial_x^2 q(x, t) . \qquad (5.2.72)$$

We now wish to consider solutions of the form

$$p(x, t) = P_\lambda(x)e^{-\lambda t} \qquad (5.2.73)$$

$$q(x, t) = Q_\lambda(x)e^{-\lambda t} \qquad (5.2.74)$$

which obey the eigenfunction equations

$$-\partial_x[A(x)P_\lambda(x)] + \tfrac{1}{2} \partial_x^2[B(x)P_\lambda(x)] = -\lambda P_\lambda(x) \qquad (5.2.75)$$

$$A(x)\partial_x Q_{\lambda'}(x) + \tfrac{1}{2} B(x)\partial_x^2 Q_{\lambda'}(x) = -\lambda' Q_{\lambda'}(x) . \qquad (5.2.76)$$

Then we can straightforwardly show by partial integration that

$$(\lambda' - \lambda) \int_a^b dx P_\lambda(x)Q_{\lambda'}(x) = [Q_{\lambda'}(x)\{-A(x)P_\lambda(x) + \tfrac{1}{2} \partial_x[B(x)P_\lambda(x)]\}$$
$$- \tfrac{1}{2} B(x)P_\lambda(x)\partial_x Q_{\lambda'}(x)]_a^b, \qquad (5.2.77)$$

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

$$\tfrac{1}{2} B(x)\partial_x Q_{\lambda'}(x) = -A(x)P_{\lambda'}(x) + \tfrac{1}{2} \partial_x[B(x)P_{\lambda'}(x)] \qquad (5.2.78)$$

so that term vanishes also. Hence, the $Q_\lambda(x)$ and $P_\lambda(x)$ form a bi-orthogonal system

$$\int_a^b dx \, P_\lambda(x)Q_{\lambda'}(x) = \delta_{\lambda\lambda'} \qquad (5.2.79)$$

or, there are two alternative orthogonality systems,

$$\int_a^b dx \, p_s(x)Q_\lambda(x)Q_{\lambda'}(x) = \delta_{\lambda\lambda'} \qquad (5.2.80)$$

$$\int_a^b dx[p_s(x)]^{-1}P_\lambda(x)P_{\lambda'}(x) = \delta_{\lambda\lambda'} . \qquad (5.2.81)$$

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

$$\langle x(t) x(0) \rangle = \int dx \int dx_0 \, x x_0 p(x, t | x_0, 0) p_s(x) \tag{5.2.89}$$

$$= \sum_\lambda [\int dx \, x P_\lambda(x)]^2 e^{-\lambda t} , \tag{5.2.90}$$

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 $p(x, t | x_0, 0) \to 0$ 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 $p(x, t | x_0, 0)$]

$$p(x, t | x_0, 0) = 2 \sum_{n=1}^{\infty} \exp(-\lambda_n t) \sin(n\pi x_0) \sin(n\pi x) . \tag{5.2.100}$$

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

$$p(x, t \mid x_0, 0) = 1 + 2 \sum_{n=1}^{\infty} \cos(n\pi x_0) \cos(n\pi x) \exp(-\lambda_n t) . \tag{5.2.107}$$

As $t \to \infty$, the process becomes stationary, with stationary distribution

$$p_s(x) = \lim_{t \to \infty} p(x, t \mid x_0, 0) = 1 . \tag{5.2.108}$$

We can compute the stationary autocorrelation function by

$$\langle x(t)x(0) \rangle_s = \int_0^1 \int_0^1 dx \, dx_0 \, x x_0 p(x, t \mid x_0, 0) p_s(x) \tag{5.2.109}$$

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

$$\langle x(t)x(0) \rangle_s \to \frac{1}{4} = [\langle x \rangle_s]^2 \tag{5.2.111}$$

and as $t \rightarrow 0$,

$$\langle x(t)x(0) \rangle_s \rightarrow \frac{1}{4} + \frac{8}{\pi^4} \sum_{n=0}^{\infty} (2n+1)^{-4} = \frac{1}{3} = \langle x_s^2 \rangle \qquad (5.2.113)$$

when one takes account of the identity (from the theory of the Riemann zeta-function)

$$\sum_{n=0}^{\infty} (2n+1)^{-4} = \frac{\pi^4}{96}. \qquad (5.2.114)$$

c) Ornstein-Uhlenbeck Process

As in Sect.3.8.4, the Fokker-Planck equation is

$$\partial_t p(x, t) = \partial_x[kxp(x, t)] + \frac{1}{2} D\partial_x^2 p(x, t) . \qquad (5.2.115)$$

The eigenfunction equation for Q_λ is

$$d_x^2 Q_\lambda - \frac{2kx}{d} d_x Q_\lambda + \frac{2\lambda}{D} Q_\lambda = 0 \qquad (5.2.116)$$

and this becomes the differential equation for *Hermite polynomials* [5.6] on making the replacement

$$y = x\sqrt{k/D}$$
$$d_y^2 Q_\lambda - 2yd_y Q_\lambda + (2\lambda/k)Q_\lambda = 0 . \qquad (5.2.117)$$

We can write

$$Q_\lambda = (2^n n!)^{-1/2} H_n(x\sqrt{k/D}) , \qquad (5.2.118)$$

where

$$\lambda = nk , \qquad (5.2.119)$$

and these solutions are normalised as in (5.2.79–81).

The stationary solution is, as previously found,

$$p_s(x) = (k/\pi D)^{1/2} \exp(-kx^2/D) \qquad (5.2.120)$$

and a general solution can be written as

$$p(x, t) = \sum_n \sqrt{[k/(2^n n! \pi D)]} \exp(-kx^2/D) H_n(x\sqrt{k/D}) e^{-nkt} A_n \qquad (5.2.121)$$

with

$$A_n = \int_{-\infty}^{\infty} dx \, p(x, 0) H_n(x\sqrt{k/D})(2^n n!)^{-1/2} . \qquad (5.2.122)$$

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

$$H_1(y) = 2y \tag{5.2.124}$$

so that the orthogonality property means that only the eigenfunction corresponding to $n = 1$ has a nonzero coefficient. We must compute

$$\int x P_{\lambda_1}(x)dx = \int_{-\infty}^{\infty} \sqrt{k/(2\pi D)} \exp(-kx^2/D)(2x\sqrt{k/D})x \tag{5.2.125}$$
$$= \sqrt{D/2k}$$

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

$$zd_z^2 Q_\lambda + (1 - z)d_z Q_\lambda + (\lambda/2\gamma)Q_\lambda = 0 . \qquad (5.2.132)$$

This is the differential equation for the *Laguerre polynomials* [5.6] provided

$$\lambda = 2n\gamma . \qquad (5.2.133)$$

We can write

$$Q_\lambda(x) = L_n(\gamma x^2/2\mu) \qquad (5.2.134)$$

which is normalised. Hence, the conditional probability is

$$p(x, t \,|\, x_0, 0) = \sum_{n=0}^{\infty} \frac{\gamma x}{\mu} \exp\left(-\frac{\gamma x^2}{2\mu}\right) L_n\left(\frac{\gamma x_0^2}{2\mu}\right) L_n\left(\frac{\gamma x^2}{2\mu}\right) e^{-2n\gamma t} . \qquad (5.2.135)$$

We can compute the autocorrelation function by the method of (5.2.90):

$$\langle x(t)x(0) \rangle = \sum_{n=0}^{\infty} \left[\int_0^{\infty} x \, dx \frac{\gamma x}{\mu} \exp\left(\frac{-\gamma x^2}{2\mu}\right) L_n\left(\frac{\gamma x^2}{2\mu}\right)\right]^2 \exp(-2n\gamma t) \qquad (5.2.136)$$

and using

$$\int_0^{\infty} dz \, z^\alpha e^{-z} L_n(z) = (-1)^n \Gamma(\alpha + 1)\binom{\alpha}{n} , \qquad (5.2.137)$$

we find for the autocorrelation function

$$\langle x(t)x(0) \rangle = \frac{2\mu}{\gamma} \sum_{n=0}^{\infty} \frac{\pi}{4} \binom{\frac{1}{2}}{n}^2 \exp(-2n\gamma t). \qquad (5.2.138)$$

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:

$$a \leqslant x \leqslant b \qquad (5.2.139)$$

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

$$\int_a^b dx'p(x', t \,|\, x, 0) \equiv G(x, t) \,.$$ (5.2.140)

Let the time that the particle leaves (a, b) be T. Then we can rewrite (5.2.140) as

$$\text{Prob}(T \geqslant t) = \int_a^b dx'p(x', t \,|\, x, 0)$$ (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

$$p(x', t \,|\, x, 0) = p(x', 0 \,|\, x, -t)$$ (5.2.142)

and the backward Fokker-Planck equation can be written

$$\partial_t p(x', t \,|\, x, 0) = A(x)\partial_x p(x', t \,|\, x, 0) + \tfrac{1}{2} B(x)\partial_x^2 p(x', t \,|\, x, 0)$$ (5.2.143)

and hence, $G(x, t)$ obeys the equation

$$\partial_t G(x, t) = A(x)\partial_x G(x, t) + \tfrac{1}{2} B(x)\partial_x^2 G(x, t) \,.$$ (5.2.144)

The boundary conditions are clearly that

$$p(x', 0 \,|\, x, 0) = \delta(x - x')$$

and hence,

$$G(x, 0) = 1 \qquad a \leqslant x \leqslant b$$
$$\qquad\quad = 0 \qquad \text{elsewhere}$$ (5.2.145)

and if $x = a$ or b, the particle is absorbed immediately, so

$$\text{Prob}(T \geqslant t) = 0 \qquad \text{when } x = a \text{ or } b \,, \qquad \text{i.e.,}$$
$$G(a, t) = G(b, t) = 0.$$ (5.2.146)

Since $G(x, t)$ is the probability that $T \geqslant t$, the mean of any function of T is

$$\langle f(T) \rangle = - \int_0^\infty f(t)dG(x, t) \,.$$ (5.2.147)

Thus, the *mean first passage time*

$$T(x) = \langle T \rangle$$ (5.2.148)

is given by

$$T(x) = - \int_0^\infty t\, \partial_t G(x, t)dt$$ (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

$$\psi(x) = \exp\left\{\int_a^x dx'[2A(x')/B(x')]\right\} . \tag{5.2.157}$$

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

$$T(x) = 2 \int_x^b \frac{dy}{\psi(y)} \int_a^y \frac{\psi(z)}{B(z)} dz \qquad \begin{array}{l} a \quad \text{reflecting} \\ b \quad \text{absorbing} \\ a < b \end{array} \tag{5.2.160}$$

Similarly, one finds

$$T(x) = 2 \int_a^x \frac{dy}{\psi(y)} \int_y^b \frac{\psi(z)}{B(z)} dz \qquad \begin{array}{l} b \quad \text{reflecting} \\ a \quad \text{absorbing} \\ a < b \end{array} \tag{5.2.161}$$

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

$$p_s(x) = \mathcal{N} \exp[-U(x)/D] \tag{5.2.163}$$

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

$$
\begin{aligned}
b &\rightarrow x_0 \\
a &\rightarrow -\infty \\
x &\rightarrow a
\end{aligned}
\tag{5.2.164}
$$

so that

$$T(a \rightarrow x_0) = \frac{1}{D} \int_a^{x_0} dy \, \exp[U(y)/D] \int_{-\infty}^y \exp[-U(z)/D] dz . \tag{5.2.165}$$

(a)

U(x)

x

(b)

$p_s(x)$

x

(c)

T(a→x)

x
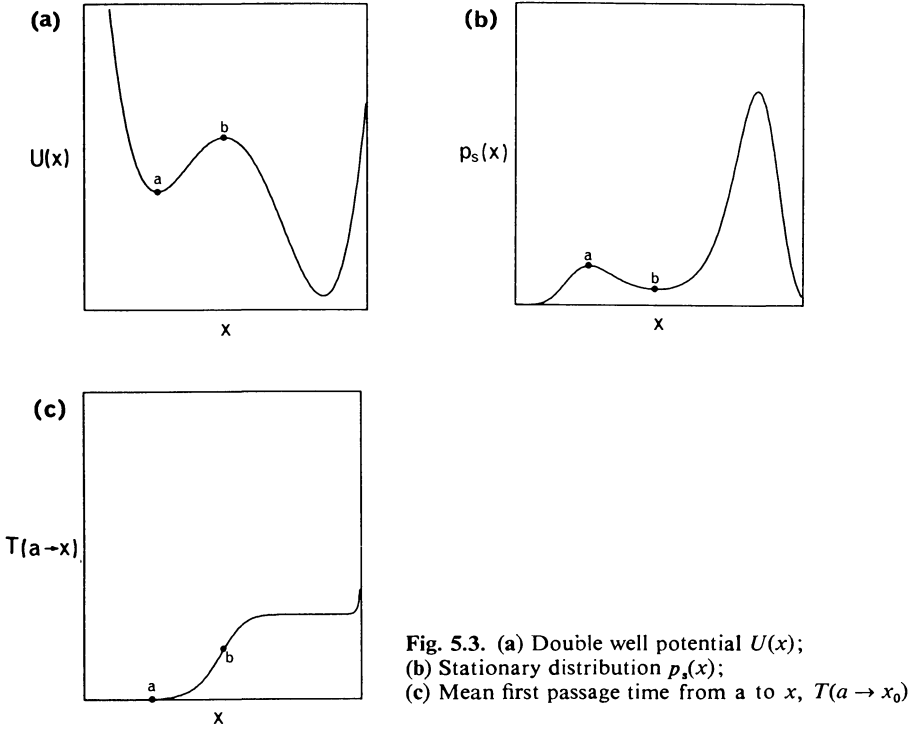
Fig. 5.3. (a) Double well potential $U(x)$;
(b) Stationary distribution $p_s(x)$;
(c) Mean first passage time from a to x, $T(a \rightarrow x_0)$

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

$$T(a \rightarrow x_0) \simeq \left\{\frac{1}{D} \int_{-\infty}^{b} dy \exp[-U(z)/D]\right\} \int_{a}^{x_0} dy \exp[U(y)/D] . \qquad (5.2.166)$$

Notice that by the definition of $p_s(x)$ in (5.2.163), we can say that

$$\int_{-\infty}^{b} dy \exp[-U(z)/D] = n_a/\mathcal{N} \qquad (5.2.167)$$

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 $T(a \rightarrow x_0)$ 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

$$U(x) \simeq U(a) + \frac{1}{2_s}\left(\frac{x-a}{\alpha}\right)^2 . \tag{5.2.169}$$

The constant factor in (5.2.166) is evaluated as

$$\int_{-\infty}^{b} dz \exp\left[-U(z)/D\right] \sim \int_{-\infty}^{\infty} dz \exp\left[-\frac{U(a)}{D} - \frac{(z-a)^2}{2D\alpha^2}\right] \tag{5.2.170}$$

$$\sim \alpha\sqrt{2\pi D}\,\exp\left[-\,U(a)/D\right] \tag{5.2.171}$$

and the inner factor becomes, on assuming x_0 is well to the right of the central point b,

$$\int_{a}^{x_0} dy \exp U(y)/D \sim \int_{-\infty}^{\infty} dy \exp\left[\frac{U(b)}{D} - \frac{(y-b)^2}{2D\delta^2}\right] \tag{5.2.172}$$

$$= \delta\sqrt{2\pi D}\,\exp\left[U(b)/D\right] . \tag{5.2.173}$$

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

$$\exp\left(\Delta E/kT\right) \tag{5.2.176}$$

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,

$$A(x)\partial_x g_a(x, t) + \tfrac{1}{2}B(x)\partial_x^2 g_a(x, t) = -\int_t^\infty dt' \partial_{t'} J(a, t' | x, 0)$$

$$= J(a, t | x, 0)$$

$$= \partial_t g_a(x, t) . \tag{5.2.181}$$

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

$$\boxed{\begin{aligned} \pi_a(x) &= [\int_x^b dy\, \psi(y)] / \int_a^b dy\, \psi(y) \\[4pt] \pi_b(x) &= [\int_a^x dy\, \psi(y)] / \int_a^b dy\, \psi(y) \end{aligned}} \tag{5.2.189} \tag{5.2.190}$$

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,

$$\partial_t p(\mathbf{x}, t) = - \sum_i \partial_i [A_i(\mathbf{x}) p(\mathbf{x}, t] + \tfrac{1}{2} \sum_{i,j} \partial_i \partial_j [B_{ij}(\mathbf{x}) p(\mathbf{x}, t)] \tag{5.3.1}$$

and we want to know the corresponding equation for the variables

$$y_i = f_i(\mathbf{x}), \tag{5.3.2}$$

where f_i are certain differentiable independent functions. Let us denote by $\bar{p}(\mathbf{y}, t)$ the probability density for the new variable, which is given by

$$\bar{p}(\mathbf{y}, t) = p(\mathbf{x}, t) \left| \frac{\partial(x_1, x_2 \ldots)}{\partial(y_1, y_2 \ldots)} \right|. \tag{5.3.3}$$

The simplest way to effect the change of variables is to use Ito's formula on the corresponding SDE

$$d\mathbf{x}(t) = A(\mathbf{x}) dt + \sqrt{\mathbf{B}(\mathbf{x})} \, d\mathbf{W}(t) \tag{5.3.4}$$

and then recompute the corresponding FPE for $\bar{p}(\mathbf{y}, t)$ 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

$$E_1 = a \cos\phi \tag{5.3.6}$$
$$E_2 = a \sin\phi.$$

The Jacobian is

$$|J| = \frac{\partial(E_1, E_2)}{\partial(a, \phi)} = \begin{vmatrix} \cos\phi & -a \sin\phi \\ \sin\phi & a \cos\phi \end{vmatrix}$$

$$= a. \tag{5.3.7}$$

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)

$$\left. \begin{array}{l} a = \sqrt{E_1^2 + E_2^2} \\ \phi = \tan^{-1}(E_2/E_1) , \end{array} \right\} \tag{5.3.9}$$

we note

$$\left. \begin{array}{l} \dfrac{\partial a}{\partial E_1} = \dfrac{E_1}{\sqrt{E_1^2 + E_2^2}} = \cos \phi . \\[2mm] \text{Similarly,} \\[2mm] \dfrac{\partial a}{\partial E_2} = \sin \phi \end{array} \right\} \tag{5.3.10}$$

and

$$\left. \begin{array}{l} \dfrac{\partial \phi}{\partial E_2} = \dfrac{E_1}{E_1^2 + E_2^2} = \cos \phi/a . \\[2mm] \text{Similarly,} \\[2mm] \dfrac{\partial \phi}{\partial E_1} = - \sin \phi/a . \end{array} \right\} \tag{5.3.11}$$

Hence,

$$\frac{\partial}{\partial E_1} E_1 p + \frac{\partial}{\partial E_2} E_2 p$$

$$= 2p + E_1 \left(\frac{\partial p}{\partial a} \frac{\partial a}{\partial E_1} + \frac{\partial p}{\partial \phi} \frac{\partial \phi}{\partial E_1} \right) + E_2 \left(\frac{\partial p}{\partial a} \frac{\partial a}{\partial E_2} + \frac{\partial p}{\partial \phi} \frac{\partial \phi}{\partial E_2} \right)$$

$$= 2p + a \frac{\partial p}{\partial a} = \frac{1}{a} \frac{\partial}{\partial a} (a^2 p) . \tag{5.3.12}$$

Let us use the symbol $\bar{p}(a, \phi)$ for the density function in terms of a and ϕ. The Jacobian formula (5.5.3) tells us that

$$\bar{p}(a, \phi) = \left| \frac{\partial(E_1, E_2)}{\partial(a, \phi)} \right| p(E_1, E_2) = a p(E_1, E_2) . \tag{5.3.13}$$

Putting together (5.3.5, 8, 12, 13), we get

$$\frac{\partial \bar{p}}{\partial t} = - \frac{\partial}{\partial a} \left[\left(-\gamma a + \frac{\varepsilon^2}{2a} \right) \bar{p} \right] + \frac{\varepsilon^2}{2} \left(\frac{1}{a^2} \frac{\partial^2 \bar{p}}{\partial \phi^2} + \frac{\partial^2 \bar{p}}{\partial a^2} \right) \tag{5.3.14}$$

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

$$p(\boldsymbol{x}, t) = 0 \quad \text{for } \boldsymbol{x} \text{ on } \boldsymbol{S}. \tag{5.3.17}$$

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

$$\boldsymbol{n} \cdot \boldsymbol{J}_1 = \boldsymbol{n} \cdot \boldsymbol{J}_2 \quad \text{on } \boldsymbol{S} \tag{5.3.18}$$

$$p_1(\boldsymbol{x}) = p_2(\boldsymbol{x}) \quad \boldsymbol{x} \text{ on } \boldsymbol{S}.$$

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

Absorbing Boundary $\quad p(\boldsymbol{x}, t | \boldsymbol{y}, t') = 0 \quad \boldsymbol{y} \in \boldsymbol{S}$ (5.3.19)

Reflecting Boundary $\quad \sum_{i,j} n_i B_{ij}(\boldsymbol{y}) \frac{\partial}{\partial y_j} p(\boldsymbol{x}, t | \boldsymbol{y}, t') = 0 \quad \boldsymbol{y} \in \boldsymbol{S}$. (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 $B_{ij}(x)$ has an inverse for all x, we can rewrite (5.3.21)

$$\frac{\partial}{\partial x_i} \log [p_s(x)] = \sum_k B_{ik}^{-1}(x) \left[2A_k(x) - \sum_j \frac{\partial}{\partial x_j} B_{kj}(x) \right] \tag{5.3.22}$$

$$\equiv Z_i[A, B, x] . \tag{5.3.23}$$

This equation cannot be satisfied for arbitrary $B_{ij}(x)$ and $A_i(x)$ 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):

$$p_s(x) = \exp \{ \int^x dx' \cdot Z[A, B, x'] \} . \tag{5.3.25}$$

The conditions (5.3.24) are known as *potential conditions* since we derive the quantities Z_i from derivatives of $\log [p_s(x)]$, which, therefore, is often thought of as a potential $-\phi(x)$ so that more precisely,

$$p_s(x) = \exp [-\phi(x)] \tag{5.3.26}$$

and

$$\phi(x) = -\int^x dx' \cdot Z[A, B, x'] . \tag{5.3.27}$$

Example: Rayleigh Process in Polar Coordinates. From (5.3.14) we find

$$A = \begin{bmatrix} -\gamma a + \varepsilon^2/2a \\ 0 \end{bmatrix} \tag{5.3.28}$$

$$B = \begin{bmatrix} \varepsilon^2 & 0 \\ 0 & \varepsilon^2/a^2 \end{bmatrix} \tag{5.3.29}$$

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

$$Z = 2 B^{-1}A = \begin{bmatrix} -2\gamma a/\varepsilon^2 + 1/a \\ 0 \end{bmatrix} \tag{5.3.31}$$

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 r and velocities v. Then a transition corresponds to a particle at some time t with position velocity (r, v) having acquired by a later time $t + \tau$ position and velocity (r', v'). The probability density of this transition is the *joint probability density* $p(r', v', t + \tau; r, v, t)$.

We may symbolically write this transition as

$$(r, v, t) \to (r', v', t + \tau) \, . \tag{5.3.36}$$

The reversed transition is not given simply by interchanging primed and unprimed quantities Rather, it is

$$(r', -v', t) \to (r, -v, t + \tau) \, . \tag{5.3.37}$$

It corresponds to the *time reversed transition* and requires the velocities to be reversed because the motion from r' to r is in the opposite direction from that from r to r'.

The probability density for the reversed transition is thus the joint probability density

$$p(r, -v, t + \tau; r', -v', t) \, . \tag{5.3.38}$$

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

$$p_s(\mathbf{r}', \mathbf{v}', \tau; \mathbf{r}, \mathbf{v}, 0) = p_s(\mathbf{r}, -\mathbf{v}, \tau; \mathbf{r}', -\mathbf{v}', 0) \tag{5.3.39}$$

(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)

$$p(\mathbf{r}', \mathbf{v}', \tau | \mathbf{r}, \mathbf{v}, 0)p_s(\mathbf{r}, \mathbf{v}) = p(\mathbf{r}, -\mathbf{v}, \tau | \mathbf{r}', -\mathbf{v}', 0)p_s(\mathbf{r}', -\mathbf{v}') \,, \tag{5.3.40}$$

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 reversal, transform to the reversed variables according to the rule

$$x_i \rightarrow \varepsilon_i x_i \tag{5.3.41}$$

$$\varepsilon_i = \pm 1 \tag{5.3.42}$$

depending on whether the variable is odd or even under time reversal. In the above, $\mathbf{r}$ is even, $\mathbf{v}$ is odd.

Then by detailed balance we require

$$p_s(\mathbf{x}, t + \tau; \mathbf{x}', t) = p_s(\boldsymbol{\varepsilon}\mathbf{x}', t + \tau; \boldsymbol{\varepsilon}\mathbf{x}, t) \,. \tag{5.3.43}$$

By $\boldsymbol{\varepsilon}\mathbf{x}$, we mean $(\varepsilon_1 x_1, \varepsilon_2 x_2, \ldots)$.
Notice that setting $\tau = 0$ in (5.3.43) we obtain

$$\delta(\mathbf{x} - \mathbf{x}')p_s(\mathbf{x}') = \delta(\boldsymbol{\varepsilon}\mathbf{x} - \boldsymbol{\varepsilon}\mathbf{x}')p_s(\boldsymbol{\varepsilon}\mathbf{x}) \,. \tag{5.3.44}$$

The two delta functions are equal since only sign changes are involved. Hence,

$$p_s(\mathbf{x}) = p_s(\boldsymbol{\varepsilon}\mathbf{x}) \tag{5.3.45}$$

is a consequence of the formulation of detailed balance by (5.3.43). Rewriting now in terms of conditional probabilities, we have

$$\boxed{p(\mathbf{x}, \tau | \mathbf{x}', 0)p_s(\mathbf{x}') = p(\boldsymbol{\varepsilon}\mathbf{x}', \tau | \boldsymbol{\varepsilon}\mathbf{x}, 0)p_s(\mathbf{x}) \,.} \tag{5.3.46}$$

b) General Consequences of Detailed Balance
An important consequence of (5.3.45) is that

$$\langle \mathbf{x} \rangle_s = \boldsymbol{\varepsilon}\langle \mathbf{x} \rangle_s \tag{5.3.47}$$

(hence all odd variables have zero stationary mean), and for the autocorrelation function

$$G(\tau) \equiv \langle \mathbf{x}(\tau)\mathbf{x}^{\mathrm{T}}(0) \rangle_s$$

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

$$S(\omega) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{-i\omega\tau} G(\tau) \,,$$

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

$$
\begin{array}{ll}
\text{(i)} & W(x\,|\,x')p_s(x') = W(\varepsilon x'\,|\,\varepsilon x)p_s(x) \\[2mm]
\text{(ii)} & \varepsilon_i A_i(\varepsilon x)p_s(x) = -A_i(x)p_s(x) + \sum_j \dfrac{\partial}{\partial x_j}[B_{ij}(x)p_s(x)] \\[2mm]
\text{(iii)} & \varepsilon_i \varepsilon_j B_{ij}(\varepsilon x) = B_{ij}(x)\,.
\end{array}
\tag{5.3.53}
$$

The specialisation to a FPE is simply done by setting the jump probabilities $W(x\,|\,x')$ 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 $W(x\,|\,x')$, $A_i(x)$ and $B_{ij}(x)$ (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)

$$
W(x\,|\,x')p_s(x') = W(\varepsilon x'\,|\,\varepsilon x)p_s(x)\,.
\tag{5.3.54}
$$

Consider now the drift coefficient. For simplicity write

$$
x' = x + \boldsymbol{\delta}\,.
\tag{5.3.55}
$$

Then from (5.3.46) we have

$$
\begin{aligned}
&\int\limits_{|\delta|<K} d\boldsymbol{\delta}\; \delta_i p(\varepsilon x + \varepsilon\boldsymbol{\delta},\, \Delta t\,|\,\varepsilon x,\, 0)p_s(x) \\
&= \int\limits_{|\delta|<K} d\boldsymbol{\delta}\; \delta_i p(x,\, \Delta t\,|\,x + \boldsymbol{\delta},\, 0)p_s(x + \boldsymbol{\delta})
\end{aligned}
\tag{5.3.56}
$$

(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

$$
\varepsilon_i A_i(\varepsilon x)p_s(x) + O(K)\,.
\tag{5.3.57}
$$

On the right-hand side we write

$$
\begin{aligned}
p(x + \boldsymbol{\delta} - \boldsymbol{\delta},\, \Delta t\,|\,x + \boldsymbol{\delta},\, 0)p_s(x + \boldsymbol{\delta}) &= p(x - \boldsymbol{\delta},\, \Delta t\,|\,x,\, 0)p_s(x) \\
&+ \sum_j \delta_j \frac{\partial}{\partial x_j}[p(x - \boldsymbol{\delta},\, \Delta t\,|\,x,\, 0)p_s(x)] + O(\delta^2)
\end{aligned}
\tag{5.3.58}
$$

so that the right-hand side is

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int d\boldsymbol{\delta} \left\{ \delta_i p(\boldsymbol{x} - \boldsymbol{\delta}, \Delta t \,|\, \boldsymbol{x}, 0) p_s(\boldsymbol{x}) + \sum_j \delta_i \delta_j \frac{\partial}{\partial x_j} [p(\boldsymbol{x} - \boldsymbol{\delta}, \Delta t \,|\, \boldsymbol{x}, 0) p_s(\boldsymbol{x})] \right\} + O(K)$$

$$= -A_i(\boldsymbol{x}) p_s(\boldsymbol{x}) + \sum_j \frac{\partial}{\partial x_j} [B_{ij}(\boldsymbol{x}) p_s(\boldsymbol{x})] + O(K), \qquad (5.3.59)$$

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

$$\varepsilon_i A_i(\varepsilon \boldsymbol{x}) p_s(\boldsymbol{x}) = -A_i(\boldsymbol{x}) p_s(\boldsymbol{x}) + \sum_j \frac{\partial}{\partial x_j} [B_{ij}(\boldsymbol{x}) p_s(\boldsymbol{x})]. \qquad (5.3.60)$$

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

$$\varepsilon_i \varepsilon_j B_{ij}(\varepsilon \boldsymbol{x}) = B_{ij}(\boldsymbol{x}). \qquad (5.3.61)$$

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

$$\bar{p}(\boldsymbol{x}, t \,|\, \boldsymbol{x}', 0) \equiv p(\varepsilon \boldsymbol{x}', t \,|\, \varepsilon \boldsymbol{x}, 0) p_s(\hat{\boldsymbol{x}}) / p_s(\boldsymbol{x}'). \qquad (5.3.62)$$

Clearly

$$\bar{p}(\boldsymbol{x}, 0 \,|\, \boldsymbol{x}', 0) = \delta(\boldsymbol{x} - \boldsymbol{x}') = p(\boldsymbol{x}, 0 \,|\, \boldsymbol{x}', 0). \qquad (5.3.63)$$

We substitute $\bar{p}$ 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 $\bar{p}$ is a solution of the *forward differential Chapman-Kolmogorov equation.*

We do this explicitly. The notation is abbreviated for clarity, so that we write

$$\left. \begin{array}{lll} \bar{p} & \text{for} & \bar{p}(\boldsymbol{x}, t \,|\, \boldsymbol{x}', 0) \\ p_s & \text{for} & p_s(\boldsymbol{x}) \\ p_s' & \text{for} & p_s(\boldsymbol{x}') \\ p(\boldsymbol{x}) & \text{for} & p(\boldsymbol{x}', t \,|\, \boldsymbol{x}, 0). \end{array} \right\} \qquad (5.3.64)$$

We proceed term by term.

i) **Drift Term:**

$$-\sum_i \frac{\partial}{\partial x_i} (A_i \bar{p}) = -\sum_i \frac{\partial}{\partial x_i} (A_i p(\varepsilon \boldsymbol{x}) p_s / p_s') \qquad (5.3.65)$$

$$= -\sum_i \left\{ \frac{\partial}{\partial x_i} [A_i p_s] p(\boldsymbol{\varepsilon x}) + A_i p_s \frac{\partial}{\partial x_i} p(\boldsymbol{\varepsilon x}) \right\} / p_s' \ .$$

ii) Diffusion Term:

$$\frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (B_{ij}\bar{p}) = \frac{1}{2} \sum_{i,j} \left[\frac{\partial^2}{\partial x_i \partial x_j} (B_{ij} p_s) p(\boldsymbol{\varepsilon x}) \right.$$

$$\left. + 2 \frac{\partial}{\partial x_i} (B_{ij} p_s) \frac{\partial}{\partial x_j} p(\boldsymbol{\varepsilon x}) + B_{ij} p_s \frac{\partial^2}{\partial x_i \partial x_j} p(\boldsymbol{\varepsilon x}) \right] / p_s' \ . \tag{5.3.66}$$

iii) Jump Term:

$$\int dz [W(\boldsymbol{x}|\boldsymbol{z})\bar{p}(\boldsymbol{z}, t|\boldsymbol{x}', 0) - W(\boldsymbol{z}|\boldsymbol{x})\bar{p}(\boldsymbol{x}, t|\boldsymbol{x}', 0)]$$

$$= \int dz [W(\boldsymbol{x}|\boldsymbol{z})p_s(\boldsymbol{z})p(\boldsymbol{\varepsilon x}', t|\boldsymbol{\varepsilon z}, 0) - W(\boldsymbol{z}|\boldsymbol{x})p_s(\boldsymbol{x})p(\boldsymbol{\varepsilon x}', t|\boldsymbol{\varepsilon x}, 0)]/p_s \ . \tag{5.3.67}$$

We now use the fact that $p_s(\boldsymbol{x})$ is a solution of the stationary differential Chapman-Kolmogorov equation to write

$$-\sum_i \left[\frac{\partial}{\partial x_i} (A_i p_s) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_i} (B_{ij} p_s) \right] - \int dz\, W(\boldsymbol{z}|\boldsymbol{x})p_s(\boldsymbol{x})$$

$$= -\int dz\, W(\boldsymbol{x}|\boldsymbol{z})p_s(\boldsymbol{z}) \tag{5.3.68}$$

and using the detailed balance condition (5.3.53(i)) for W

$$= -\int dz\, W(\boldsymbol{\varepsilon z}|\boldsymbol{\varepsilon x})p_s(\boldsymbol{x}) \ . \tag{5.3.69}$$

Now substitute

$$\boldsymbol{y} = \boldsymbol{\varepsilon x} \tag{5.3.70}$$

$$\boldsymbol{y}' = \boldsymbol{\varepsilon x}'$$

and all up all three contributions, taking care of (5.3.68, 69):

$$= \left\{ -\sum_i \varepsilon_i A_i(\boldsymbol{\varepsilon y})p_s(\boldsymbol{y}) \left[\frac{\partial}{\partial y_i} p(\boldsymbol{y}) \right] + \sum_{i,j} \varepsilon_i \varepsilon_j \frac{\partial}{\partial y_j} [B_{ij}(\boldsymbol{\varepsilon y})p_s(\boldsymbol{y})] \left[\frac{\partial}{\partial y_i} p(\boldsymbol{y}) \right] \right.$$

$$+ \frac{1}{2} \sum_{i,j} \varepsilon_i \varepsilon_j B_{ij}(\boldsymbol{\varepsilon y})p_s(\boldsymbol{y}) \left[\frac{\partial^2}{\partial y_i \partial y_j} p(\boldsymbol{y}) \right] \tag{5.3.71}$$

$$\left. + \int dz [W(\boldsymbol{\varepsilon y}|\boldsymbol{z})p_s(\boldsymbol{z})p(\boldsymbol{y}', t|\boldsymbol{\varepsilon z}, 0) - W(\boldsymbol{\varepsilon y}|\boldsymbol{z})p_s(\boldsymbol{z})p(\boldsymbol{y}', t|\boldsymbol{y}, 0)] \right\} / p_s(\boldsymbol{y}').$$

We now substitute the detailed balance conditions (5.3.53).

$$= \left\{ \sum_i A_i(\boldsymbol{y}) \frac{\partial}{\partial y_i} p(\boldsymbol{y}', t|\boldsymbol{y}, 0) + \frac{1}{2} \sum_{i,j} B_{ij}(\boldsymbol{y}) \frac{\partial^2}{\partial y_i \partial y_j} p(\boldsymbol{y}', t|\boldsymbol{y}, 0) \right.$$

$$\left. + \int dz [W(\boldsymbol{z}|\boldsymbol{y})p(\boldsymbol{y}', t|\boldsymbol{z}, 0) - W(\boldsymbol{z}|\boldsymbol{y})p(\boldsymbol{y}', t|\boldsymbol{y}, 0)] \right\} p_s(\boldsymbol{y})/p_s(\boldsymbol{y}') \ . \tag{5.3.72}$$

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

$$p(y', t|y, 0) = p(y, 0|y, -t) .$$

We see that

$$(5.3.72) = \frac{\partial}{\partial t} [p(y', t|, 0)p_s(y)/p_s(y')] = \frac{\partial}{\partial t} \hat{p}(x, t|x', 0) \qquad (5.3.73)$$

which means that $\hat{p}(x, t|x', 0)$, defined in (5.3.62), satisfies the forward differential Chapman-Kolmogorov equation. Since the initial condition of $p(x, t|x', 0)$ and $\hat{p}(x, t|x', 0)$ 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

$$W(x|x')p_s(x') = W(x'|x)p_s(x) \qquad (5.3.74)$$

$$A_i(x)p_s(x) = \frac{1}{2} \sum_j \frac{\partial}{\partial x_j} [B_{ij}(x)p_s(x)] \qquad (5.3.75)$$

$$B_{ij}(x) = B_{ij}(x) , \qquad (5.3.76)$$

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 J, the probability current in the stationary state.

The conditions (5.3.74, 75) taken together imply that $p_s(x)$ 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

$$D_i(x) = \frac{1}{2} [A_i(x) + \varepsilon_i A_i(\varepsilon x)] \qquad (5.3.77)$$

and the reversible drift

$$I_i(x) = \frac{1}{2} [A_i(x) - \varepsilon_i A_i(\varepsilon x)] . \qquad (5.3.78)$$

Using again the potential defined by

$$p_s(x) = \exp [-\phi(x)] , \qquad (5.3.79)$$

we see that in the case of a Fokker-Planck equation, we can write the conditions for detailed balance as

$$\varepsilon_i \varepsilon_j B_{ij}(\varepsilon x) = B_{ij}(x) \tag{5.3.80}$$

$$D_i(x) - \frac{1}{2} \sum_j \frac{\partial}{\partial x_j} [B_{ij}(x)] = -\frac{1}{2} \sum_i B_{ij}(x) \frac{\partial \phi(x)}{\partial x_j} \tag{5.3.81}$$

$$\sum_i \left[\frac{\partial}{\partial x_i} I_i(x) - I_i(x) \frac{\partial \phi(x)}{\partial x_i} \right] = 0 \tag{5.3.82}$$

where the last equation is simply the stationary FPE for $p_s(x)$, 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 $D_i(x)$ and $B_{ij}(x)$ are satisfied. If $B_{ij}(x)$ 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

$$\hat{Z}_i = \sum_k B_{ik}^{-1}(x) \left[2D_k(x) - \sum_j \frac{\partial}{\partial x_j} B_{kj}(x) \right] \tag{5.3.84}$$

and we have

$$p_s(x) = \exp[-\phi(x)] = \exp\left(\int^x dx' \cdot \hat{Z}\right). \tag{5.3.85}$$

Thus, as in the case of a vanishing probability current, $p_s(x)$ 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 $f(x, t)$ is a solution of the forward differential Chapman-Kolmogorov equation, then

$$\tilde{f}(x, t) = f(\varepsilon x, -t)/p_s(x) \tag{5.3.86}$$

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

$$A(y, u) = \begin{bmatrix} u \\ - U'(y) - \gamma u \end{bmatrix}, \tag{5.3.98}$$

$$\underset{\approx}{B}(y, u) = \begin{bmatrix} 0 & 0 \\ 0 & 2\gamma \end{bmatrix} \tag{5.3.99}$$

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 $\underset{\approx}{B}$ 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.,

$$p_s(y, u) = \exp\left(-\tfrac{1}{2}u^2\right)f(y) \tag{5.3.104}$$

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

$$f(y) = \mathcal{N}\exp\left[-U(y)\right] \tag{5.3.106}$$

and

$$p_s(y, u) = \mathcal{N} \exp \left[-U(y) - \tfrac{1}{2}u^2\right].$$ (5.3.107)

In terms of the original (x, v) variables,

$$p_s(x, v) = \mathcal{N} \exp \left[- \frac{V(x)}{kT} - \frac{mv^2}{2kT}\right]$$ (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 $B_{ij}(x)$ and $W(x|x')$ equal to zero, so the detailed balance conditions are simply

$$\varepsilon_i A_i(\varepsilon x) = -A_i(x).$$ (5.3.109)

Since we are now dealing with a Liouville equation (Sect. 3.5.3), the motion of a point whose coordinates are x is described by the ordinary differential equation

$$\frac{d}{dt} x(t) = A[x(t)].$$ (5.3.110)

Suppose a solution of (5.3.110) which passes through the point y at $t = 0$ is

$$q[t, y]$$ (5.3.111)

which therefore satisfies

$$q[0, y] = y.$$ (5.3.112)

Then the relation (5.3.109) implies that the reversed solution

$$\varepsilon q(-t, \varepsilon y)$$ (5.3.113)

is also a solution of (5.3.110), and since

$$\varepsilon q(0, \varepsilon y) = \varepsilon \varepsilon y = y,$$ (5.3.114)

i.e., the initial conditions are the same, these solutions must be identical, i.e.,

$$\varepsilon q(-t, \varepsilon y) = q(t, y).$$ (5.3.115)

Now the joint probability in the stationary state can be written as

$$p_s(x, t; x', t') = \int dy \, p_s(x, t; x', t'; y, 0)$$
$$= \int dy \, \delta[x - q(t, y)]\delta[x' - q(t', y)]p_s(y) \qquad (5.3.116)$$

and

$$p_s(\varepsilon x', -t'; \varepsilon x, -t) = \int dy \, \delta[\varepsilon x - q(-t, y)]\delta[\varepsilon x' - q(-t', y)]p_s(y). \qquad (5.3.117)$$

Change the variables from y to εy and note that $p_s(y) = p_s(\varepsilon y)$, and $d\varepsilon y = dy$, so that

$$(5.3.117) = \int dy \, \delta[x - \varepsilon q(-t, \varepsilon y)]\delta[x' - \varepsilon q(-t', \varepsilon y)]p_s(y) \qquad (5.3.118)$$

and using (5.3.115),

$$= \int dy \, \delta[x - q(t, y)]\delta[x' - q(t', y)]p_s(y) \qquad (5.3.119)$$
$$= p_s(x, t; x', t'). \qquad (5.3.120)$$

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 $A(x)$ 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 x.

Thus, we assume x can be written

$$x = (a, \hat{x}) \qquad (5.3.121)$$

where the vector a represents the macroscopically observable quantities and $\hat{x}$ is all the others. Then, in practice, we are interested in

$$\bar{p}(a_1, t_1; a_2, t_2; a_3, t_3; \ldots)$$
$$= \int\!\!\int \ldots \int d\hat{x}_1, d\hat{x}_2 \ldots p(x_1, t_1; x_2, t_2; x_3, t_3; \ldots). \qquad (5.3.122)$$

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

$$A_i(x) = \sum_j A_{ij} x_j \tag{5.3.123}$$

$$B_{ij}(x) = B_{ij} . \tag{5.3.124}$$

The detailed balance conditions are not trivial, however. Namely,

$$\sum_j (\varepsilon_i \varepsilon_j A_{ij} + A_{ij}) x_j = \sum_j B_{ij} \frac{\partial}{\partial x_j} \log p_s(x) \tag{5.3.125}$$

and

$$\varepsilon_i \varepsilon_j B_{ij} = B_{ij} . \tag{5.3.126}$$

Equation (5.3.125) has the qualitative implication that $p_s(x)$ is a Gaussian since derivative of $\log p_s(x)$ is linear in x. Furthermore, since the left-hand side contains no constant term, this Gaussian must have zero mean, hence, we can write

$$p_s(x) = \mathcal{N} \exp\left(- \tfrac{1}{2} x^T \sigma^{-1} x\right) . \tag{5.3.127}$$

One can now substitute (5.3.127) in the stationary Fokker-Planck equation and rearrange to obtain

$$-\sum_i A_{ii} - \tfrac{1}{2} \sum_{i,j} B_{ij} \sigma_{ij}^{-1} + \sum_{k,j}\left(\sum_i \sigma_{ki}^{-1} A_{ij} + \tfrac{1}{2} \sum_{i,l} \sigma_{ki}^{-1} B_{il} \sigma_{ij}^{-1}\right) x_k x_j = 0 \tag{5.3.128}$$

(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

$$\sigma^{-1} A + A^T \sigma^{-1} = -\sigma^{-1} B \sigma^{-1} \tag{5.3.129}$$

or

$$A\sigma + \sigma A^T = -B . \tag{5.3.130}$$

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 $\phi = \log[p_s(x)]$:

$$F(x) = -\nabla\phi(x) = \sigma^{-1}x \tag{5.3.140}$$

(in physics, ϕ/kT is the entropy of the system). Because of the linear form of the $A_i(x)$ [(5.3.123)], the exact equations of motion for $\langle x \rangle$ are

$$\frac{d}{dt} \langle x \rangle = A\langle x \rangle = A\sigma F(\langle x \rangle) . \tag{5.3.141}$$

Thus, if the *fluxes* $d/dt\langle x \rangle$ are related linearly to the *forces* $F(\langle x \rangle)$ by a matrix L defined by

$$L = A\sigma , \tag{5.3.142}$$

then (5.3.139) says

$$\varepsilon L \varepsilon = L^T \tag{5.3.143}$$

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

$$L = L^T \tag{5.3.146}$$

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.
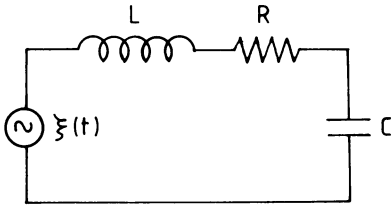
Fig. 5.4. Electric circuit used in the derivation of Nyquist's formula

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

$$A = - \begin{bmatrix} \gamma & -1 \\ \dfrac{1}{LC} & \dfrac{R}{L} \end{bmatrix}. \tag{5.3.154}$$

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,

$$\boldsymbol{\varepsilon} = \mathrm{diag}\,(1, -1) \tag{5.3.159}$$

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),

$$\frac{dE}{dt} = -i^2 R - \frac{\gamma q^2}{C} \tag{5.3.166}$$

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 $P_\lambda(x)$ of the forward Fokker-Planck and a set $Q_\lambda(x)$ of the backward Fokker-Planck equation. Thus,

$$-\sum_i \partial_i[A_i(x)P_\lambda(x)] + \tfrac{1}{2}\sum_{i,j} \partial_i\partial_j[B_{ij}(x)P_\lambda(x)] = -\lambda P_\lambda(x) \tag{5.3.167}$$

$$\sum A_i(x)\partial_i Q_{\lambda'}(x) + \tfrac{1}{2}\sum B_{ij}(x)\partial_i\partial_j Q_{\lambda'}(x) = -\lambda' Q_{\lambda'}(x) . \tag{5.3.168}$$

Whether $Q_{\lambda'}(x)$ and $P_\lambda(x)$ satisfy absorbing or reflecting boundary conditions, one can show, in a manner very similar to that used in Sect.5.2.4, that

$$-(\lambda - \lambda') \int_R dx\, P_\lambda(x)Q_{\lambda'}(x) = 0 \tag{5.3.169}$$

so that the $P_\lambda(x)$ and $Q_{\lambda'}(x)$ 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

$$\int dx\, P_\lambda(x)Q_\lambda(x) = \delta_{\lambda\lambda'} , \tag{5.3.170}$$

if the spectrum of eigenvalues λ is discrete. [The Kronecker $\delta_{\lambda\lambda'}$ is to be replaced

by $\delta(\lambda - \lambda_r)$ when the spectrum is continuous, except where we have reflecting boundary conditions, and $\lambda = 0$ is stationary. The normalisation of $p_s(x)$ then gives

$$\int dx\, P_0(x)Q_0(x) = \int dx\, P_0(x) = 1 \tag{5.3.171}$$

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 $p(\varepsilon x, -t)/p_s(x)$ is a solution of the backward Fokker-Planck equation so that, from the uniqueness of solutions, we can say

$$Q_\lambda(x) = \eta_\lambda P_\lambda(\varepsilon x)/p_s(x). \tag{5.3.172}$$

Here, $\eta_\lambda = \pm 1$ but is otherwise undetermined. For, if $\lambda = \lambda'$, we can then write (5.3.170) in the form

$$\int dx\, \frac{P_\lambda(x)\, P_\lambda(\varepsilon x)}{P_s(x)} = \eta_\lambda. \tag{5.3.173}$$

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 $P_\lambda(x)$ 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

$$Q_\lambda(x) = P_\lambda(x)/p_s(x) \tag{5.3.174}$$

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

$$\mathscr{L}(x) = -\sum_i \partial_i A_i(x) + \tfrac{1}{2} \sum_{i,j} \partial_i \partial_j B_{ij}(x) \tag{5.3.175}$$

and the backward Fokker-Planck operator as

$$\mathscr{L}^*(x) = \sum_i A_i(x)\partial_i + \tfrac{1}{2} \sum_{i,j} B_{ij}(x)\partial_i \partial_j. \tag{5.3.176}$$

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 $p_s(x)$ arises from the fact that for any $f(x)$,

$$\mathscr{L}(x)[f(x)p_s(x)] = p_s(x)\mathscr{L}^*(x)[f(x)]. \tag{5.3.177}$$

Define a Hilbert space of functions $f(x)$, $g(x)$ ... by the scalar product

$$(f, g) = \int_R dx \frac{f(x)g(x)}{p_s(x)}. \tag{5.3.178}$$

Then from (5.3.177),

$$(f, \mathscr{L}g) = \int_R dx \frac{f(x)}{p_s(x)} \mathscr{L}(x)\left[\frac{g(x)}{p_s(x)}p_s(x)\right] \tag{5.3.179}$$

$$= \int_R dx\, f(x)\mathscr{L}^*(x)\left[\frac{g(x)}{p_s(x)}\right] \tag{5.3.180}$$

and integrating by parts, discarding surface terms by use of either reflecting or absorbing boundary conditions

$$= \int_R dx \frac{g(x)}{p_s(x)} \mathscr{L}(x)[f(x)]. \tag{5.3.181}$$

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 $\mathscr{L}(x)$ are real. To prove negative semi-definiteness, notice that for even variables only, [see(5.3.75)]

$$A_i(x) = \sum_j \partial_j[B_{ij}(x)p_s(x)]/2p_s(x) \tag{5.3.183}$$

so that for any $p(x)$,

$$\mathscr{L}(x)p(x) = \sum_{i,j} \partial_i \left\{-\frac{\partial_j[B_{ij}(x)p_s(x)]p(x)}{2p_s(x)} + \tfrac{1}{2}\partial_j[B_{ij}(x)p(x)]\right\}$$

$$= \frac{1}{2} \sum_{i,j} \partial_i \{B_{ij}(x)p_s(x)\partial_j[p(x)/p_s(x)]\}. \tag{5.3.184}$$

Hence,

$$(p, \mathscr{L}p) = \frac{1}{2} \int dx\, p(x)/p_s(x) \sum_{i,j} \partial_i \{B_{ij}(x)p_s(x)\partial_j[p(x)/p_s(x)]\}$$

and integrating by parts (discarding surface terms),

$$(p, \mathscr{L}p) = - \int dx\, B_{ij}(x)p_s(x)\partial_i[p(x)/p_s(x)]\partial_j[p(x)/p_s(x)] \tag{5.3.185}$$
$$\leqslant 0$$

since $B_{ij}(x)$ is positive semi-definite.

Hence, we conclude for any eigenfunction $P_\lambda(x)$ 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 $f(x)$ and expand it in eigenfunctions

$$f(x) = \sum_\lambda \alpha_\lambda P_\lambda(x) . \tag{5.3.187}$$

Further, let us fix $(f, f) = \int dx \, f(x)^2/p_s(x) = 1$.
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.,

$$\begin{aligned} \alpha_0 &\neq 0 \\ \alpha_\lambda &= 0 \quad \text{for} \quad \lambda \neq 0 . \end{aligned} \tag{5.3.189}$$

Now choose $f(x)$ 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)

$$\alpha_\lambda = 0 \quad \text{for all other } \lambda,$$

where λ_1 is the next smallest eigenvalue. This means that $P_{\lambda_1}(x)$ 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

$$p(x, t \,|\, x_0, 0) = \sum_\lambda P_\lambda(x) Q_\lambda(x_0) e^{-\lambda t} . \tag{5.3.192}$$

iv) **Autocorrelation Matrix.** If detailed balance is valid, for the stationary autocorrelation matrix [using (5.3.172)] we have

$$G(t) = \langle x(t)x^{\mathrm{T}}(0)\rangle = \sum \eta_\lambda e^{-\lambda t} [\int_R dx \; xP_\lambda(x)] [\int_R dx \; xP_\lambda(x)]^{\mathrm{T}}\varepsilon \tag{5.3.193}$$

which explicitly satisfies the condition

$$G(t) = \varepsilon G^{\mathrm{T}}(t)\varepsilon \tag{5.3.194}$$

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,

$$U_\lambda = \int_R dx \; xP_\lambda(x) \,, \tag{5.3.196}$$

then

$$S(\omega) = \frac{1}{2\pi} \sum_\lambda \left[\frac{\eta_\lambda}{\lambda + i\omega} U_\lambda U_\lambda^{\mathrm{T}}\varepsilon + \frac{\eta_\lambda}{\lambda - i\omega} \varepsilon U_\lambda U_\lambda^{\mathrm{T}} \right]. \tag{5.3.197}$$

If any of the λ are complex, from the reality of $\mathscr{L}(x)$ we find that the eigenfunction belonging to λ^* is $[P_\lambda(x)]^*$, $\eta_\lambda = \eta_\lambda^*$ and $[U_\lambda]^* = U_{\lambda*}$. The spectrum is then obtained by adding the complex conjugate of those terms involving complex eigenvalues to (5.3.197).

In the case where $\varepsilon = 1$ and hence $\eta_\lambda = 1$, the spectrum matrix has the simpler form

$$S(\omega) = \frac{1}{\pi} \sum_\lambda \frac{\lambda U_\lambda U_\lambda^{\mathrm{T}}}{\lambda^2 + \omega^2} \tag{5.3.198}$$

which is explicitly a positive definite matrix. The spectrum of a single variable q made up as a linear combination of the x by a formula such as

$$q = m \cdot x \tag{5.3.199}$$

is given by

$$S_q(\omega) = m^{\mathrm{T}} S(\omega) m = \frac{1}{\pi} \sum_\lambda \frac{\lambda (m \cdot U_\lambda)^2}{\lambda^2 + \omega^2} \tag{5.3.200}$$

and is a strictly decreasing function of ω.

In the case where $\varepsilon \neq 1$, 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,

$$p(\mathbf{x}', t \,|\, \mathbf{x}, 0) = 0 \qquad (\mathbf{x} \in S) . \tag{5.4.1}$$

The probability that the particle, initially at $\mathbf{x}$, is somewhere within R after a time t is

$$G(\mathbf{x}, t) = \int_R d\mathbf{x}' p(\mathbf{x}', t \,|\, \mathbf{x}, 0) \tag{5.4.2}$$

and if T is the time at which the particle leaves R, then

$$\text{Prob}(T \geqslant t) = G(\mathbf{x}, t) . \tag{5.4.3}$$

Since the process is homogeneous, we find that $G(\mathbf{x}, t)$ obeys the backward Fokker-Planck equation

$$\partial_t G(\mathbf{x}, t) = \sum_i A_i(\mathbf{x}) \partial_i G(\mathbf{x}, t) + 2 \sum_{i,j} B_{ij}(\mathbf{x}) \partial_i \partial_j G(\mathbf{x}, t) . \tag{5.4.4}$$

The initial conditions on (5.4.4) will arise from:

i) $p(\mathbf{x}', 0 \,|\, \mathbf{x}, 0) = \delta(\mathbf{x} - \mathbf{x}')$ \hfill (5.4.5)

so that

$$\begin{aligned} G(\mathbf{x}, 0) &= 1 \qquad \mathbf{x} \in R \\ &= 0 \qquad \text{elsewhere;} \end{aligned} \tag{5.4.6}$$

ii) the boundary condition (5.4.1) requires

$$G(\mathbf{x}, t) = 0 \qquad \mathbf{x} \in S . \tag{5.4.7}$$

As in Sect. 5.2.7, we find that these imply that the *mean exit time* from R starting at x, for which we shall use the symbol $T(x)$, satisfies

$$\sum_i A_i(x)\partial_i T(x) + 2\sum_{i,j} B_{ij}(x)\partial_i\partial_j T(x) = -1 \tag{5.4.8}$$

with the boundary condition

$$T(x) = 0 \quad x \in S \tag{5.4.9}$$

and the nth moments

$$T_n(x) = \langle T^n \rangle = \int_0^\infty t^{n-1} G(x, t)dt \tag{5.4.10}$$

satisfy

$$-n\, T_{n-1}(x) = \sum_i A_i(x)\partial_i T_n(x) + \sum_{i,j} B_{ij}(x)\partial_i\partial_j T_n(x) \tag{5.4.11}$$

with the boundary conditions

$$T_n(x) = 0 \quad x \in S\,. \tag{5.4.12}$$

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 $G(x, t)$ are then, from those derived in Sect. 5.2.4,

$$\sum_{i,j} n_i B_{ij}(x)\partial_j G(x, t) = 0 \quad (x \in S_r) \tag{5.4.13}$$

$$G(x, t) = 0 \quad (x \in S_a) \tag{5.4.14}$$

and hence,

$$\sum_{i,j} n_i B_{ij}(x)\partial_j T_n(x) = 0 \quad (x \in S_r) \tag{5.4.15}$$

$$T_n(x) = 0 \quad (x \in S_a)\,. \tag{5.4.16}$$

5.4.1   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

$$dx = -kx\,dt + \sqrt{D}\,dW_1(t)$$
$$dy = -ky\,dt + \sqrt{D}\,dW_2(t)$$
(5.4.17)

and want to know the mean exit time from the region

$$x^2 + y^2 \leqslant a^2$$
(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

$$r = \sqrt{x^2 + y^2}$$
(5.4.19)

by changing variables as in Sect.4.4.5, namely,

$$dr = (-kr + \tfrac{1}{2} D/r)dt + \sqrt{D}\,dW(t)$$
(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

$$
\left.
\begin{aligned}
U(x) &\to \tfrac{1}{2}kr^2 - \tfrac{1}{2}D \log r \\
D &\to \tfrac{1}{2}D \\
x_0 &\to a \\
a &\to \sqrt{x_0^2 + y_0^2} \equiv r_0 \\
-\infty &\to 0\,.
\end{aligned}
\right\}
$$
(5.4.21)

Thus, $T(r_0 \to a)$

$$= \frac{1}{D} \int_{r_0}^{a} y^{-1} \exp[ky^2/D]dy \int_0^y z \exp(-kz^2/D)dz \,.$$
(5.4.22)

The problem is thus essentially one dimensional. This does not often happen.

b) Application of Eigenfunctions
Suppose we use the eigenfunctions $Q_\lambda(x)$ and $P_\lambda(x)$ to expand the mean first passage time as

$$T(x) = \sum t_\lambda Q_\lambda(x).$$
(5.4.23)

   We suppose that the $P_\lambda(x)$ and $Q_\lambda(x)$ satisfy the boundary conditions required for the particular exit problem being studied, so that $T(x)$ as written in (5.4.23) satisfies the appropriate boundary conditions.
   We then expand

$$-1 = \sum_\lambda I_\lambda Q_\lambda(x)\,,$$
(5.4.24)

where

$$I_\lambda = -\int_R d\boldsymbol{x}\, P_\lambda(\boldsymbol{x})\,. \tag{5.4.25}$$

Inserting (5.4.23) into (5.4.8) yields

$$\lambda t_\lambda = -I_\lambda \tag{5.4.26}$$

so that

$$T(\boldsymbol{x}) = \sum_\lambda \frac{1}{\lambda}\, Q_\lambda(\boldsymbol{x}) \int_R d\boldsymbol{x}'\, P_\lambda(\boldsymbol{x}')\,. \tag{5.4.27}$$

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

$$\sum_i A_i(\boldsymbol{x})\partial_i f(\boldsymbol{x}) + \frac{1}{2}\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\partial_j f(\boldsymbol{x}) = 0 \tag{5.4.28}$$

since λ_1 is very small. Hence,

$$Q_1(\boldsymbol{x}) \sim K \tag{5.4.29}$$

where K is a constant. Taking account of the bi-orthonormality of P_λ and Q_λ, we see

$$1 = \int d\boldsymbol{x}\, P_1(\boldsymbol{x})Q_1(\boldsymbol{x}) \sim K \int d\boldsymbol{x}\, P_1(\boldsymbol{x}) \tag{5.4.30}$$

so that

$$T(\boldsymbol{x}) \sim 1/\lambda_1\,. \tag{5.4.31}$$

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

$$\int_R dx\, dy\, P_{n,m}(x, y) = 0 \qquad \text{(either n or m even)} \tag{5.4.36}$$

$$= \frac{4}{mn\pi^2} \quad \text{(m and n both odd)} .$$

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 $dS(a)$ at a of the boundary S of the region R. We assume absorption on all S.

The probability that the particle exits through $dS(a)$ after time t is

$$g(a, x, t)|dS(a)| = -\int_t^\infty dt'\, J(a, t'|x, 0) \cdot dS(a) . \tag{5.4.38}$$
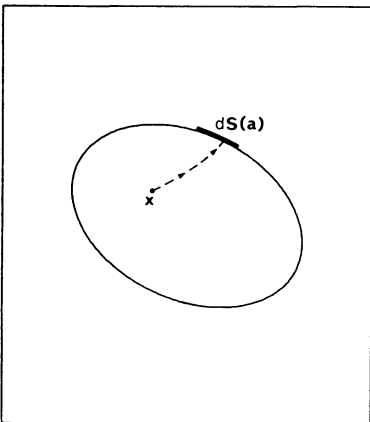
Fig. 5.5. Region and surface considered in Sect. 5.4.2

Similar reasoning to that in Sect.5.2.8 shows that $g(a, x, t)$ obeys the backward Fokker-Planck equation

$$\sum_i A_i(x)\partial_i g(a, x, t) + \tfrac{1}{2} \sum_{i,j} B_{ij}(x)\partial_i\partial_j g(a, x, t) = \partial_t g(a, x, t) . \qquad (5.4.39)$$

The boundary conditions follow by definition. Initially we have

$$g(a, x, 0) = 0 \qquad \text{for} \quad x \neq a, x \in R \qquad (5.4.40)$$

and at all times

$$g(a, x, t) = 0 \qquad \text{for} \quad x \neq a, x \in S . \qquad (5.4.41)$$

If $x = a$, then exit through $dS(a)$ is certain, hence,

$$g(a, a, t)\, dS(a) = 1 \qquad \text{for all } t \qquad (5.4.42)$$

or effectively

$$g(a, x, t) = \delta_s(a - x) \qquad x \in S, \quad \text{for all } t, \qquad (5.4.43)$$

where $\delta_s(a - x)$ is an appropriate surface delta function such that

$$\int |\, dS(a)|\, \delta_s(a - x) = 1 . \qquad (5.4.44)$$

The probability of ultimate exit through $dS(a)$ is

$$\pi(a, x)|\, dS(a)| = g(a, x, 0) |\, dS(a)| . \qquad (5.4.45)$$

The mean exit time given that exit occurred at a is

$$T(a, x) = \int_0^\infty dt\, g(a, x, t)/\pi(a, x) \qquad (5.4.46)$$

and in the same way as in Sect.5.2.8, we show that this satisfies

$$\boxed{\sum_i A_i(x)\partial_i[\pi(a, x)T(a, x)] + \tfrac{1}{2} \sum_{i,i} B_{ij}(x)\partial_i\partial_j[\pi(a, x)T(a, x)] = -\pi(a, x)}$$

$$(5.4.47)$$

and the boundary conditions is

$$\pi(a, x)T(a, x) = 0, \qquad x \in S . \qquad (5.4.48)$$

Further, by letting $t \to \infty$ in the corresponding Fokker-Planck equation for $g(a, x, t)$, we obtain the equation for $\pi(a, x)$:

$$\sum A_i(x)\partial_i[\pi(a, x)] + \tfrac{1}{2} \sum_{i,j} B_{ij}(x)\partial_i\partial_j[\pi(a, x)] = 0 .$$

(5.4.49)

The boundary condition for (5.4.49) is

$$\pi(a, x) = 0 \qquad a \neq x, \qquad x \in S$$

(5.4.50)

and

$$\int |dS(a)|\pi(a, x) = 1 .$$

(5.4.51)

Thus, we can summarise this as

$$\pi(a, x) = \delta_s(a - x) \qquad x \in S ,$$

(5.4.52)

where $\delta_s(a - x)$ is the surface delta function for the boundary 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

$$\langle x(t) \rangle = \alpha(t) \equiv c\,e^{-kt} \tag{6.1.6}$$

$$\mathrm{var}\{x(t)\} = \varepsilon^2 \beta(t) \equiv \varepsilon^2 (1 - e^{-2kt})/2k \tag{6.1.7}$$

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

$$a(x) = a(x_0 + \sum_{m=1}^{\infty} \varepsilon^m x_m)$$

$$= \sum_{p=0}^{\infty} \frac{1}{p!} \frac{d^p a(x_0)}{dx_0^p} \left(\sum_{m=1}^{\infty} \varepsilon^m x_m \right)^p . \tag{6.2.5}$$

Formally resumming is not easy, but it is straightforward to compute the first few powers of ε and to obtain

$$a_0(x_0) = a(x_0)$$

$$a_1(x_0, x_1) = x_1 \frac{da(x_0)}{dx_0}$$

$$a_2(x_0, x_1, x_2) = x_2 \frac{da(x_0)}{dx_0} + \frac{1}{2} x_1^2 \frac{d^2 a(x_0)}{dx_0^2} \tag{6.2.6}$$

$$a_3(x_0, x_1, x_2, x_3) = x_3 \frac{da(x_0)}{dx_0} + x_1 x_2 \frac{d^2 a(x_0)}{dx_0^2} + \frac{1}{6} x_1^3 \frac{d^3 a(x_0)}{dx_0^3} .$$

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

$$b(x) = b_0(x_0) + \varepsilon\, b_1(x_0, x_1) + \varepsilon^2 b_2(x_0, x_1, x_2) + \ldots, \tag{6.2.8}$$

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

$$dx_0 = a(x_0)dt \tag{6.2.10a}$$

$$dx_n = [-k(x_0)x_n + A_n(x_0, \ldots x_{n-1})]dt + b_{n-1}(x_0, \ldots x_{n-1})dW(t). \tag{6.2.10b}$$

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)

$$x_0(0) = x(0)$$
$$x_n(0) = 0 \qquad n \geqslant 1. \tag{6.2.11}$$

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

$$\exp\{\int\limits_0^t dt'k[\alpha(t')]\}$$

and is

$$x_1(t) = \int\limits_0^t b[\alpha(t')] \exp\{-\int\limits_{t'}^t k[\alpha(s)ds\} dW(t') , \tag{6.2.14}$$

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

$$x_N(t) = \int\limits_0^t [A_N(t')dt' + b_{N-1}(t')dW(t')] \exp\{-\int\limits_{t'}^t k[\alpha(s)]ds\} . \tag{6.2.16}$$

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

$$y_n(\varepsilon, t) = [x(t) - \sum\limits_{r=0}^n \varepsilon^r x_r(t)]/\varepsilon^{n+1} , \tag{6.2.17}$$

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

$$a[x(t)] = a[\sum\limits_{r=0}^n \varepsilon^r x_r(t) + y_n(\varepsilon, t)\varepsilon^{n+1}] \tag{6.2.18}$$

and we define a function $\hat{a}_{n+1}[X_0, X_1, X_2 \ldots X_n, Y, \varepsilon]$ by

$$\hat{a}_{n+1}[X_0, X_1, \ldots X, Y, \varepsilon] = \varepsilon^{-n-1} \{ a [\sum_{r=0}^{n} \varepsilon^r X_r + \varepsilon^{n+1} Y]$$

$$- \sum_{r=0}^{n} \varepsilon^r a_r (X_0, X_1, \ldots X_r) \} .$$

(6.2.19)

We require that for all fixed $X_0, X_1, \ldots X_N, Y$,

$$\lim_{\varepsilon \to 0} \hat{a}_{n+1}[X_0, X_1, \ldots X_n, Y, \varepsilon]$$

(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

$$dy_n = \hat{a}_{n+1}[x_0(t), x_1(t), \ldots x_n(t), y_n, \varepsilon] \, dt$$
$$+ \hat{b}_n[x_0(t), x_1(t), \ldots x_{n-1}(t), y_n, \varepsilon] \, dW(t) .$$

(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

$$\text{st-}\lim_{\varepsilon \to 0} \hat{a}_{n+1}[x_0(t), x_1(t), \ldots x_n(t), y_n, \varepsilon] \equiv \tilde{a}_{n+1}(t, y_n)$$

(6.2.22)

and

$$\text{st-}\lim_{\varepsilon \to 0} \hat{b}_n [x_0(t), x_1(t), \ldots x_{n-1}(t), y_n, \varepsilon] \equiv \tilde{b}_n(t, y_n)$$

(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

$$dy_n(0, t) = \tilde{a}_{n+1}[t, y_n(0, t)] + \tilde{b}_n[t, y_n(0, t)]dW(t)$$

(6.2.24)

which, from the definition (6.2.17) shows that

$$x(t) - \sum_{r=0}^{n} \varepsilon^r x_r(t) \sim \varepsilon^{n+1} .$$

(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

$$\text{var}\{x(t)\} = \varepsilon^2 \text{var}\{x_1(t)\} + 2\varepsilon^3 \langle x_1(t), x_2(t) \rangle$$
$$+ \varepsilon^4[2\langle x_1(t), x_3(t) \rangle + \text{var}\{x_2(t)\}] + \ldots \tag{6.2.31}$$

and similarly,

$$\langle x(t), x(s) \rangle = \varepsilon^2 \langle x_1(t), x_1(s) \rangle + \varepsilon^3[\langle x_1(t), x_2(s) \rangle + \langle x_1(s), x_2(t) \rangle]$$
$$+ \varepsilon^4[\langle x_1(t), x_3(s) \rangle + \langle x_1(s), x_3(t) \rangle + \langle x_2(t), x_2(s) \rangle]$$
$$+ \ldots . \tag{6.2.32}$$

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

$$p_s(x) = \mathcal{N} \exp(-x^4/2\varepsilon^2) \tag{6.2.37}$$

and the moments are

$$\langle x^n \rangle = (2\varepsilon^2)^{n/4} \, \Gamma\left(\frac{n+1}{4}\right) / \Gamma\left(\frac{1}{4}\right) \qquad (n \text{ even}) \tag{6.2.38}$$
$$= 0 \qquad\qquad\qquad\qquad\qquad\qquad (n \text{ odd}).$$

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

$$x = \sqrt{\varepsilon}\, y$$
$$t = \tau/\varepsilon$$
(6.2.39)

and use

$$dW(\tau/\varepsilon) = dW(\tau)/\sqrt{\varepsilon}\,,$$
(6.2.40)

then the cubic process can be recuced to a parameterless form

$$dy = -y^3 d\tau + dW(\tau)\,.$$
(6.2.41)

Regarding the solution of (6.2.41) as a known quantity, we can write

$$x(t) = \sqrt{\varepsilon}\, y(\varepsilon t)$$
(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

$$a(x) = -x^3 c(x)\,,$$
(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

$$dy = -y^3 c(y\sqrt{\varepsilon})d\tau + b(y\sqrt{\varepsilon})dW(\tau)\,.$$
(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

$$y(t) = \sum_{n=0}^{\infty} \varepsilon^{n/2} y_n(t)\,,$$
(6.2.45)

then we get for the first two terms

$$dy_0 = -y_0^3 c(0)d\tau + b(0)dW(\tau)$$
(6.2.46)

$$dy_1 = -y_1\left[3y_0^2 c(0) + y_0^3 \frac{dc}{dx}(0)\right]d\tau + \left[\frac{db}{dx}(0)y_0\right]dW(\tau)\,.$$
(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 v Processes (v odd)
If instead we have the Stochastic differential equation

$$dx = -x^v dt + \varepsilon \, dW(t), \tag{6.2.48}$$

we find a stable deterministic state only if v is an odd integer. In this case we make the transformation to scaled variables

$$x = y\varepsilon^{2/(1+v)} \tag{6.2.49}$$

$$t = \tau \, \varepsilon^{2(1-v)/(1+v)} \tag{6.2.50}$$

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

$$\frac{\partial\hat{p}_0}{\partial s} = -\tilde{A}_1(s)\frac{\partial}{\partial y}(y\hat{p}_0) + \frac{1}{2}B_0(s)\frac{\partial^2\hat{p}_0}{\partial y^2} \tag{6.3.12}$$

$$\frac{\partial \tilde{p}_r}{\partial s} = -\frac{\partial}{\partial y}[\tilde{A}_1(s)y\tilde{p}_1 + \tilde{A}_2(s)y^2\tilde{p}_0] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[\tilde{B}_0(s)\tilde{p}_1 + \tilde{B}_1(s)y\tilde{p}_0] \qquad (6.3.13)$$

and, in general,

$$\frac{\partial \tilde{p}_r}{\partial s} = -\frac{\partial}{\partial y}\left[\sum_{m=0}^{r} y^{r-m+1}\tilde{A}_{r-m+1}(s)\tilde{p}_m\right] + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[\sum_{m=0}^{r} y^{r-m}\tilde{B}_{r-m}(s)\tilde{p}_m\right]. \qquad (6.3.14)$$

Only the equation for $\tilde{p}_0$ is a FPE and, as mentioned in Sect.6.1, only $\tilde{p}_0$ 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 $\tilde{p}_r$ 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

$$y = \frac{[a - \alpha(s)]}{\varepsilon} \qquad (6.3.15)$$

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

$$\langle [y(t)]^n \rangle = \sum_{r=0}^{\infty} \varepsilon^r M_r^n(t). \qquad (6.3.16)$$

Then clearly

$$M_r^n(t) = \int dy\, y^n \tilde{p}_r(y, t) \qquad (6.3.17)$$

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

$$M_0^1(\infty) = 0$$
$$M_0^2(\infty) = -\tfrac{1}{2}\tilde{B}_0/\tilde{A}_1$$

$$M_0^3(\infty) = 0$$
$$M_1^1(\infty) = -\tilde{A}_2 M_0^2(\infty)/\tilde{A}_1 = \tfrac{1}{2}\,\tilde{A}_2\tilde{B}_0/(\tilde{A}_1)^2 \tag{6.3.29}$$
$$M_1^2(\infty) = -\tilde{A}_2 M_0^3(\infty)/\tilde{A}_1 = 0$$
$$M_2^1(\infty) = -[\tilde{A}_2 M_1^2(\infty) + \tilde{A}_3 M_0^3(\infty)]/\tilde{A}_1 = 0 \;.$$

Thus, the stationary mean, variance, etc., are

$$\langle x \rangle_s = \alpha + \varepsilon[M_0^1(\infty) + \varepsilon M_1^1(\infty) + \varepsilon^2 M_2^1(\infty)]$$
$$= \alpha + \tfrac{1}{2}\varepsilon^2 \tilde{A}_2 \tilde{B}_0/(\tilde{A}_1)^2 \tag{6.3.30}$$

$$\mathrm{var}\{x\}_s = \langle x^2 \rangle_s - \langle x_s \rangle^2$$
$$= \langle (\alpha + \varepsilon y)^2 \rangle_s - \langle \alpha + \varepsilon y \rangle_s^2 = \varepsilon^2 \mathrm{var}\{y\}_s$$
$$= -\frac{\varepsilon^2}{2}\,\tilde{B}_0/\tilde{A}_1 \qquad \text{to order } \varepsilon^2 \;. \tag{6.3.31}$$

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

$$\frac{d}{dt}\langle y(t)^n y(0) \rangle_s = \left\langle \left\{ \frac{A[\alpha + \varepsilon y(t)] - A(\alpha)}{\varepsilon} ny(t)^{n-1} \right. \right.$$
$$\left. \left. + \tfrac{1}{2}n(n-1)B[\alpha + \varepsilon y(t)]y(t)^{n-2} \right\} y(0) \right\rangle_s \tag{6.3.33}$$

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

$$\frac{d}{dt}\langle y(t)^n y(0) \rangle_s = \sum_{q=0}^{\infty} \varepsilon^q \left[n\tilde{A}_{q+1}\langle y(t)^{q+n} y(0) \rangle_s \right.$$
$$\left. + \frac{n(n-1)}{2}\tilde{B}_q\langle y(t)^{q+n-2} y(0) \rangle_s \right]. \tag{6.3.34}$$

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

$$\begin{aligned}
\langle x\rangle_s &= 0 \\
\text{var}\{x\}_s &= \varepsilon^2/6 .
\end{aligned} \tag{6.3.38}$$

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

$$\begin{bmatrix} \dfrac{dc_1}{dt} \\[2mm] \dfrac{dc_3}{dt} \end{bmatrix} = \begin{bmatrix} -1 & -\varepsilon^2 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_3 \end{bmatrix} \tag{6.3.40}$$

[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

$$\lambda_1 = -2 + \sqrt{1 - \varepsilon^2}$$
$$\lambda_2 = -2 - \sqrt{1 - \varepsilon^2}$$

(6.3.41)

and the corresponding eigenvectors

$$\boldsymbol{u}_1 = \begin{bmatrix} 1 + \sqrt{1 - \varepsilon^2} \\ 1 \end{bmatrix}$$
$$\boldsymbol{u}_2 = \begin{bmatrix} 1 - \sqrt{1 - \varepsilon^2} \\ 1 \end{bmatrix}.$$

(6.3.42)

The solution of (6.3.40) can then be written

$$\begin{bmatrix} c_1(t) \\ c_3(t) \end{bmatrix} = \alpha_1 e^{-\lambda_1 t} \boldsymbol{u}_1 + \alpha_2 e^{-\lambda_2 t} \boldsymbol{u}_2 \qquad (t > 0).$$

(6.3.43)

The initial condition is

$$c_1(0) = \langle y^2 \rangle_s$$
$$c_3(0) = \langle y^4 \rangle_s .$$

(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

$$M_0^4 = -\frac{3}{2} \frac{\tilde{B}_0 M_0^2}{\tilde{A}_1} = \frac{1}{12}$$

(6.3.45)

then $\langle y^4 \rangle = 1/12$.

Hence, we obtain

$$\begin{bmatrix} 1/6 \\ 1/12 \end{bmatrix} = \alpha_1 \boldsymbol{u}_1 + \alpha_2 \boldsymbol{u}_2$$

(6.3.46)

which have the solutions

$$\alpha_1 = \frac{1}{24}(1 + \sqrt{1 - \varepsilon^2})/\sqrt{1 - \varepsilon^2}$$
$$\alpha_2 = \frac{1}{24}(-1 + \sqrt{1 - \varepsilon^2})/\sqrt{1 - \varepsilon^2} .$$

(6.3.47)

The correlation function is, to 2nd order in ε (many terms cancel)

$$c_1(t) = \frac{1}{6} e^{-|\lambda_1|t} .$$

(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

$$[\sum_i A_i(\boldsymbol{x})\partial_i\phi + \tfrac{1}{2}\sum_{i,j} B_{ij}(\boldsymbol{x})\partial_i\phi\partial_j\phi] + \varepsilon^2[-\sum_i \partial_i A_i(\boldsymbol{x}) + \sum_{i,j} \partial_i B_{ij}\partial_j\phi$$
$$+ \sum_{i,j} \partial_i\partial_i B_{ij}(\boldsymbol{x})] = 0 . \tag{6.3.54}$$

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 $\phi_0(\boldsymbol{x})$ 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

$$\frac{dx}{dt} = v \qquad (6.4.1)$$

$$m\frac{dv}{dt} = -\beta v + \sqrt{2k\beta T}\,\xi(t)\,. \qquad (6.4.2)$$

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

$$v(t) = v(0)\exp(-\beta t/m) + \frac{\sqrt{2kT\beta}}{m}\int_0^t \exp[-\beta(t-t')/m]\xi(t')dt'\,. \qquad (6.4.3)$$

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

$$t \gg m/\beta \equiv \tau\,, \qquad (6.4.4)$$

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,

$$v(t) \to \frac{\sqrt{2kT\beta}}{m}\int_{-\infty}^t \exp[-(t-t')/\tau]\xi(t')dt'\,. \qquad (6.4.5)$$

Here τ will be called *relaxation time* since it determines the time scale of relaxation to (6.4.5).
Let us define

$$\eta(t,\tau) = \tau^{-1}\int_{-\infty}^t \exp[-(t-t')/\tau]dW(t') \qquad (6.4.6)$$

which is, from Sect.4.4.4, a stationary Ornstein-Uhlenbeck process. The correlation function is

$$\langle\eta(t,\tau)\eta(t',\tau)\rangle = \frac{1}{2\tau}\exp(-|t-t'|/\tau) \qquad (6.4.7)$$

$$\xrightarrow[\tau\to0]{} \delta(t-t')\,. \qquad (6.4.8)$$

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)

$$\frac{dx}{dt} = v$$

$$m\frac{dv}{dt} = -\beta v - V'(x) + \sqrt{2\beta kT}\,\xi(t)\,. \tag{6.4.15}$$

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

$$(2\pi)^{-1/2} \exp\left(-\tfrac{1}{2}u^2\right) . \tag{6.4.24}$$

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

$$(Pf)(u, y) = (2\pi)^{-1/2} \exp\left(-\tfrac{1}{2}u^2\right) \int du \, f(u, y) \, . \tag{6.4.25}$$

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

$$g(u, y) = (2\pi)^{-1/2} \exp\left(-\tfrac{1}{2}u^2\right) \hat{g}(y) \, , \tag{6.4.27}$$

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

$$\lim_{t \to \infty} [\exp (L_1 \, t) \, f(u, y)] = \sum_\lambda A_\lambda(y) \lim_{t \to \infty} e^{-\lambda t} P_\lambda(u) \tag{6.4.32}$$

$$= P_0(u) \int du \, Q_0(u) f(u, y) \tag{6.4.33}$$

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

$$\boxed{PL_2 P = 0 ,} \tag{6.4.36}$$

For, considering $PL_2 Pf(u, y)$, we see from the definition of L_2, that $L_2 Pf(u, y)$ 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

$$\frac{\partial v}{\partial t} = P(\gamma L_1 + L_2)p$$

$$= P(\gamma L_1 + L_2)[Pp + (1 - P)p]$$

$$= PL_2(1 - P)p \, ,$$

[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,

$$\frac{\partial w}{\partial t} = (1 - P)(\gamma L_1 + L_2)p = (1 - P)(\gamma L_1 + L_2)[Pp + (1 - P)p]$$

$$= \gamma L_1(1 - P)p + (1 - P)L_2(1 - P)p + (1 - P)L_2Pp \tag{6.4.43}$$

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

$$\bar{f}(s) = \int\limits_0^\infty e^{-st} f(t) \tag{6.4.45}$$

and it may quite readily be defined for operators and abstract vectors in exactly the same way. Our fundamental equations become, on using

$$\int\limits_0^\infty e^{-st} \frac{df}{dt} = s\,\bar{f}(s) - f(0)\,, \tag{6.4.46}$$

$$\begin{aligned} s\,\bar{v}(s) &= PL_2\bar{w}(s) + v(0)\,, \\ s\,\bar{w}(s) &= [\gamma L_1 + (1 - P)L_2]\bar{w}(s) + L_2\bar{v}(s) + w(0)\,. \end{aligned} \tag{6.4.47}$$

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

$$p(u, y, 0) = (2\pi)^{-1/2} \exp\left(-\tfrac{1}{2}u^2\right) \bar{p}(y, 0)\,, \tag{6.4.49}$$

that is, the initial thermalisation of the velocity u is assumed. Then we have formally

$$\bar{w}(s) = [s - \gamma L_1 - (1 - P)L_2]^{-1} L_2 \bar{v}(s) \tag{6.4.50}$$

and hence,

$$\boxed{\; s\,\bar{v}(s) = PL_2[s - \gamma L_1 - (1 - P)L_2]^{-1} L_2 \bar{v}(s) + v(0)\,. \;} \tag{6.4.51}$$

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

$$s\,\bar{v}(s) \simeq -\gamma^{-1} PL_2 L_1^{-1} L_2 \bar{v}(s) + v(0)\,. \tag{6.4.52}$$

Notice that L_1^{-1} does not always exist. However, we know that

$$PL_2\bar{v}(s) = PL_2 P\,\bar{p}(s) = 0 \tag{6.4.53}$$

from (6.4.36). Hence, $L_2\bar{v}(s)$ contains no component in the null space of L_1 and thus $L_1^{-1} L_2 \bar{v}(s)$ exists.

In this case of Kramers' equation, let us now see what (6.4.52) looks like. It is equivalent to the differential 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

$$P_n(u) = (2\pi)^{-1/2} \exp\left(-\tfrac{1}{2}u^2\right) Q_n(u) \tag{6.4.57}$$

with

$$Q_n(u) = (2^n n!)^{-1/2} H_n(u/\sqrt{2}) \,. \tag{6.4.58}$$

Using

$$L_1 P_n(u) = -n P_n(u) \tag{6.4.59}$$

and the recursion formulae for Hermite polynomials

$$x H_n(x) = \tfrac{1}{2} H_{n+1}(x) + n H_{n-1}(x) \tag{6.4.60}$$

$$\frac{d}{dx}[e^{-x^2} H_n(x)] = -e^{-x^2} H_{n+1}(x) \,. \tag{6.4.61}$$

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

$$L_2 P_1(u) = -\left[\sqrt{2} \, P_2(u) + P_0(u) \right] \frac{\partial}{\partial y} - \sqrt{2} \, P_2(u) U'(y) \tag{6.4.64}$$

$$P L_2 L_1^{-1} L_2 v = -\frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y} \right] \hat{p}(y) P_0(u) \tag{6.4.65}$$

and the equation of motion (6.4.54) takes the form [after cancelling the factor $P_0(u)$]

$$\frac{\partial \bar{p}}{\partial t} = \gamma^{-1} \frac{\partial}{\partial y}\left[U'(y)\bar{p} + \frac{\partial \bar{p}}{\partial y}\right] \tag{6.4.66}$$

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

$$\frac{\partial v}{\partial t} = \int_0^t dt' \exp\left[\gamma(t' - t)\right]PL_2^2v(t')dt . \tag{6.4.75}$$

Alternatively, we can rewrite (6.4.74) in the form

$$\frac{1}{\gamma}[s^2\bar{v} - sv(0)] + [s\bar{v} - v(0)] = \gamma^{-1}PL_{\frac{1}{2}}^2 v \tag{6.4.76}$$

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

$$\int_0^\infty e^{-st}f''(t) = s^2\tilde{f}(s) - s f(0) - f'(0)$$

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 process.

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

for $u > 0$, $y = a$ is an exit boundary

for $u < 0$, $y = a$ is an entrance boundary.

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

$$
p(a, t) = 0
\tag{6.4.83}
$$

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

$$
u \rightarrow -u \, ,
$$

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

$$\boldsymbol{J} = \left\{-up, \; [\gamma u + U'(y)] \, p + \gamma \frac{\partial p}{\partial u}\right\} \tag{6.4.85}$$

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)

$$= - P_1(u)\gamma^{-1}\left[U'(y)\tilde{p}(y) + \frac{\partial \tilde{p}(y)}{\partial y}\right]_{y=a}. \tag{6.4.88}$$

Since this is proportional to the *odd* eigenfunction $P_1(u)$, it vanishes. Hence we derive

$$\left[U'(y)\tilde{p}(y) + \frac{\partial \tilde{p}(y)}{\partial y}\right]_{y=a} = 0 \tag{6.4.89}$$

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)

$$A = -PL_2L_1^{-1}L_2 = \frac{\partial}{\partial y}\left[U'(y) + \frac{\partial}{\partial y}\right]. \tag{6.4.102}$$

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

$$L_2P_2(u) = \left[-\frac{\partial}{\partial y}u + U'(y)\frac{\partial}{\partial u}\right]P_2(u) \tag{6.4.103}$$

$$= -\sqrt{3}\,U'(y)P_3(u) - \frac{\partial}{\partial y}\left[\sqrt{3}\,P_3(u) + \sqrt{2}\,P_1(u)\right].$$

Finally, multiplying by P annihilates all terms since $P_0(u)$ does not occur. Hence,

$$B = PL_2L_1^{-1}(1 - P)L_2L_1^{-1}L_2 = 0 . \tag{6.4.104}$$

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

$$\frac{\partial \hat{p}}{\partial t} = \gamma^{-1}\frac{\partial}{\partial y}[1 + \gamma^{-2}U''(y)]\left[U'(y)\hat{p} + \frac{\partial \hat{p}}{\partial y}\right] \tag{6.4.108}$$

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 $t \sim \gamma^{-1}$.
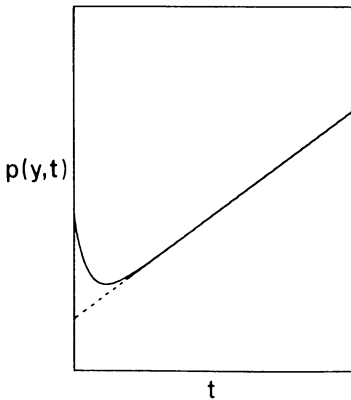
Fig. 6.1. Formation of a layer at a boundary. The exact solution (———) changes rapidly near the boundary on the left. The approximation (------) is good except near the boundary. The appropriate boundary condition for the approximation is thus given by the smaller value, where the dashed line meets the 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

$$\psi_0(u, z) = (2\pi)^{-1/2} \exp\left(-\tfrac{1}{2} u^2\right) \tag{6.4.112}$$

$$\psi_0'(u, z) = (2\pi)^{-1/2} (z - u) \exp\left(-\tfrac{1}{2} u^2\right) \tag{6.4.113}$$

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

$$L_1 = -\frac{\partial}{\partial \alpha}A(\alpha) + \frac{1}{2}\frac{\partial^2}{\partial \alpha^2}B(\alpha)$$

$$L_2 = -\frac{\partial}{\partial x}b(x)\,\alpha \tag{6.5.15}$$

$$L_3 = -\frac{\partial}{\partial x}a(x) .$$

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

$$PL_2 P = 0 . \tag{6.5.18}$$

Since

$$(PL_2 Pf)(x, \alpha) = p_s(\alpha) \int d\alpha \left[-\frac{\partial}{\partial x}b(x)\alpha p_s(\alpha) \right] \int d\alpha' f(x, \alpha')$$

$$= -p_s(\alpha)\langle \alpha \rangle_s \frac{\partial}{\partial x}b(x) \int d\alpha' f(x, \alpha') = 0 . \tag{6.5.19}$$

Also, it is obvious that

$$PL_3 = L_3 P \tag{6.5.20}$$

and, as before,

$$PL_1 = L_1 P = 0 . \tag{6.5.21}$$

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

$$s\,\tilde{v}(s) = P(\gamma^2 L_1 + \gamma L_2 + L_3)\bar{p}(s) + v(0) \tag{6.5.24}$$
$$= \gamma PL_2[P\bar{p}(s) + (1 - P)\bar{p}(s)] + L_3 P\,\bar{p}(s) + v(0)$$

so that

$$\boxed{s\,\tilde{v}(s) = \gamma PL_2\tilde{w}(s) + L_3\tilde{v}(s) + v(0)} \tag{6.5.25}$$

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

$$s\tilde{v}(s) = L_3\tilde{v}(s) - \gamma PL_2[-s + \gamma^2 L_1 + \gamma(1 - P)L_2 + L_3]^{-1}\gamma L_2\tilde{v}(s) + v(0). \tag{6.5.27}$$

Now the limit $\gamma \to \infty$ gives

$$\boxed{s\tilde{v}(s) \simeq (L_3 - PL_2 L_1^{-1}L_2)\tilde{v}(s) + v(0)\,.} \tag{6.5.28}$$

We now compute $PL_2 L_1^{-1}L_2\tilde{v}$. We write

$$\tilde{v}(s) = \bar{p}(x)p_s(\alpha) \tag{6.5.29}$$

$$PL_2 L_1^{-1}L_2\tilde{v} = p_s(\alpha) \int d\alpha' \left[-\frac{\partial}{\partial x}b(x)\alpha'\right]L_1^{-1}\left[-\frac{\partial}{\partial x}b(x)\alpha'\right]p_s(\alpha')\bar{p}(x)\,. \tag{6.5.30}$$

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

$$\frac{\partial \tilde{p}}{\partial t} = -\frac{\partial}{\partial x}[\{a(x) + E(x)\}\tilde{p}] + \frac{\partial^2}{\partial x^2}[D(x)\tilde{p}]. \tag{6.5.53}$$

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

$$\frac{dx}{dt} = a(x, t) + b(x, t)\alpha_0(t), \qquad \tag{6.5.54}$$

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

$$s\, \tilde{v}(s) = PL_3\tilde{v}(s) + P(\gamma L_2 + L_3)[s - \gamma^2 L_1 - \gamma(1 - P)L_2 - (1 - P)L_3]^{-1}$$
$$\times [\gamma L_2 + (1 - P)L_3]\tilde{v}(s) + v(0) . \tag{6.5.69}$$

We see that because L_2 is multiplied by γ and L_3 is not, we get in the limit of large γ

$$s\, \tilde{v}(s) \sim (PL_3 - PL_2 L_1^{-1} L_2)\tilde{v}(s) + v(0) \tag{6.5.70}$$

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)

$$dx = (-x + \gamma y)dt + \varepsilon B_{11}dW_1(t) + \varepsilon B_{12}dW_2(t) \tag{6.6.5}$$
$$dy = (-2\gamma y + x + a)dt + \varepsilon B_{21}dW_1(t) + \varepsilon B_{22}dW_2(t),$$

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

$$\frac{\partial p}{\partial t} = \left[\frac{\partial}{\partial x}(x - \gamma y) + \frac{\partial^2}{\partial x^2}\varepsilon^2 a - 2\frac{\partial^2}{\partial x \partial y}\varepsilon^2 a \right.$$
$$\left. + \frac{\partial}{\partial y}(2\gamma y - x - a) + 2\varepsilon^2 a \frac{\partial^2}{\partial y^2} \right] p \,. \tag{6.6.7}$$

It seems reasonable to define a new variable z by

$$z = 2\gamma y - x - a \qquad (6.6.8)$$

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

$$x_1 = x \qquad\qquad x = x_1 \qquad\qquad (6.6.9)$$
$$z = 2\gamma y - x - a \qquad y = (z + x_1 + a)/2\gamma$$

so that we can transform the FPE using

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial x_1} - \frac{\partial}{\partial z}$$
$$\qquad\qquad\qquad (6.6.10)$$
$$\frac{\partial}{\partial y} = 2\gamma \frac{\partial}{\partial z}$$

to obtain

$$\frac{\partial p}{\partial t} = \left[\frac{\partial}{\partial x_1}\left[\frac{x_1 - a}{2} - \frac{z}{2} \right] + \varepsilon^2 a \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_1 \partial z}(-2\varepsilon^2 a - 4\gamma\varepsilon^2 a) \right.$$
$$\left. + \frac{\partial}{\partial z}\left(2\gamma z - \frac{x-a}{2} + \frac{z}{z} \right) + \frac{\partial^2}{\partial z^2}(8\varepsilon^2\gamma^2 a + \varepsilon^2 a + 4\gamma\varepsilon^2 a) \right] p . \qquad (6.6.11)$$

The limit of $\gamma \rightarrow \infty$ 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

$$\alpha = z\gamma^{-1/2} \qquad (6.6.12)$$

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)

$$\frac{\partial p}{\partial t} = [\gamma L_1 + \gamma^{1/2} L_2(\gamma) + L_3] p \qquad (6.6.13)$$

in which

$$L_1 = 2 \frac{\partial}{\partial \alpha}\alpha + 8\varepsilon^2 a \frac{\partial^2}{\partial \alpha^2} \qquad (6.6.14)$$

$$L_2(\gamma) = \frac{\partial}{\partial x}\left[-4\varepsilon^2 a \frac{\partial}{\partial \alpha} - \tfrac{1}{2}\alpha \right] + 4\gamma^{-1/2} \varepsilon^2 a \frac{\partial^2}{\partial \alpha^2} - \gamma^{-1} \frac{\partial}{\partial \alpha}\left[\frac{x-a}{2} \right]$$
$$+ \left\{ -\gamma^{-1/2} \frac{\partial}{\partial \alpha}\alpha - \gamma^{-1} 2\varepsilon^2 a \frac{\partial^2}{\partial x \partial \alpha} + \gamma^{-3/2}\varepsilon^2 a \frac{\partial^2}{\partial \alpha^2} \right\} \qquad (6.6.15)$$

$$L_3 = \frac{\partial}{\partial x}\left(\frac{x-a}{2} \right) + \varepsilon^2 a \frac{\partial^2}{\partial x^2} . \qquad (6.6.16)$$

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

$$P\tilde{p}(s) = \tilde{v}(s) \tag{6.6.20}$$
$$(1 - P)\tilde{p}(s) = \tilde{w}(s)$$

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

$$PL_1 = L_1P = 0$$
$$PL_2P = 0 \tag{6.6.22}$$
$$PL_3 = L_3P ,$$

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

$$s\,\tilde{v}(s) = \{L_3 + \gamma PL_2(\gamma)[s - \gamma L_1 - \gamma^{1/2}(1 - P)L_2(\gamma) - L_3]^{-1}L_2\}\,\tilde{v}(s)$$
$$+ v(0)\,. \tag{6.6.25}$$

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 $L_1 p_s(\alpha) = 0$, we see from (6.6.14) that

$$\frac{\partial}{\partial\alpha}p_s(\alpha) = -\alpha p_s(\alpha)/4\varepsilon^2 a \qquad \text{✦} \tag{6.6.29}$$

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),

$$-PL_2L_1^{-1}L_2v = -\frac{1}{4}p_s(\alpha)\frac{\partial^2\hat{p}(x)}{\partial x^2}4\varepsilon^2 a\int_0^\infty dt\,e^{-2t} \tag{6.6.32}$$

$$= -\frac{1}{2}\varepsilon^2 a\frac{\partial^2\hat{p}(x)}{\partial x^2}p_s(\alpha)\,.$$

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

$$\text{var}\{x(t)\}_s = \varepsilon^2 \langle x \rangle_s . \tag{6.6.35}$$

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

$$\dot{x} = -\varepsilon x - ax\alpha \tag{6.6.36}$$

$$\dot{\alpha} = -\kappa\alpha + bx^2 . \tag{6.6.37}$$

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):

$$dx = -(\varepsilon x + ax\alpha)dt + C\,dW_1(t)$$
$$da = (-\kappa\alpha + bx^2)dt + D\,dW_2(t)$$

(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

$$\frac{\partial p}{\partial t} = \left[\frac{\partial}{\partial x}(\varepsilon x + ax\alpha) + \tfrac{1}{2}C^2\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial \alpha}(\kappa\alpha - bx^2) + \tfrac{1}{2}D^2\frac{\partial^2}{\partial \alpha^2}\right]p\,.$$

(6.6.42)

We wish to eliminate α. It is convenient to define a new variable β by

$$\beta = \alpha - \frac{b}{\kappa}x^2$$

(6.6.43)

so that, for fixed x, the quantity β has zero mean. In terms of this variable, we can write a FPE:

$$\frac{\partial p}{\partial t} = (L_1^0 + L_2^0 + L_3^0)p$$

(6.6.44)

$$L_1^0 = \frac{\partial}{\partial\beta}\kappa\beta + \frac{D^2}{2}\frac{\partial^2}{\partial\beta^2}$$

(6.6.45)

$$L_2^0 = \frac{\partial}{\partial\beta}a\beta x - \frac{2bx}{\kappa}\frac{\partial}{\partial\beta}\left(\varepsilon x + \frac{ab}{\kappa}x^3 + ax\beta\right)$$
$$- C^2\left(\frac{bx}{\kappa}\frac{\partial^2}{\partial x\partial\beta} + \frac{\partial^2}{\partial x\partial\beta}\frac{bx}{\kappa}\right) + \frac{2b^2x^2C^2}{\kappa^2}\frac{\partial^2}{\partial\beta^2}$$

(6.6.46)

$$L_3^0 = \left[\frac{\partial}{\partial x}\left(\varepsilon x + \frac{ab}{\kappa}x^3\right) + \frac{C^2}{2}\frac{\partial^2}{\partial x^2}\right]$$

(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

$$\frac{ab}{\kappa} = \varepsilon A, \quad \text{as} \quad \varepsilon \to 0\,.$$

(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

$$\boxed{\frac{\partial \bar{p}}{\partial t} = L_3 \bar{p} = \left[\frac{\partial}{\partial x}(x + Ax^3) + B \frac{\partial^2}{\partial x^2} \right] \bar{p}} \tag{6.6.55}$$

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

$$a = \tilde{a}\varepsilon^{1/2} \tag{6.6.56}$$
$$b = \tilde{b}\varepsilon^{1/2} ,$$

where

$$\tilde{a}\tilde{b} = \kappa A . \tag{6.6.57}$$

L_1^0 stays constant, L_3^0 is proportional to ε and

$$L_2^0 = \varepsilon^{1/2}L_2 + \text{higher order terms in } \varepsilon , \tag{6.6.58}$$

where

$$L_2 = \tilde{a}\beta\frac{\partial}{\partial x} x . \tag{6.6.59}$$

Thus, the limiting equation is

$$\frac{\partial \hat{p}}{\partial \tau} = (L_3 - PL_2L_1^{-1}L_2)\hat{p} . \tag{6.6.60}$$

The term $PL_2L_1^{-1}L_2$ can be worked out as previously; we find

$$- PL_2L_1^{-1}L_2 = \tilde{a}^2 \frac{D^2}{2\kappa^2} \frac{\partial}{\partial x} x \frac{\partial}{\partial x} x \tag{6.6.61}$$

so

$$\frac{\partial \hat{p}}{\partial \tau} = \left\{ \frac{\partial}{\partial x}\left[x\left[1 - \frac{\tilde{a}^2 D^2}{2\kappa^2} \right] + Ax^3 \right] + \frac{\partial^2}{\partial x^2}\left[B + \frac{\tilde{a}^2 D^2 x^2}{2\kappa^2} \right] \right\} \hat{p} . \tag{6.6.62}$$

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 $r < \frac{1}{2}$)

$$\frac{\partial \tilde{p}}{\partial \tau} = \varepsilon^{2r-1}\tilde{a}^2 \frac{D^2}{2\kappa^2} \left(\frac{\partial}{\partial x} x \frac{\partial}{\partial x} x\right) \tilde{p} \ . \tag{6.6.63}$$

These are quite distinct differences, all of which can be incorporated in the one formula, namely, in general

$$\frac{\partial \tilde{p}}{\partial \tau} = \left[\frac{\partial}{\partial x}(x + Ax^3) + \frac{\partial^2}{\partial x^2} B + \varepsilon^{2r-1}\tilde{a}^2 \frac{D^2}{2\kappa^2} \frac{\partial}{\partial k} x \frac{\partial}{\partial x} x\right] \tilde{p} \ . \tag{6.6.64}$$

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

$$dx = [a(x) + b(x)u_0(x)]dt + c(x)dW_1(t).$$ (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

$$\frac{\partial p}{\partial t} = (\gamma^2 L_1 + L_2 + L_3)p,$$ (6.6.70)

where

$$L_1 = \frac{\partial}{\partial \alpha}[f(x) - A(\alpha)] + \frac{\partial^2}{\partial \alpha^2}B(\alpha)$$ (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

$$L_1 p_x(\alpha) = 0$$ (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

$$(PF)(x, \alpha) = p_x(\alpha) \int d\alpha' F(x, \alpha')$$ (6.6.73)

for any function $F(x,\alpha)$.

We now define the function $u(x)$ as

$$u(x) = \int d\alpha \, \alpha p_x(\alpha) \equiv \langle \alpha \rangle_x.$$ (6.6.74)

Then we define

$$L_2 = -\frac{\partial}{\partial x}\left\{b(x)[\alpha - u(x)]\right\}$$ (6.6.75)

$$L_3 = -\frac{\partial}{\partial x}[a(x) + b(x)u(x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)]^2$$ (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

$$PL_2PF = -p_x(\alpha) \int d\alpha' \frac{\partial}{\partial x}\{b(x)[\alpha' - u(x)]\}p_x(\alpha') \int d\alpha'' F(x, \alpha'')$$ (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,

$$P\,\tilde{p}(s) = \tilde{v}(s)$$
$$(1 - P)\tilde{p}(s) = \tilde{w}(s) \tag{6.6.80}$$

and assuming $w(0) = 0$, we find

$$s\,\tilde{v}(s) = P(L_2 + L_3)\tilde{w}(s) + PL_3\tilde{v}(s) + v(0)$$
$$s\,\tilde{w}(s) = [\gamma^2 L_1 + (1 - P)L_2 + (1 - P)L_3]\tilde{w}(s) + L_2\tilde{v}(s) + (1 - P)L_3\tilde{v}(s) \tag{6.6.81}$$

so that

$$s\,\tilde{v}(s) = PL_3\tilde{v}(s) + P(L_2 + L_3)[s - \gamma^2 L_1 - (1 - P)L_2 - (1 - P)L_3]^{-1}$$
$$\times [L_2 + (1 - P)L_3]\tilde{v}(s) + v(0) . \tag{6.6.82}$$

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

$$v(t) = p_x(\alpha)\tilde{p}(x) ,$$

we find

$$PL_3v(t) = p_x(\alpha) \int d\alpha' \left\{ -\frac{\partial}{\partial x}[a(x) + b(x)u(x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)^2] \right\} p_x(\alpha)\tilde{p}(x) \tag{6.6.84}$$

and since

$$\int d\alpha\, p_x(\alpha) = 1 ,$$

$$PL_3v(t) = p_x(\alpha) \left\{ -\frac{\partial}{\partial x}[a(x) + b(x)u(x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)^2] \right\} \tilde{p}(x) \tag{6.6.85}$$

so the lowest-order differential equation is

$$\frac{\partial\tilde{p}(x)}{\partial t} = -\frac{\partial}{\partial x}[a(x) + b(x)u(x)]\,\tilde{p}(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}c(x)^2\tilde{p}(x) \tag{6.6.86}$$

which is equivalent to the stochastic differential equation

$$dx = [a(x) + b(x)u(x)]dt + c(x)dW(t) .$$
(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

$$P(L_2 + L_3)L_1^{-1}[L_2 + (1 - P)L_3]\bar{v}(s) .$$
(6.6.88)

We know

i) $P\bar{v}(s) = \bar{v}(s)$

ii) $(1 - P)L_2 P\bar{v}(s) = L_2 P\bar{v}(s) ,$
(6.6.89)

where we have used $PL_2 P = 0$.

Thus, (6.6.88) becomes

$$P(L_2 + L_3)L_1^{-1}(1 - P)(L_2 + L_3)P\bar{v}(s) = P\{PL_2 + [P, L_3] + L_3 P\}$$
$$\times (1 - P)L_1^{-1}(1 - P)\{L_2 P + [L_3, P] + PL_3\} \bar{v}(s)$$
(6.6.90)

where the *commutator* $[A, B]$ is defined by

$$[A, B] = AB - BA .$$
(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 $P\{PL_2 + [P, L_3]\} L_1^{-1} (1 - P)\{L_2 + [L_3, P]\} \bar{v}(s) .$
(6.6.92)

We will now compute $[P, L_3]$:

$$(PL_3 f)(x, \alpha) = p_x(\alpha) \left\{ -\frac{\partial}{\partial x}[a(x) + b(x)u(x)] \right.$$
$$\left. + \frac{1}{2}\frac{\partial^2}{\partial x^2}[c(x)]^2 \right\} \int f(x, \alpha') \, d\alpha'$$
(6.6.93)

and

$$(L_3 P) f(x, \alpha) = \left\{ -\frac{\partial}{\partial x}]a(x) + b(x)u(x)] \right.$$

$$\left. + \frac{1}{2} \frac{\partial^2}{\partial x^2} [a(x)]^2 \right\} p_x(\alpha) \int d\alpha' f(x, \alpha') . \tag{6.6.94}$$

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

$$([P, L_3] f) (x, \alpha) = r_x(\alpha)[a(x) + b(x)u(x)] P f(x, \alpha)$$

$$- \tfrac{1}{2} s_x(\alpha) c(x)^2 P f(x, \alpha) - r_x(\alpha) P \frac{\partial}{\partial x} [c(x)^2 f(x, \alpha)] . \tag{6.6.97}$$

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

$$PL_2L_1^{-1}\{L_2 + [L_3, P]\}\,\bar{v}(s) \tag{6.6.105}$$

and

$$[L_3, P]\bar{v}(s) = -p_x(\alpha)\{r_x(\alpha)[a(x) + b(x)u(x)] - \tfrac{1}{2}s_x(\alpha)c(x)^2\}\,\hat{p}(x)$$

$$+ p_x(\alpha)r_x(\alpha)\frac{\partial}{\partial x}[c(x)^2\hat{p}(x)]\,. \tag{6.6.106}$$

6.6.4 An Example with Arbitrary Nonlinear Coupling

We consider the pair of equations

$$dx = \gamma b(x)\alpha dt \tag{6.6.107}$$

$$d\alpha = -\gamma^2 A(x, \alpha, \gamma)dt + \gamma\sqrt{2B(x, \alpha, \gamma)}\,dW(t)$$

and assume the existence of the following limits and asymptotic expansions

$$A(x, \alpha, \gamma) \sim \sum_{n=0}^{\infty} A_n(x, \alpha)\gamma^{-n}$$

$$B(x, \alpha, \gamma) \sim \sum_{n=0}^{\infty} B_n(x, \alpha)\gamma^{-n} \tag{6.6.108}$$

These expansions imply that there is an asymptotic stationary distribution of α at fixed x given by

$$p_s(\alpha, x) = \lim_{\gamma\to\infty} p_s(\alpha, x, \gamma) \tag{6.6.109}$$

$$p_s(\alpha, x) \propto B_0(x, \alpha)^{-1} \exp\{\textstyle\int d\alpha[A_0(x, \alpha)/B_0(x, \alpha)]\}\,. \tag{6.6.110}$$

We assume that $A_0(x, \alpha)$ and $B_0(x, \alpha)$ are such that

$$\langle\alpha(x)\rangle_s = \textstyle\int d\alpha\,\alpha p_s(\alpha, x) = 0 \tag{6.6.111}$$

so that we deduce from (6.6.108) that, for finite γ

$$\langle\alpha(x, \gamma)\rangle_s = \textstyle\int d\alpha\,\alpha p_s(\alpha, x, \gamma) \sim \alpha_0(x)\gamma^{-1} \tag{6.6.112}$$

where $\alpha_0(x)$ can be determined from (6.6.108).

We define the new variables

$$\beta = \alpha - \frac{1}{\gamma}\,\alpha_0(x) \tag{6.6.113}$$

$$x_1 = x\,.$$

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

$$PL_2L_1^{-1}L_2\hat{p} = \left[\frac{\partial}{\partial x}b(x)D(x)\frac{\partial}{\partial x}b(x) + \frac{\partial}{\partial x}b(x)E(x)\right]\hat{p} \qquad (6.6.122)$$

with

$$D(x) = \int_0^\infty dt \langle \beta(t), \beta(0)|x\rangle \qquad (6.6.123)$$

$$E(x) = \int_0^\infty dt \langle \beta(t), G(\beta, x)|x\rangle$$

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

$$\partial_t P(x, t \mid x', t') = t^+(x - 1)P(x - 1, t \mid x', t') + t^-(x + 1)P(x + 1, t \mid x', t')$$
$$- [t^+(x) + t^-(x)]P(x, t \mid x', t') \, . \tag{7.1.4}$$

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

(i) $t^-(0) = 0$:   no probability of an individual dying if there are none present;   (7.1.7)

(ii) $P(x, t \mid x', t') = 0$ for $x < 0$ or $x' < 0$.   (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

$$0 = \sum_{z=0}^{x-1} [J(z+1) - J(z)] = J(x) - J(0).$$ (7.1.10)

Hence,

$$J(x) = 0$$ (7.1.11)

and thus

$$P_s(x) = \frac{t^+(x-1)}{t^-(x)} P_s(x-1)$$ (7.1.12)

so that

$$P_s(x) = P_s(0) \prod_{z=1}^{x} \frac{t^+(z-1)}{t^-(z)}.$$ (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

$$P(x, \tau | x', 0) P_s(x') = P(x', \tau | x, 0) P_s(x),$$ (7.1.14)

Setting $x' = x \pm 1$ and taking the limit $\tau \to 0$, and noting that by definition (3.4.1),

$$W(x | x', t) = \lim_{\tau \to 0} P(x, t + \tau | x', t)/\tau,$$ (7.1.15)

the necessity of this condition is easily proved.

b) Rate Equations

We notice that the mean of x satisfies

$$\partial_t \langle x(t) \rangle = \partial_t \sum_{x=0}^{\infty} x P(x, t | x', t')$$ (7.1.16)

$$= \sum_{x=0}^{\infty} x[t^+(x-1)P(x-1, t | x', t') - t^+(x)P(x, t | x', t')]$$

$$+ \sum_{x=0}^{\infty} x[t^-(x+1)P(x+1, t | x', t') - t^-(x)P(x, t | x', t')]$$ (7.1.17)

$$= \sum_{x=0}^{\infty} [(x+1)t^+(x) - xt^+(x) + (x-1)t^-(x)$$

$$- xt^-(x)]P(x, t | x', t'),$$ (7.1.18)

i.e.,

$$\frac{d}{dt} \langle x(t) \rangle = \langle t^+[x(t)] \rangle - \langle t^-[x(t)] \rangle.$$ (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 $P(x, t | x', t')$ 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

$$\phi(s, t) = G(s, t) \exp(-k_2 as/k_1) , \tag{7.1.29}$$

(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 $t' = 0$)

$$P(x, 0 | N, 0) = \delta_{x, N} \tag{7.1.34}$$

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

$$\boxed{\begin{aligned}
P(x, t | N, 0) = \exp&\left[-\frac{k_2 a}{k_1}(1 - e^{-k_1 t})\right] \sum_{r=0}^{x} \frac{N!}{(N-r)! r! (x-r)!}\left(\frac{k_2 a}{k_1}\right)^{x-r} \\
&\times (1 - e^{-k_1 t})^{N+x-2r} e^{-k_1 tr} .
\end{aligned}} \tag{7.1.37}$$

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

$$\langle x(t) \rangle \qquad = \partial_s G(s = 1, t) = \frac{k_2 a}{k_1} (1 - e^{-k_1 t}) + N e^{-k_1 t} \qquad (7.1.38)$$

$$\langle x(t)[x(t) - 1] \rangle = \partial_s^2 G(s = 1, t) = \langle x(t) \rangle^2 - N e^{-2k_1 t} \qquad (7.1.39)$$

$$\mathrm{var}\{x(t)\} \qquad = \left(N e^{-k_1 t} + \frac{k_2 a}{k_1} \right)(1 - e^{-k_1 t}) . \qquad (7.1.40)$$

c) Moment Equations

From the differential equation (7.1.28) we have

$$\partial_t [\partial_s^n G(s, t)] = \{n[k_2 a \partial_s^{n-1} - k_1 \partial_s^n] + (s - 1)[k_2 a \partial_s^n - k_1 \partial_s^{n+1}]\} G(s, t) . \qquad (7.1.41)$$

Setting $s = 1$ and using

$$\partial_s^n G(s, t)|_{s=1} = \langle x(t)^n \rangle_f , \qquad (7.1.42)$$

we find

$$\frac{d}{dt} \langle x(t)^n \rangle_f = n[k_2 a \langle x(t)^{n-1} \rangle_f - k_1 \langle x(t)^n \rangle_f] \qquad (7.1.43)$$

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)

$$G(s, t \to \infty) = \exp[(s - 1)k_2 a/k_1] \qquad (7.1.44)$$

corresponding to the *Poissonian solution:*

$$P_s(x) = \exp(-k_2 a/k_1) (k_2 a/k_1)^x / x! . \qquad (7.1.45)$$

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,

$$\langle x(t) \rangle_s \qquad = k_2 a/k_1 \qquad (7.1.46)$$

$$\mathrm{var}\{x(t)\}_s \qquad = k_2 a/k_1 \qquad (7.1.47)$$

$$\langle x(t), x(0) \rangle_s = e^{-k_1 t} k_2 a/k_1 . \qquad (7.1.48)$$

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

$$P(x, 0) = \frac{e^{-\alpha_0}\alpha_0^x}{x!},$$ (7.1.49)

then

$$G(s, 0) = \exp\left[(s - 1)\alpha_0\right]$$ (7.1.50)

and from (7.1.32) we find

$$G(s, t) = \exp\{(s - 1)[\alpha_0 e^{-k_1 t} + (k_2 a/k_1)(1 - e^{-k_1 t})]\}$$ (7.1.51)

corresponding to

$$P(x, t) = \frac{e^{-\alpha(t)}\alpha(t)^x}{x!}$$ (7.1.52)

with

$$\alpha(t) = \alpha_0 e^{-k_1 t} + (k_2 a/k_1)(1 - e^{-k_1 t}).$$ (7.1.53)

Here $\alpha(t)$ is seen to be the solution of the deterministic equation

$$\frac{dx}{dt} = k_2 a - k_1 x$$ (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

$$A + 2X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3X$$ (7.1.56)

$$A \underset{k_4}{\overset{k_3}{\rightleftharpoons}} X$$ (7.1.57)
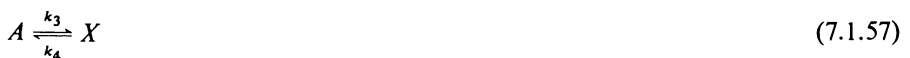
which has been studied by many authors [7.1]. The concentration of A is held fixed so that we have

$$t^+(x) = k_1 Ax(x - 1) + k_3 A$$
$$t^-(x) = k_2 x(x - 1)(x - 2) + k_4 x .$$

(7.1.58)

The corresponding deterministic equation is, of course,

$$\frac{dx}{dt} = t^+(x) - t^-(x)$$
$$\simeq -k_2 x^3 + k_1 Ax^2 - k_4 x + k_3 A ,$$

(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

$$\left(\frac{x - x_1}{x_0 - x_1}\right)^{x_3 - x_2} \left(\frac{x - x_2}{x_0 - x_2}\right)^{x_1 - x_3} \left(\frac{x - x_3}{x_0 - x_3}\right)^{x_2 - x_1}$$
$$= \exp\left[-k_2(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)t\right] .$$

(7.1.60)

Here, x_1, x_2, x_3 are roots of

$$k_2 x^3 - k_1 Ax^2 + k_4 x - x_3 A = 0$$

(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

$$x < x_1 \qquad \Longrightarrow \frac{dx}{dt} > 0$$

$$x_2 > x > x_1 \Longrightarrow \frac{dx}{dt} < 0$$

(7.1.62)

$$x_3 > x > x_2 \Longrightarrow \frac{dx}{dt} > 0$$

$$x > x_3 \qquad \Longrightarrow \frac{dx}{dt} < 0 .$$

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)

$$P_s(x) = P_s(0) \prod_{z=1}^{x} \left\{\frac{B[(z - 1)(z - 2) + P]}{z(z - 1)(z - 2) + Rz}\right\} ,$$

(7.1.63)

where

$$B = k_1 A/k_2$$
$$R = k_4/k_2 \qquad\qquad\qquad (7.1.64)$$
$$P = k_3/k_1 \ .$$

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:

$$x(B) \sim B \qquad \text{large } B$$
$$\qquad\qquad\qquad\qquad\qquad (7.1.66)$$
$$x(B) \sim PB/R \quad \text{small } B$$

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.
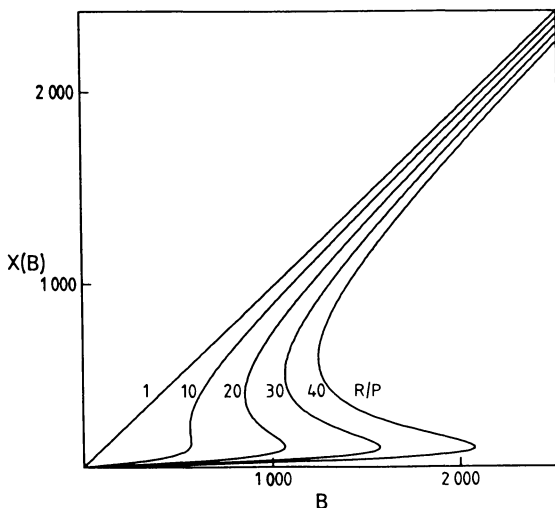
Fig. 7.1. Plot of $x(B)$ against B, as given by the solution of (7.1.65) for various values of R/P, and $P = 10,000$

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,

$$k_1 A \sim 1/V$$
$$k_3 A \sim V \tag{7.1.67}$$
$$k_2 \sim 1/V^2$$
$$k_4 \sim 1$$

which means that

$$B \sim V$$
$$R \sim V^2 \tag{7.1.68}$$
$$P \sim V^2.$$

We then write

$$B = \tilde{B}V$$
$$R = \tilde{R}V^2$$
$$P = \tilde{P}V^2$$

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,

$$\log\left[P_s(y_2)/P_s(y_1)\right] \simeq \sum_{z=y_1 V}^{y_2 V} \{\log \tilde{B}V + \log(z^2 + \tilde{P}V^2)$$
$$- \log[z(z^2 + \tilde{R}V^2)]\} \tag{7.1.69}$$

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,

$$\langle x^k \rangle = \left[\sum_{x=0}^{\infty} x^k B^x G(x) \right] \left[\sum_{x=0}^{\infty} B^x G(x) \right]^{-1} \left.\begin{array}{c} \\ \\ \\ \\ \\ \end{array}\right\}$$

and

$$B \frac{\partial}{\partial B} \langle x^k \rangle = \langle x^{k+1} \rangle - \langle x \rangle \langle x^k \rangle \qquad \qquad (7.1.72)$$

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),

$$\langle x(B) \rangle = \mathrm{var}\{x(B)\} = B \qquad \text{large } B$$
$$\langle x(B) \rangle = \mathrm{var}\{x(B)\} = PB/R \quad \text{small } B$$

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 proces⁻ 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

$$W_\delta(x'|x) = \Phi\left(\frac{x' - x - A(x)\delta}{\sqrt{\delta}}, x\right)\delta^{-3/2}, \tag{7.2.1}$$

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

$$\begin{aligned}
\alpha_0(x) &\equiv \int dx' W_\delta(x'|x) = Q/\delta \\
\alpha_1(x) &\equiv \int dx'(x' - x)W_\delta(x'|x) = A(x)Q \\
\alpha_2(x) &\equiv \int dx'(x' - x)^2 W_\delta(x'|x) = \int dy\, y^2 \Phi(y, x).
\end{aligned} \tag{7.2.4}$$

We further assume that $\Phi(y, x)$ vanishes sufficiently rapidly as $y \to \infty$, so that

$$\lim_{\delta \to 0} W_\delta(x'|x) = \lim_{y \to \infty}\left[\left(\frac{y}{x' - x}\right)^3 \Phi(y, x)\right] = 0 \qquad \text{for} \quad x' \neq x. \tag{7.2.5}$$

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

$$\lim_{\delta \to 0} \left\langle \frac{\partial f(z)}{\partial t} \right\rangle = \left\langle \alpha_1(z) \frac{\partial f}{\partial z} + \frac{1}{2} \alpha_2(z) \frac{\partial^2 f}{\partial z^2} \right\rangle \qquad (7.2.6)$$

implying that in the limit $\delta \to 0$, the Master equation

$$\frac{\partial P(x)}{\partial t} = \int dx' [W(x|x')P(x') - W(x'|x)P(x)] \qquad (7.2.7)$$

becomes the FPE

$$\frac{\partial P(x)}{\partial t} = -\frac{\partial}{\partial x} \alpha_1(x)P(x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \alpha_2(x)P(x) . \qquad (7.2.8)$$

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

$$W(x|x') = d(\delta_{x',x-l} + \delta_{x',x+l}) . \qquad (7.2.9)$$

Then

$$\alpha_0(x) = 2d$$
$$\alpha_1(x) = 0 \qquad (7.2.10)$$
$$\alpha_2(x) = 2l^2 d ;$$

let

$$\delta = l^2 \qquad (7.2.11)$$

and

$$D = l^2 d . \qquad (7.2.12)$$

Then all requirements are met, so the limiting equation is

$$\frac{\partial P}{\partial t} = D \frac{\partial^2 P}{\partial x^2} \qquad (7.2.13)$$

as found in Sect.3.8.2.

b) Poisson Process (Sect.3.8.3)
Here, letting $x = nl$,

$$W(x|x') = d\delta_{x,x'+l} \qquad (7.2.14)$$

and

$$\alpha_0(x) = d$$
$$\alpha_1(x) = ld \qquad\qquad (7.2.15)$$
$$\alpha_2(x) = l^2 d.$$

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

$$W_\delta(x'|x) = \left(\frac{A(x)}{2\delta} + \frac{B(x)}{2\delta^2}\right)\delta_{x',x+\delta} + \left(-\frac{A(x)}{2\delta} + \frac{B(x)}{2\delta^2}\right)\delta_{x',x-\delta} \qquad (7.2.16)$$

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

$$\alpha_0(x) = B(x)/\delta^2 \qquad\qquad (7.2.17a)$$

$$\alpha_1(x) = A(x) \qquad\qquad (7.2.17b)$$

$$\alpha_2(x) = B(x) \qquad\qquad (7.2.17c)$$

and

$$\lim_{\delta \to 0} W_\delta(x'|x) = 0 \qquad \text{for} \qquad x' \neq x. \qquad (7.2.17d)$$

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

$$\frac{\partial P(x)}{\partial t} = -\frac{\partial}{\partial x} A(x)P(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2} B(x)P(x). \qquad (7.2.18)$$

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

$$P_s(x) = P_s(0)\prod_{z=\delta}^{x}\left[\frac{\delta A(z-\delta) + B(z-\delta)}{-\delta A(z) + B(z)}\right]$$

$$= P_s(0)\left[\frac{-\delta A(0) + B(0)}{\delta A(x) + B(x)}\right]\prod_{z=0}^{x}\left[\frac{1 + \delta A(z)/B(z)}{1 - \delta A(z)/B(z)}\right] \qquad (7.2.19)$$

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

$$y = x - x' \quad \text{in the first term, and}$$
$$y = x' - x \quad \text{in the second term.}$$

Defining

$$t(y, x) = W(x + y \,|\, x) \,, \tag{7.2.22}$$

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

$$\alpha_n(x) = \int dx'(x' - x)^n W(x' \,|\, x) = \int dy \, y^n \, t(y, x) \,. \tag{7.2.26}$$

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

$$\boxed{\frac{\partial P(z, t)}{\partial t} = - \tilde{\alpha}_1^1[\phi(t)] \frac{\partial}{\partial z} z P(z, t) + \frac{1}{2} \tilde{\alpha}_2[\phi(t)] \frac{\partial^2}{\partial z^2} P(z, t) .} \tag{7.2.35}$$

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

$$W(x|x') = \delta_{x,x'+1}k_2 a + \delta_{x,x'-1}k_1 x' . \tag{7.2.41}$$

The assumption is

$$\begin{aligned} a &= a_0 V \\ x &= x_0 V , \end{aligned} \tag{7.2.42}$$

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,

$$W(x_0'; \Delta x) = V[k_2 a_0 \delta_{\Delta x,1} + k_1 x_0' \delta_{\Delta x,-1}] \tag{7.2.43}$$

which is in the form of (7.2.29), with $\Omega \to V$, $a \to x$, etc.
Thus,

$$\begin{aligned} \alpha_1(x) &= \sum (x' - x)W(x'|x) = k_2 a - k_1 x = V(k_2 a_0 - k_1 x_0) \\ \alpha_2(x) &= \sum (x' - x)^2 W(x'|x) = k_2 a + k_1 x = V(k_2 a_0 + k_1 x_0) . \end{aligned} \tag{7.2.44}$$

The deterministic equation is

$$\phi'(t) = [k_2 a_0 - k_1 \phi(t)] \tag{7.2.45}$$

whose solutions is

$$\phi(t) = \phi(0)e^{-k_1 t} + \frac{k_2 a_0}{k_1}(1 - e^{-k_1 t}) . \tag{7.2.46}$$

The Fokker-Planck equation is

$$\frac{\partial P(z)}{\partial t} = k_1 \frac{\partial}{\partial z} z P(z) + \frac{1}{2} \frac{\partial^2}{\partial z^2} [k_2 a_0 + k_1 \phi(t)] P(z) . \tag{7.2.47}$$

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

$$\mathrm{var}\,\{z(t)\} = \left[\frac{k_2 a_0}{k_1} + \phi(0) e^{-k_1 t}\right] (1 - e^{-k_1 t}) \tag{7.2.49}$$

so that

$$\langle x(t) \rangle = V \phi(t) = V \phi(0) e^{-k_1 t} + \frac{k_2 a}{k_1} (1 - e^{-k_1 t}) \tag{7.2.50}$$

$$\mathrm{var}\,\{x(t)\} = V\,\mathrm{var}\,\{z(t)\} = \left[\frac{k_2 a}{k_1} + V\phi(0) e^{-k_1 t}\right](1 - e^{-k_1 t}) . \tag{7.2.51}$$

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

$$\begin{aligned} P_s(x) &= \frac{\mathcal{N}}{\alpha_2(x)} \exp\left[\int_0^x 2 \frac{\alpha_1(x')}{\alpha_2(x')}\,dx'\right] \\ &= \mathcal{N}(k_2 a + k_1 x)^{-1 + 4 k_2 a/k_1} e^{-2x} . \end{aligned} \tag{7.2.53}$$

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

$$(7.2.53) \simeq \mathcal{N}(2 V k_2 a_0 + k_1 \delta)^{-1 + 4 V k_2 a_0/k_1} e^{-2 V k_2 a_0/k_1 - 2\delta} . \tag{7.2.55}$$

Then,

$$\log P_s(x) \simeq \mathrm{const} - \frac{k_2}{2 k_2 a_0 V} (\delta - \delta^2) . \tag{7.2.56}$$

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)

$$\tilde{\alpha}_1(y) = (\tilde{B}y^2 + \bar{p} - y^3 - y\tilde{R})\tilde{k}_2\,,$$

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
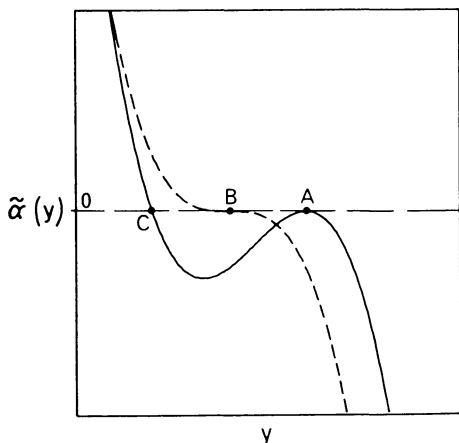
Fig. 7.2. Graph showing different kinds of behaviour of $\alpha_1(y)$ which lead to $\alpha_1'(y) = 0$

$$\dot{y} = -\tilde{k}_2(y - \phi_s)^3 \tag{7.2.69}$$

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	$t^-(a) = 0$	$t^+(a - 1) = 0$	(7.3.1)
Boundary at b	$t^+(b) = 0$	$t^-(b + 1) = 0.$	

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

$$\partial_t P(x, t | x', t') = t^+(x - 1)P(x - 1, t | x', t') + t^-(x + 1)P(x + 1, t | x', t')$$
$$- [t^+(x) + t^-(x)]P(x, t | x', t') . \tag{7.3.2}$$

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)

$$\partial_{t'} P(x, t | x', t') = t^+(x')[P(x, t | x' + 1, t') - P(x, t | x', t')]$$
$$+ t^-(x')[P(x, t | x' - 1, t') - P(x, t | x', t')] . \tag{7.3.8}$$

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	$P(\cdot \| b + 1) = P(\cdot \| b)$	$P(\cdot \| b + 1) = 0$

$$\tag{7.3.12}$$

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 absorbed or reflected at either end, as the case may be. For definiteness we take a system with

reflecting barrier at $\quad x = a$;

absorbing state at $\quad x = b + 1$.

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

$$\phi(x) = \prod_{z=a+1}^{x} \frac{t^-(z)}{t^+(z)} \quad \text{and} \tag{7.4.6}$$

$$S(x) = U(x)/\phi(x) \tag{7.4.7}$$

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

$$\boxed{T(x) = \sum_{y=x}^{b} \phi(y) \sum_{z=a}^{y} 1/[t^+(z)\phi(z)]} \quad \begin{matrix} a & \text{reflecting} \\ b & \text{absorbing} \\ b > a \end{matrix} \tag{7.4.12}$$

which also satisfies the boundary condition $T(b + 1) = 0$, (7.4.3b).
Similarly, if a is absorbing and b reflecting

$$\boxed{T(x) = \sum_{y=a}^{x} \phi(y) \sum_{z=y}^{b} 1/[t^+(z)\phi(z)]} \quad \begin{matrix} a & \text{absorbing} \\ b & \text{reflecting} \\ b > a \end{matrix} \tag{7.4.13}$$

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 $X \rightarrow 2Y$ 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

$$t(x \rightarrow x - 1, y \rightarrow y + 2) = k_1 x . \tag{7.5.1}$$

For the reverse reaction, one can assemble pairs of molecules of Y in $y(y-1)/2$ different ways. Hence

$$t(x \rightarrow x + 1, y \rightarrow y - 2) = k_2 y(y - 1) . \tag{7.5.2}$$

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:

$$\sum_a N_a^A X_a \underset{k_A^-}{\overset{k_A^+}{\rightleftharpoons}} \sum_a M_a^A X_a \qquad (A = 1, 2, \dots, s) . \tag{7.5.3}$$

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

$$
\begin{aligned}
\boldsymbol{x} &= (x_1, x_2, \dots, x_n) \\
\boldsymbol{N}^A &= (N_1^A, N_2^A, \dots, N_n^A) \\
\boldsymbol{M}^A &= (M_1^A, M_2^A, \dots, M_n^A)
\end{aligned}
\tag{7.5.4}
$$

and we also define

$$r^A = M^A - N^A .$$ (7.5.5)

Clearly, as reaction A proceeds one step in the forward direction,

$$x \to x + r^A$$ (7.5.6)

and in the backward direction,

$$x \to x - r^A .$$ (7.5.7)

The rate constants are defined by

$$t_A^+(x) = k_A^+ \prod_a \frac{x_a!}{(x_a - N_a^A)!}$$
$$t_A^-(x) = k_A^- \prod_a \frac{x_a!}{(x_a - M_a^A)!}$$ (7.5.8)

which are proportional, respectively, to the number of ways of choosing the combination N^A or M^A from x molecules. The Master equation is thus

$$\partial_t P(x, t) = \sum_A \{[t_A^-(x + r^A)P(x + r^A, t) - t_A^+(x)P(x, t)]$$
$$+ [t_A^+(x - r^A)P(x - r^A, t) - t_A^-(x)P(x, t)]\} .$$ (7.5.9)

This form is, of course, a completely general way of writing a time-homogeneous Master equation for an integer variable x in which steps of size 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 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)

$$t_A^-(x + r^A)P_s(x + r^A) = t_A^+(x)P_s(x)$$ (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 $P_s(x_0 + nr^A)$ for any n and any initial x_0. Using this method for all available A, we can generate $P_s(x)$ on the space of all x which can be written

$$x = x_0 + \sum n_A r^A ; \qquad (n_A \text{ integral})$$ (7.5.11)

but the solutions obtained may be ambiguous since, for example, from (7.5.10) we may write

$$P_s[(x + r^A) + r^B] = \frac{P_s(x) \, t_A^+(x) \, t_B^+(x + r^A)}{t_A^-(x + r^A) \, t_B^-(x + r^A + r^B)}$$

but

$$(7.5.12)$$

$$P_s[(x + r^B) + r^A] = \frac{P_s(x) \, t_B^+(x) \, t_A^+(x + r^B)}{t_B^-(x + r^B) \, t_A^-(x + r^A + r^B)}$$

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,

$$N^A \neq N^B$$
$$M^A \neq M^B \tag{7.5.13}$$

but $r^A = r^B \equiv r$.

In this case, uniqueness of $P_s(x + r^A)$ in (7.5.10) requires

$$\frac{t_A^+(x)}{t_A^-(x + r)} = \frac{t_B^+(x)}{t_B^-(x + r)} \tag{7.5.14}$$

and this means

$$\frac{k_A^+}{k_A^-} = \frac{k_B^+}{k_B^-} . \tag{7.5.15}$$

If there are two chains $A, B, C, \dots$, and $A', B', C', \dots$, of reactions such that

$$r^A + r^B + r^C + \dots = r^{A'} + r^{B'} + r^{C'} + \dots . \tag{7.5.16}$$

Direct substitution shows that

$$P_s(x + r^A + r^B + r^C + \dots) = P_s(x + r^{A'} + r^{B'} + r^{C'} + \dots) \tag{7.5.17}$$

only if

$$\boxed{\frac{k_A^+}{k_A^-} \frac{k_B^+}{k_B^-} \frac{k_C^+}{k_C^-} \dots = \frac{k_{A'}^+}{k_{A'}^-} \frac{k_{B'}^+}{k_{B'}^-} \frac{k_{C'}^+}{k_{C'}^-} \dots} \tag{7.5.18}$$

which is, therefore, the condition for detailed balance in a Master equation with combinatorial kinetics.

A solution for $P_s(x)$ in this case is a multivariate Poisson

$$P_s(x) = \prod_a \frac{\alpha_a^{x_a} \, e^{-\alpha_a}}{x_a !} \tag{7.5.19}$$

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

$$A \rightarrow 2Y: \quad t^+(y) = k_1 a \tag{7.5.26}$$
$$2Y \rightarrow A \quad t^-(y) = k_2 y(y - 1).$$
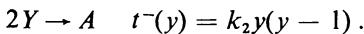
In this case, the conservation law is now simply that y is always even, or always odd. The stationary solution is of the form

$$P_s(y) = \frac{\alpha^y}{y!} \psi(y, \alpha) \tag{7.5.27}$$

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

$$x \sim \Omega \rho, \tag{7.5.28}$$

where ρ is the set of chemical concentrations. Thus, we must have $t_A^{\pm}(x)$ proportional to Ω as $\Omega \to \infty$, so that this requires

$$\begin{aligned}
k_A^+ &\sim \kappa_A^+ \, \Omega^{-\sum_a N_a^A + 1} \\
k_A^- &\sim \kappa_A^- \, \Omega^{-\sum_a M_a^A + 1} .
\end{aligned} \tag{7.5.29}$$

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

$$\partial_t P(x, t) = \sum_{A,n} \left\{ \frac{(r^A \cdot \nabla)^n}{n!} [t_A^-(x) P(x, t)] + \frac{(-r^A \cdot \nabla)^n}{n!} [t_A^+(x) P(x, t)] \right\} \tag{7.5.30}$$

and we now truncate this to second order to obtain

$$\partial_t P(x, t) = -\sum_a \partial_a [A_a(x) P(x, t)] + \tfrac{1}{2} \sum_{a,b} \partial_a \partial_b [B_{ab}(x) P(x, t)] , \tag{7.5.31}$$

$$\left.\begin{aligned}
A_a(x) &= \sum_A r_a^A [t_A^+(x) - t_A^-(x)] \\
B_{ab}(x) &= \sum_A r_a^A r_b^A [t_A^+(x) + t_A^-(x)] .
\end{aligned}\right\} \tag{7.5.32}$$

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,

$$t^+(x) = k_1 a x \tag{7.6.1}$$
$$t^-(x) = k_2 x(x - 1) \ .$$

Hence,

$$A(x) = k_1 a x - k_2 x(x-1) \sim k_1 a x - k_2 x^2 \quad \text{to order } 1/\Omega \tag{7.6.2}$$
$$B(x) = k_1 a x + k_2 x(x-1) \sim k_1 a x + k_2 x^2 \quad \text{to order } 1/\Omega \ .$$

7.6.2 $X \underset{k}{\overset{\gamma}{\rightleftharpoons}} Y \underset{\gamma}{\overset{k}{\rightleftharpoons}} A$

Here we have

$$\left.\begin{array}{l} t_1^+(x) = \gamma y \\ t_1^-(x) = kx \end{array}\right\} \quad r^1 = (1, -1) \tag{7.6.3}$$

$$\left.\begin{array}{l} t_2^+(x) = ka \\ t_2^-(x) = \gamma y \end{array}\right\} \quad r^2 = (0, 1) \ . \tag{7.6.4}$$

Hence,

$$A(x) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} (\gamma y - kx) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} (ka - \gamma y) \tag{7.6.5}$$

$$= \begin{bmatrix} \gamma y - kx \\ kx + ka - 2\gamma y \end{bmatrix} \tag{7.6.6}$$

$$\underline{B}(x) = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} (1, -1) \right\} (\gamma y + kx) + \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} (0, 1) \right\} (ka + \gamma y) \tag{7.6.7}$$

$$= \begin{bmatrix} \gamma y + kx & -\gamma y - kx \\ -\gamma y - kx & 2\gamma y + kx + ka \end{bmatrix} \ . \tag{7.6.8}$$

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

i) $X + A \rightarrow 2X$ $r^1 = (1, 0)$

ii) $X + Y \rightarrow 2Y$ $r^2 = (-1, 1)$ $\qquad\qquad$ (7.6.11)

iii) $Y \rightarrow B$ $r^3 = (0, -1)$.
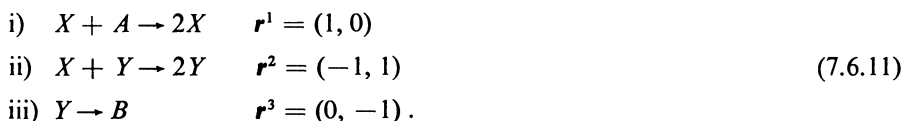
The reactions are all irreversible (though reversibility may be introduced) so we have

$$t_A^-(x) = 0 \qquad (A = 1, 2, 3)$$

but

$$t_1^+(x) = k_1 a \frac{x!}{(x-1)!} \frac{y!}{y!} = k_1 a x$$

$$t_2^+(x) = k_2 \frac{x!}{(x-1)!} \frac{y!}{(y-1)!} = k_2 xy \tag{7.6.12}$$

$$t_3^+(x) = k_3 \frac{x!}{x!} \frac{y!}{(y-1)!} = k_3 y .$$

The Master equation can now be explicitly written out using (7.5.9): one obtains

$$\partial_t P(x, y) = k_1 a(x-1) P(x-1, y) + k_2 (x+1)(y-1) P(x+1, y-1)$$
$$+ k_3 (y+1) P(x, y+1) - (k_1 a x + k_2 xy + k_3 y) P(x, y). \tag{7.6.13}$$

There are no exact solutions of this equation, so approximation methods must be used.

Kramers-Moyal. From (7.5.32)

$$A(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} k_1 a x + \begin{bmatrix} -1 \\ 1 \end{bmatrix} k_2 xy + \begin{bmatrix} 0 \\ -1 \end{bmatrix} k_3 y \tag{7.6.14}$$

$$= \begin{bmatrix} k_1 a x - k_2 xy \\ k_2 xy - k_3 y \end{bmatrix}, \tag{7.6.15}$$

$$\boldsymbol{B(x)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} (1, 0)k_1ax + \begin{bmatrix} -1 \\ 1 \end{bmatrix} (-1, 1)k_2xy + \begin{bmatrix} 0 \\ -1 \end{bmatrix} (0, -1)k_3y \tag{7.6.16}$$

$$= \begin{bmatrix} k_1ax + k_2xy & -k_2xy \\ -k_2xy & k_2xy + k_3y \end{bmatrix}. \tag{7.6.17}$$

The deterministic equations are

$$\frac{d}{dt} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} k_1ax - k_2xy \\ k_2xy - k_3y \end{bmatrix}. \tag{7.6.18}$$

Stationary State at   ·

$$\begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} k_3/k_2 \\ k_1a/k_2 \end{bmatrix}. \tag{7.6.19}$$

To determine the stability of this state, we check the stability of the iinearised deterministic equation

$$\frac{d}{dt} \begin{bmatrix} \delta x \\ \delta y \end{bmatrix} = \frac{\partial A(\boldsymbol{x}_s)}{\partial x_s} \delta x + \frac{\partial A(\boldsymbol{x}_s)}{\partial y_s} \delta y$$

$$= \begin{bmatrix} k_1a - k_2y_s \\ k_2y_s \end{bmatrix} \delta x + \begin{bmatrix} -k_2x_s \\ k_2x_s - k_3 \end{bmatrix} \delta y \tag{7.6.20}$$

$$= \begin{bmatrix} 0 & -k_3 \\ k_1a & 0 \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}. \tag{7.6.21}$$

The eigenvalues of the matrix are

$$\lambda = \pm \, i(k_1k_3a)^{1/2} \tag{7.6.22}$$

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

$$V = k_2(x + y) - k_3 \log x - k_1a \log y \tag{7.6.23}$$

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

$$\langle dV(x,y)\rangle = \left\langle \frac{\partial V}{\partial x}(k_1 ax - k_2 xy) + \frac{\partial V}{\partial y}(k_2 xy - k_3 y)\right\rangle dt \tag{7.6.29}$$

$$+ \left\langle B_{11}\frac{k_3}{2x^2} + B_{22}\frac{k_1}{2y^2}\right\rangle dt.$$

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

$$x = \langle x\rangle + \delta x \tag{7.6.34}$$
$$y = \langle y\rangle + \delta y$$

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

$$\frac{d}{dt}\begin{bmatrix}\langle x\rangle \\ \langle y\rangle\end{bmatrix} = \begin{bmatrix}k_1a\langle x\rangle - k_2\langle x\rangle\langle y\rangle \\ k_2\langle x\rangle\langle y\rangle - k_3\langle x\rangle\langle y\rangle\end{bmatrix} + \begin{bmatrix}-k_2\langle \delta x\delta y\rangle \\ k_2\langle \delta x\delta y\rangle\end{bmatrix} \tag{7.6.35}$$

$$\frac{d}{dt}\begin{bmatrix}\langle \delta x^2\rangle \\ \langle \delta x\delta y\rangle \\ \langle \delta y^2\rangle\end{bmatrix} = \begin{bmatrix}k_1a\langle x\rangle + k_2\langle x\rangle\langle y\rangle \\ -k_2\langle x\rangle\langle y\rangle \\ k_2\langle x\rangle\langle y\rangle + k_3\langle y\rangle\end{bmatrix}$$

$$+ \begin{bmatrix}2k_1a - 2k_2\langle y\rangle, & -2k_2\langle x\rangle & , 0 \\ k_2\langle y\rangle & , k_1a - k_3 + k_2(\langle x\rangle - \langle y\rangle), & -k_2\langle x\rangle \\ 0 & , 2k_2\langle y\rangle & , 2k_2\langle x\rangle - 2k_3\end{bmatrix}\begin{bmatrix}\langle \delta x^2\rangle \\ \langle \delta x\delta y\rangle \\ \langle \delta y^2\rangle\end{bmatrix} \tag{7.6.36}$$

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

$$x_s = y_s \tag{7.6.42}$$
$$g_s = \alpha^{-1}(x_s - \alpha x_s^2)$$

so that

$$f_s = x_s(-2\alpha x_s^2 + x_s(2 - \alpha) - 1)/(2 - 2\alpha x_s) \tag{7.6.43}$$
$$h_s = x_s(-2\alpha x_s^2 + x_s(2 + \alpha) + 1)/(2 - 2\alpha x_s)$$

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.

$$\left. \begin{array}{l} x_s = y_s = \tfrac{1}{2} \\ f_s = \tfrac{1}{2}\alpha/(\alpha - 2) \\ g_s = \tfrac{1}{4}(2 - \alpha)/\alpha \\ h_s = -1/(\alpha - 2) \end{array} \right\} \tag{7.6.45}$$

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,

$$f(t) = \frac{1}{2\alpha}(\cos 2t - 1) + \frac{2t}{\alpha}$$

$$g(t) = -\frac{1}{2\alpha}\sin 2t \tag{7.6.46}$$

$$h(t) = -\frac{1}{2\alpha}(\cos 2t - 1) + \frac{2t}{\alpha}.$$

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:

$$G(s, t) = \sum_{x} \left(\prod_a s_a^{x_a}\right) P(x, t) . \tag{7.6.47}$$

For we note that

$$\partial_t G(s, t) = \partial_t^+ G(s, t) + \partial_t^- G(s, t) \tag{7.6.48}$$

where the two terms correspond to the t^+ and t^- parts of the master equation. Thus

$$\partial_t^+ G(s, t) = \sum_{A,x} k_A^+ \left\{\prod_a \left[\frac{(x_a - r_a^A)!}{(x_a - r_a^A - N_a^A)!} s_a^{x_a}\right] P(x - r^A, t)\right.$$
$$\left. - \prod_a \left[\frac{x_a!}{(x_a - N_a^A)!} s_a^{x_a}\right] P(x, t)\right\} . \tag{7.6.49}$$

Changing the summation variable to $x - r^A$ and renaming this as x, in the first term we find

$$\partial_t^+ G(s, t) = \sum_{A,x} k_A^+ \left[\prod_a \frac{x_a!}{(x_a - N_a^A)!} s_a^{x_a + r_a^A} - \prod_a \frac{x_a!}{(x_a - N_a^A)!} s_a^{x_a}\right] P(x, t) . \tag{7.6.50}$$

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

$$\prod_a \frac{s^{x_a+r_a^A} x_a!}{(x_a - N_a^A)!} = \prod_a \left(\partial_a^{N_a^A} s_a^{x_a}\right) s_a^{M_a^A} \tag{7.6.52}$$

so that

$$\partial_t^+ G(s, t) = \sum_A k_A^+ \left(\prod_a s_a^{M_a^A} - \prod_a s_a^{N_a^A}\right) \partial_a^{N_a^A} G(s, t). \tag{7.6.53}$$

Similarly, we derive a formula for $\partial_t^- G(s, t)$ and put these together to get

$$\boxed{\partial_t G(s, t) = \sum_A \left(\prod_a s_a^{M_a^A} - \prod_a s_a^{N_a^A}\right)\left(k_A^+ \prod_a \partial_a^{N_a^A} - k_A^- \prod_a \partial_a^{M_a^A}\right) G(s, t)} \tag{7.6.54}$$

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)

$$A + X \xrightarrow{k_2} 2X + D \qquad N^1 = 1, M^1 = 2; r^1 = 1$$
$$k_1^+ = k_2 A \tag{7.6.55}$$
$$k_1^- = 0$$

$$B + X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} C \qquad N^2 = 1, M^2 = 0; r^2 = -1$$

$$k_2^+ = k_1 B \tag{7.6.56}$$
$$k_2^- = k_3 C .$$
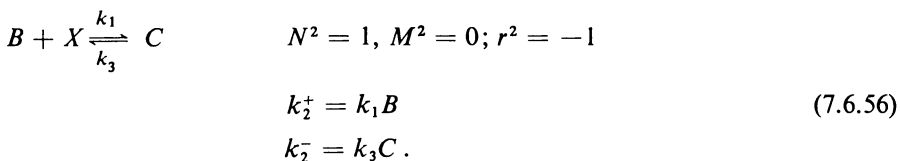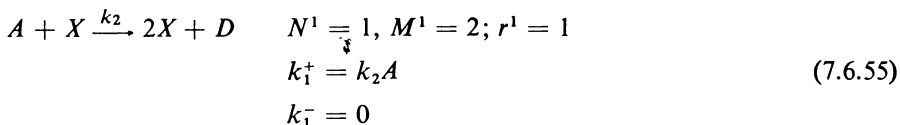
Hence, from (7.6.54), the generating function equation is

$$\partial_t G = (s^2 - s)(k_2 A \partial_s G) + (1 - s)(k_1 B \partial_s G - k_3 CG) . \tag{7.6.57}$$

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

$$G_y(s, 0) = s^y \tag{7.6.63}$$

$$\Rightarrow F(z) = (1-\beta z)^y(1-\alpha z)^{-\gamma/\alpha-y}(\beta-\alpha)^{\gamma/\alpha} \tag{7.6.64}$$

$$\boxed{\begin{aligned}\Rightarrow G_y(s, t) &= \lambda^{\gamma/\alpha}[\beta(1-e^{-\lambda t}) - s(\alpha-\beta e^{-\lambda t})]^y \\ &\times [(\beta-\alpha e^{-\lambda t}) - \alpha s(1-e^{-\lambda t})]^{-\gamma/\alpha-y}\end{aligned}} \tag{7.6.65}$$

(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

$$\frac{d}{dt}\langle x(t)\rangle = (k_2 A - k_1 B)\langle x(t)\rangle + k_3 C \tag{7.6.69}$$

and

$$\begin{aligned}\frac{d}{dt}\langle x(t)[x(t)-1]\rangle &= 2(k_2 A - k_1 B)\langle x(t)[x(t)-1]\rangle \\ &\quad + 2k_2 A\langle x(t)\rangle + 2k_3 C\langle x(t)\rangle .\end{aligned} \tag{7.6.70}$$

These equations have a stable stationary solution provided

$k_2A < k_1B,$ i.e., $\alpha < \beta$.

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 var$\{x\}_s$ both become very large as a critical point is approached and, in fact,

$$\frac{\text{var}\,\{x_s\}}{\langle x \rangle_s} = \frac{k_1B}{k_1B - k_2A} \longrightarrow \infty \,. \tag{7.6.73}$$

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

$$\langle x(t), x(0) \rangle_s = \exp\left[(k_2A - k_1B)t\right]\text{var}\,\{x\}_s \tag{7.6.74}$$

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

$$k_2ds_1 + k_1ds_2 = 0 \Longrightarrow k_2s_1 + k_1s_2 = v \,. \tag{7.6.77}$$

$$(k_1 + k_2)dt = \frac{d(s_2 - s_1)}{s_2 - s_1}$$

$$\implies (s_2 - s_1)e^{-(k_1+k_2)t} = u \tag{7.6.78}$$

$$\therefore G(s_1, s_2, t) = F[k_2 s_1 + k_1 s_2, (s_2 - s_1)e^{-(k_1+k_2)t}] . \tag{7.6.79}$$

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:

$$G(s_1, s_2, t) = \exp\left\{ \frac{k_2\beta - k_1\alpha}{k_1 + k_2}(s_2 - s_1)e^{-(k_1+k_2)t} \right.$$

$$\left. + \frac{\alpha + \beta}{k_1 + k_2}[k_1(s_2 - 1) + k_2(s_1 - 1)] \right\}. \tag{7.6.81}$$

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 $P(x, t)$ as a superposition of multivariate uncorrelated Poissons:

$$P(x, t) = \int d\alpha \prod_a \frac{e^{-\alpha_a}\alpha_a^{x_a}}{x_a!} f(\alpha, t) . \tag{7.7.1}$$

This means that the generating function $G(s, t)$ can be written

$$G(s, t) = \int d\alpha \, \exp[\sum_a (s_a - 1)\alpha_a] f(\alpha, t) . \tag{7.7.2}$$

We substitute this in the generating function equation (7.6.54) to get

$$\partial_t G(s, t) = \sum_A \int d\boldsymbol{\alpha} \left\{ \left[\prod_a \left(\frac{\partial}{\partial \alpha_a} + 1 \right)^{M_a^A} - \prod_a \left(\frac{\partial}{\partial \alpha_a} + 1 \right)^{N_a^A} \right] \right.$$

$$\left. \times \left(k_A^+ \prod_a \alpha_a^{N_a^A} - k_A^- \prod_a \alpha_a^{M_a^A} \right) \exp \left[\sum_a (s_a - 1)\alpha_a \right] \right\} f(\boldsymbol{\alpha}, t) . \qquad (7.7.3)$$

We now integrate by parts, drop surface terms and finally equate coefficients of the exponential to obtain

$$\frac{\partial f(\boldsymbol{\alpha}, t)}{\partial t} = \sum_A \left[\prod_a \left(1 - \frac{\partial}{\partial \alpha_a} \right)^{M_a^A} - \prod_a \left(1 - \frac{\partial}{\partial \alpha_a} \right)^{N_a^A} \right]$$

$$\times \left[k_A^+ \prod_a \alpha_a^{N_a^A} - k_A^- \prod_a \alpha_a^{M_a^A} \right] f(\boldsymbol{\alpha}, t) . \qquad (7.7.4)$$

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,

$$\sum_a M_a^A \leqslant 2$$
$$\sum_a N_a^A \leqslant 2 \qquad (7.7.5)$$

which indicates that only pairs of molecules at the most participate in reactions. The FPE can then be simplified as follows. Define the currents

$$J_A(\boldsymbol{\alpha}) = k_A^+ \prod_a \alpha_a^{N_a^A} - k_A^- \prod_a \alpha_a^{M_a^A}, \qquad (7.7.6)$$

the drifts

$$A_a[J(\boldsymbol{\alpha})] = \sum_A r_a^A J_A(\boldsymbol{\alpha}) , \qquad (7.7.7)$$

and the diffusion matrix elements by

$$B_{ab}[J(\boldsymbol{\alpha})] = \sum_A J_A(\boldsymbol{\alpha})(M_a^A M_b^A - N_a^A N_b^A - \delta_{a,b} r_a^A) . \qquad (7.7.8)$$

Then the Poisson representation FPE is

$$\frac{\partial f(\boldsymbol{\alpha}, t)}{\partial t} = - \sum_a \frac{\partial}{\partial \alpha_a} \{ A_a[J(\boldsymbol{\alpha})] f(\boldsymbol{\alpha}, t) \}$$

$$+ \tfrac{1}{2} \sum_{a,b} \frac{\partial^2}{\partial \alpha_a \partial \alpha_b} \{ B_{ab}[J(\boldsymbol{\alpha})] f(\boldsymbol{\alpha}, t) \} . \qquad (7.7.9)$$

Notice also that if we use the explicit volume dependence of the parameters given in Sect. 7.5.3 (7.5.29) and define

$$\eta_a = \alpha_a/V \qquad (7.7.10)$$

$$\varepsilon = V^{-1/2} \qquad (7.7.11)$$

and $F(\eta, t)$ is the quasiprobability in the η variable, then the FPE for the η variable takes the form of

$$\frac{\partial F(\eta, t)}{\partial t} = -\sum_a \frac{\partial}{\partial \eta_a}[\hat{A}_a(\eta)F(\eta, t)] + \frac{\varepsilon^2}{2}\sum_{a,b}\frac{\partial^2}{\partial \eta_a \partial \eta_b}[\hat{B}_{a,b}(\eta)F(\eta,t)] \qquad (7.7.12)$$

with

$$\hat{A}_a(\eta) = \sum_A r_a^A \hat{J}_A(\eta) \qquad (7.7.13a)$$

$$\hat{J}_A(\eta) = \kappa_A^+ \prod_a \eta_a^{N_a^A} - \kappa_A^- \prod_a \eta_a^{M_a^A} \qquad (7.7.13b)$$

$$\hat{B}_{ab}(\eta) = \sum_A \hat{J}_A(\eta)(M_a^A M_b^A - N_a^A N_b^A - \delta_{a,b}r_a^A). \qquad (7.7.13c)$$

In this form we see how the system size expansion in $V^{-1/2}$ corresponds exactly to a small noise expansion in η 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,

$$\sum_a M_a^A \leqslant 1 \qquad \text{and} \qquad \sum_a N_a^A \leqslant 1,$$

then it is easily checked that the diffusion coefficient $\hat{B}_{ab}(\eta)$ in (7.7.13) vanishes, and we have a Liouville equation. An initially Poissonian $P(x, t_0)$, corresponds to a delta function $F(\eta, t_0)$, and the time evolution generated by this Liouville equation will generate a delta function solution, $\delta(\eta - \bar{\eta}(t))$, where $\bar{\eta}(t)$ is the solution of

$$d\eta/dt = A(\eta)$$

This means that $P(x, t)$ will preserve a Poissonian form, with mean equal to $\bar{\eta}(t)$. 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 $F(\eta, t_0)$ which are not delta functions.

c) Example

As an example, consider the reaction pair

i) $\quad A + X \underset{k_4}{\overset{k_2}{\rightleftharpoons}} 2X$

ii) $\quad B + X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} C$

$$(7.7.14)$$
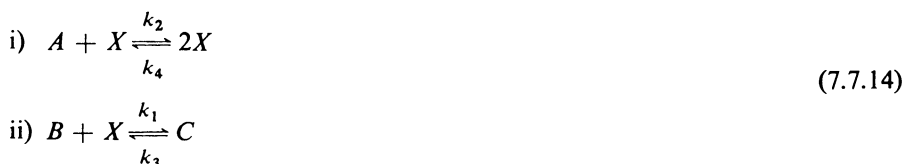
$$N^1 = 1, \qquad M^1 = 2, \qquad k_1^+ = k_2 A, \qquad k_1^- = k_4$$
$$N^2 = 1, \qquad M^2 = 0, \qquad k_2^+ = k_1 B, \qquad k_2^- = k_3 C$$

so that (7.7.4) takes the form

$$\frac{\partial f}{\partial t} = \left[\left(1 - \frac{\partial}{\partial \alpha} \right)^2 - \left(1 - \frac{\partial}{\partial \alpha} \right) \right] (k_2 A \alpha - k_4 \alpha^2) f \tag{7.7.15}$$
$$+ \left[1 - \left(1 - \frac{\partial}{\partial \alpha} \right) \right] (k_1 B \alpha - k_3 C) f$$

$$\frac{\partial f}{\partial t} = \left\{ -\frac{\partial}{\partial \alpha} [k_3 C + (k_2 A - k_1 B)\alpha - k_4 \alpha^2] + \frac{\partial^2}{\partial \alpha^2} [k_2 A \alpha - k_4 \alpha^2] \right\} f \tag{7.7.16}$$

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)

$$\langle x^r \rangle_f = \sum_x \int d\alpha \, [x(x-1) \dots (x - r + 1) \frac{e^{-\alpha} \alpha^x}{x!} f(\alpha) \tag{7.7.17}$$
$$= \int d\alpha \, \alpha^r f(\alpha) \equiv \langle \alpha^r \rangle .$$

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

$$\delta > 0 \tag{7.7.24}$$
$$k_3 > 0.$$

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

$$[\{(k_2A\alpha - k_4\alpha^2 - k_1B\alpha + k_3C)f - \partial_{\alpha}[(k_2\alpha - k_4\alpha^2)f]\} \{e^{(s-1)\alpha}\}]_b^a$$
$$+ [(k_2\alpha - k_4\alpha^2)f[(s - 1)e^{(s-1)\alpha}]]_b^a . \tag{7.7.25}$$

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 $(0, k_2 A / k_4)$. In this range, the diffusion coefficient $(k_2 A \alpha - k_4 \alpha^2)$ is positive. The deterministic mean of α, given by

$$\alpha = \frac{k_2 A - k_1 B + [(k_2 A - k_1 B)^2 + 4 k_3 k_4 C]^{1/2}}{2 k_4} \tag{7.7.36}$$

lies *within* the interval $(0, k_2 A / k_4)$. 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 $\alpha = k_2 A / k_4$ 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

$$\alpha_0 = k_2 A / k_4 .$$

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 $\alpha = k_2 A / k_4$, 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
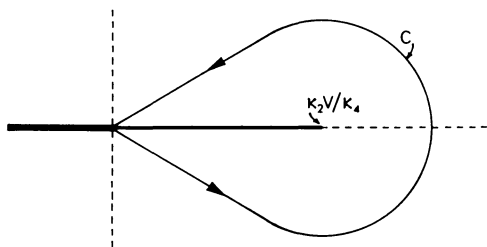
Fig. 7.3. Contour C in the complex plane for the evaluation of (7.7.38)

$$\text{var}\{\alpha\} = \langle\alpha^2\rangle - \langle\alpha\rangle^2, \tag{7.7.39}$$

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

$$\langle x \rangle = V(a/2)^{1/2} + \tfrac{1}{8} + O(1/V) \tag{7.7.45}$$

$$\mathrm{var}\{x\} = \tfrac{3}{4} V(a/2)^{1/2} - \tfrac{1}{16} + O(1/V) .$$

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(α) 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

$$f_p(\alpha) = (2\pi\sigma^2)^{-1} \exp\left(-|\alpha - \alpha_0|^2/2\sigma^2\right) \tag{7.7.48}$$

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

$$\int (2\pi\sigma^2)^{-1}d^2\alpha \exp\left(-|\alpha - \alpha_0|^2/2\sigma^2\right)g(\alpha) = g(\alpha_0)\,, \qquad (7.7.50)$$

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

$$P(x) = \int d^2\alpha f_p(\alpha)e^{-\alpha}\alpha^x/x! = e^{-\alpha_0}\alpha_0^x/x!\,. \qquad (7.7.51)$$

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

$$\partial_t G(s, t) = \int d^2\alpha f(\alpha, t)\left(\sum_a A_a \frac{\partial}{\partial\alpha_a} + \sum_{a,b} \tfrac{1}{2} B_{ab}\frac{\partial^2}{\partial\alpha_a\partial\alpha_b}\right)\exp\left[\sum_a(s_a - 1)\alpha_a\right]. \qquad (7.7.52)$$

We now take explicit account of the fact that $\boldsymbol{\alpha}$ is a complex variable

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_x + i\boldsymbol{\alpha}_y \qquad (7.7.53)$$

and also write

$$A(\boldsymbol{\alpha}) = A_x(\boldsymbol{\alpha}) + iA_y(\boldsymbol{\alpha})\,. \qquad (7.7.54)$$

We further write

$$\underline{B}(\boldsymbol{\alpha}) = \underline{C}(\boldsymbol{\alpha})\underline{C}^{\mathrm{T}}(\boldsymbol{\alpha}) \qquad (7.7.55)$$

and

$$\underline{C}(\boldsymbol{\alpha}) = \underline{C}_x(\boldsymbol{\alpha}) + i\underline{C}_y(\boldsymbol{\alpha}) \qquad (7.7.56)$$

For brevity we use

$$\partial_a = \frac{\partial}{\partial\alpha_a}$$

$$\partial_a^x = \frac{\partial}{\partial\alpha_{x,a}} \qquad (7.7.57)$$

$$\partial_a^y = \frac{\partial}{\partial\alpha_{y,a}}\,.$$

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

$$\partial_a \leftrightarrow \partial_a^x \leftrightarrow -i\partial_a^y\,. \qquad (7.7.58)$$

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 A or $\underset{\sim}{C}$ is x or y respectively. We then derive

$$\partial_t G(s, t) = \int d^2\alpha\, f(\alpha, t) \{[\sum_a (A_{a;x}\partial_a^x + A_{a;y}\partial_a^y)$$
$$+ \tfrac{1}{2}\sum_{a,b,c}(C_{a,c,x}C_{c,b;x}\partial_a^x\partial_b^x + C_{a,c;y}C_{c,b;y}\partial_a^y\partial_b^y$$
$$+ 2C_{a,c;x}C_{c,b;y}\partial_a^x\partial_b^y]\exp[\sum_a(s_a - 1)\alpha_a]\}. \qquad (7.7.59)$$

Integrating by parts and discarding the surface terms to get a FPE in the variables (α_x, α_y),

$$\partial_t f(\alpha, t) = [-\sum_a(\partial_a^x A_{a;x} + \partial_a^y A_{a;y}) + \tfrac{1}{2}\sum_{a,b,c}(\partial_a^x\partial_b^x C_{a,c;x}C_{c,b;x}$$
$$+ \partial_a^y\partial_b^y C_{a,c;\,y}C_{c,b;\,y} + 2\partial_a^x\partial_b^y C_{a,c;x}C_{c,b;\,y})]f(\alpha, t). \qquad (7.7.60)$$

In the space of doubled dimensions, this is a FPE with positive semidefinite diffusion. For, we have for the variable (α_x, α_y) the drift vector

$$\mathscr{A}(\alpha) = [A_x(\alpha), A_y(\alpha)] \qquad (7.7.61)$$

and the diffusion matrix

$$\mathscr{B}(\alpha) = \begin{bmatrix} \underset{\sim}{C}_x\underset{\sim}{C}_x^T & \underset{\sim}{C}_x\underset{\sim}{C}_y^T \\ \underset{\sim}{C}_y\underset{\sim}{C}_x^T & \underset{\sim}{C}_y\underset{\sim}{C}_y^T \end{bmatrix} = \mathscr{C}(\alpha)\mathscr{C}(\alpha)^T \qquad (7.7.62)$$

where

$$\mathscr{C}(\alpha) = \begin{bmatrix} \underset{\sim}{C}_x & 0 \\ \underset{\sim}{C}_y & 0 \end{bmatrix} \qquad (7.7.63)$$

so that $\mathscr{B}(\alpha)$ 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

$$\begin{bmatrix} d\alpha_x \\ d\alpha_y \end{bmatrix} = \begin{bmatrix} A_x(\alpha) \\ A_y(\alpha) \end{bmatrix}dt + \begin{bmatrix} \underset{\sim}{C}_x dW(t) \\ \underset{\sim}{C}_y dW(t) \end{bmatrix} \qquad (7.7.64)$$

where $W(t)$ is a Wiener process of the same dimension as α_x. Note that the same Wiener process occurs in both lines because of the two zero entries $\mathscr{C}(\alpha)$ as written in (7.7.63).

Recombining real and imaginary parts, we find the SDE for the complex variable α:

$$d\alpha = A(\alpha)dt + \underset{\sim}{C}(\alpha)dW(t). \qquad (7.7.65)$$

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 $\underset{\sim}{C}(\alpha)$

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)

$$A + X \underset{k_4}{\overset{k_2}{\rightleftharpoons}} 2X$$

$$B + X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} C. \tag{7.7.66}$$

The use of the positive Poisson representation applied to this system yields the SDE, arising from the FPE (7.7.16):

$$d\alpha = [k_3C + (k_2A - k_1B)\alpha - k_4\alpha^2]dt + [2(k_2A\alpha - k_4\alpha^2)]^{1/2}dW(t) . \tag{7.7.67}$$

In the case $\delta > 0$, we note that the noise term vanishes at $\alpha = 0$ and at $\alpha = k_2A/k_4$, is positive between these points and the drift term is such as to return α to the range $[0, k_2A/k_4]$ whenever it approaches the end points. Thus, for $\delta > 0$, (7.7.67) represents a real SDE on the real interval $[0, k_2A/k_4]$.

In the case $\delta < 0$, the stationary point lies outside the interval $[0, k_2A/k_4]$, 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 $[0, k_2A/k_4]$ again.

The case of $\delta = 0$ is not very dissimilar, except that once the point reaches the right-hand end of the interval $[0, k_2A/k_4]$, both drift and diffusion vanish so it remains there from then on.

In the case of the system

$$B \rightarrow X$$

$$2X \rightarrow A , \tag{7.7.68}$$
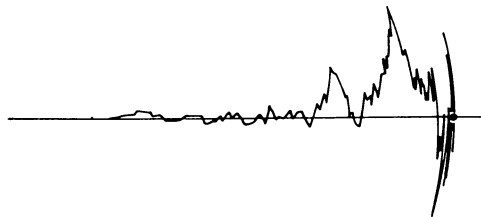
Fig. 7.4. Path followed by a point obeying the stochastic differential equation (7.7.67)
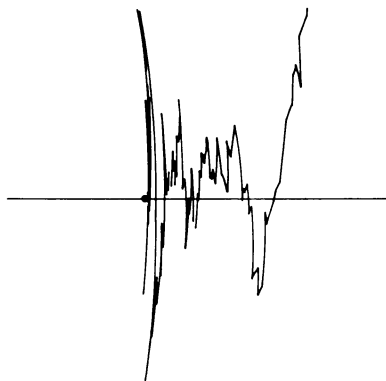
Fig. 7.5. Simulation of the path of a point obeying the stochastic differential equation (7.7.69)   ▶

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 $\text{Re}\{\alpha\} = (a/2)^{1/2}$ 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 $X \rightleftharpoons Y$ 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

$$\int d\mu(\alpha)\, \alpha f(\alpha, t | \alpha', s) = \sum_{x, x'} x P(x, t | x', s)e^{-\alpha'}\alpha'^{x'}/x'!$$

Hence,

$$\langle\alpha(t)\alpha(s)\rangle = \sum_{x, x'} x P(x, t | x', s) \int d\mu(\alpha')(\alpha'^{x'+1}e^{-\alpha'}/x'!)f(\alpha', s)$$

$$= \sum_{x, x'} x P(x, t | x', s) \int d\mu(\alpha')\left[\left(-\alpha'\frac{\partial}{\partial\alpha'} + x'\right)(\alpha'^{x'}e^{-\alpha'}/x'!)\right]f(\alpha', s)$$

$$= \sum_{x, x'} x\, x' P(x, t | x', s)P(x', s) \tag{7.7.72}$$

$$- \int d\mu(\alpha')\, f(\alpha', s)\alpha'\frac{\partial}{\partial\alpha'}\sum_{x, x'} x P(x, t | x', s)(\alpha'^{x'}e^{-\alpha'}x!) . \tag{7.7.73}$$

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

$$- \int d\mu(\alpha')\alpha' \frac{\partial}{\partial\alpha'} \langle\alpha(t)|[\alpha', s]\rangle f(\alpha', s) \equiv \left\langle \alpha' \frac{\partial}{\partial\alpha'} \langle\alpha(t)|[\alpha, s]\rangle \right\rangle \qquad (7.7.75)$$

so we have

$$\langle x(t)x(s)\rangle = \langle\alpha(t)\alpha(s)\rangle + \left\langle \alpha' \frac{\partial}{\partial\alpha'} \langle\alpha(t)|[\alpha, s]\rangle \right\rangle . \qquad (7.7.76)$$

Taking into account a many-variable situation and noting that

$$\langle x(t)\rangle = \langle\alpha(t)\rangle \qquad \text{always} ,$$

we have

$$\boxed{\langle x_a(t), x_b(s)\rangle = \langle\alpha_a(t), \alpha_b(s)\rangle + \left\langle \alpha'_b \frac{\partial}{\partial\alpha'_b} \langle\alpha_a(t)|[\alpha', s]\rangle \right\rangle .} \qquad (7.7.77)$$

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 $A, B, C, ...$, the distribution function in the grand canonical ensemble is given by

$$P(I) = \exp\{\beta[\Omega + \sum_i \mu_i x_i(I) - E(I)]\}, \qquad (7.7.78)$$

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

$$\beta = 1/kT . \qquad (7.7.79)$$

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

$$\sum_i \nu_i^A x_i(I) = \sum_i \nu_i^A x_i(J), \qquad A = 1, 2, 3, ... \qquad (7.7.80)$$

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 x of molecules of reacting components in the grand canonical ensemble of an ideal reacting system is easily evaluated, namely,

$$P(x) = \exp\left[\beta(\Omega + \sum_i \mu_i x_i)\right] \sum_I \prod_i \delta[x_i(I), x_i] \exp\left[-\beta E(I)\right] . \tag{7.7.85}$$

The sum over states is the same as that for the canonical ensemble of an ideal nonreacting mixture so that

$$P(x) = \exp\left[\beta(\Omega + \sum_i \mu_i x_i)\right] \prod_i \frac{1}{x_i!} \{\sum_k \exp\left[-\beta E_k(i)\right]\}^{x_i} , \tag{7.7.86}$$

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

$$\log\langle x_i \rangle = \beta \mu_i - \log\left[\sum_k e^{-\beta E_k(i)}\right] \tag{7.7.87}$$

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

$$P(x) \propto \left\{ \prod_i \frac{1}{x_i!} [\sum_k e^{-\beta E_k(i)}]^{x_i} \right\} \sum_A \delta[\sum_i \nu_i^A x_i, \tau^A] . \tag{7.7.88}$$

In terms of the Poisson representation, we have just shown that in equilibrium situations, the quasiprobability (in a grand canonical ensemble) is

$$f(\alpha)_{eq} = \delta[\alpha - \alpha(eq)] \tag{7.7.89}$$

since the x space distribution is Poissonian. For the time correlation functions there are two results of this.

i) The variables $\alpha(t)$ and $\alpha(s)$ are nonfluctuating quantities with values $\alpha(eq)$. 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,

$$\langle x_a(t), x_b(s) \rangle = \left[\alpha_b' \frac{\partial}{\partial \alpha_b'} \langle \alpha_a(t) | [\alpha', s] \rangle \right]_{\alpha' = \alpha(eq)} . \tag{7.7.91}$$

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

$$Z^{-1}(\mu) \exp \left[\{ \sum_i \mu_i x_i(I) - E(I) \} / kT \right] \tag{7.7.92}$$

with

$$Z(\mu) = \exp(-\Omega \beta) , \tag{7.7.93}$$

where $Z(\mu)$ 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,

$$\langle x_i, t | [\mu, s] \rangle = \sum_j \langle x_i, t | [J, s] \rangle Z^{-1}(\mu)$$

$$\times \exp \left\{ \frac{1}{kT} [\sum_j \mu_j x_j(J) - E(J)] \right\}. \tag{7.7.95}$$

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

$$\mu_i(\langle \mathbf{x} \rangle) = kT \log \langle x_i \rangle + \text{const}. \tag{7.7.97}$$

Using (7.7.97), we find that (7.7.96) becomes

$$\langle x_i(t), x_j(s) \rangle = \left[\langle x_j \rangle \frac{\partial}{\partial \langle x_j \rangle} \langle x_i, t | [\mu(\langle \mathbf{x} \rangle), s] \rangle \right]_{\langle \mathbf{x} \rangle = \langle \mathbf{x} \rangle_{eq}}. \tag{7.7.98}$$

Since the ideal solution theory gives rise to a distribution in x_i that is Poissonian, it follows that in that limit

$$\langle x_i, t | [\mu(\langle \mathbf{x} \rangle), s] \rangle = \langle \alpha_i, t | [\alpha', s] \rangle \tag{7.7.99}$$

with $\alpha' = \langle \mathbf{x} \rangle$. Thus, (7.7.98) becomes

$$\langle x_i(t), x_j(s) \rangle = \left[\alpha'_j \frac{\partial}{\partial \alpha_j} \langle \alpha_i, t | [\alpha', s] \rangle \right]_{\alpha' = \alpha(eq)}. \tag{7.7.100}$$

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

$$d\eta = \hat{A}(\eta)dt + \varepsilon\hat{C}(\eta)dW(t) \, , \tag{7.7.101}$$

where

$$\hat{C}\hat{C}^{\mathrm{T}} = \hat{B} \, . \tag{7.7.102}$$

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

$$\hat{A}(\bar{\eta}) = 0 \tag{7.7.104}$$
$$d\eta_1 = -F\eta_1 dt + G\, dW(t)$$

where

$$F_{rs} = -\frac{\partial}{\partial\eta_s}\,\hat{A}_r(\bar{\eta}) \tag{7.7.105}$$

$$G = \hat{C}(\bar{\eta}) \, . \tag{7.7.106}$$

Then,

$$\langle \alpha_r(t), \alpha_s(0)\rangle_s = V\sum_{r'} [\exp(-Ft)]_{r,r'} \langle \eta_{r',1}, \eta_{s,1}\rangle_s \tag{7.7.106}$$

and

$$\frac{\partial}{\partial\alpha_s'}\langle \alpha_r(t)\,|\,[\alpha',\,0]\rangle = \frac{\partial}{\partial\eta_{s,1}'}\langle \eta_{r,1}(t)\,|\,[\eta_1',\,0]\rangle = [\exp(-Ft)]_{r,s} \, . \tag{7.7.107}$$

Hence,

$$\langle x_r(t), x_s(0)\rangle_s = V\sum_{r'} [\exp(-Ft)]_{r,r'}[\langle \eta_{r',1}, \eta_{s,1}\rangle_s + \delta_{r',s}\bar{\eta}_s) \tag{7.7.108}$$

$$= \sum_{r'} \exp(-Ft)_{r,r'}\langle x_{r'}, x_s\rangle_s \, . \tag{7.7.109}$$

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

i) $A + Y \underset{1}{\overset{1}{\rightleftharpoons}} X + Y$ (7.7.111a)

ii) $Y \underset{1}{\overset{\gamma}{\rightleftharpoons}} 2X ,$ (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

$$\frac{dx}{dt} = ay - xy + 2(\gamma y - x^2)$$
$$\frac{dy}{dt} = x^2 - \gamma y$$
(7.7.112)

and the usual deterministic adiabatic elimination procedure sets $y = x^2/\gamma$ and gives

$$\frac{dx}{dt} = (ax^2 - x^3)/\gamma .$$ (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),

$$\frac{\partial f}{\partial t} = \frac{1}{\gamma} \left\{ \left(-\frac{\partial}{\partial \alpha} + 2 \frac{\partial^2}{\partial \alpha^2} - \frac{\partial^3}{\partial \alpha^3} \right) [\alpha^2(a - \alpha)] f \right\}$$ (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

$$
\frac{\partial f(x,y)}{\partial t} = \Bigg\{ -\left(\frac{\partial}{\partial x} - 2x\frac{\partial}{\partial y}\right)\left[\frac{(a^{\frac{1}{2}} x)(y + x^2)}{\gamma} + 2y\right] + \gamma\frac{\partial}{\partial y}y
$$
$$
+ \left(\frac{\partial}{\partial x} - 2x\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} - 2x\frac{\partial}{\partial y}\right)y + \left(\frac{\partial}{\partial x} - 2x\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial y}\right)[(y + x^2)(a - x)]\Bigg\} f.
\tag{7.7.118}
$$

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

$$
\gamma L_1(\gamma) = \gamma\left\{\frac{\partial}{\partial v}v + \frac{\partial^2}{\partial v^2}[2x^3(x - a) + (4x^2 - 2x)v\gamma^{-1/2}]\right\}
$$
$$
\xrightarrow[\gamma\to\infty]{} \gamma\left\{\frac{\partial}{\partial v}v + \frac{\partial^2}{\partial v^2}[2x^3(x - a)]\right\}
$$
$$
\equiv \gamma L_1 .
\tag{7.7.121}
$$

With this substitution, we finally identify

$$\gamma^{-1}L_3 = -\gamma^{-1}\frac{\partial}{\partial x}[x^2(a-x)] \qquad (7.7.122)$$

$$L_2(\gamma) = -\frac{\partial}{\partial x}[(a-x)v\gamma^{-3/2} + 2v\gamma^{-1/2}] - \frac{\partial}{\partial x}2x\frac{\partial}{\partial v}v - 2x\frac{\partial}{\partial v}\frac{\partial}{\partial x}v$$

$$+ \gamma^{-1/2}\frac{\partial^2}{\partial x^2}v + \gamma^{1/2}\frac{\partial}{\partial x}\frac{\partial}{\partial v}[(a-x)(x^2+v\gamma^{-1/2})] \qquad (7.7.123)$$

and

$$\frac{\partial f}{\partial t} = [\gamma^{-1}L_3 + L_2(\dot\gamma) + \gamma L_1(\gamma)]f. \qquad (7.7.124)$$

The projection operator P will be onto the null space of L_1 and because L_1 depends on x, we have

$$L_3 P \neq P L_3 . \qquad (7.7.125)$$

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

$$s\bar{g}(s) = \gamma^{-1}PL_3\bar{g}(s) + P[L_2(\gamma) + \gamma^{-1}L_3][s - \gamma L_1 - (1-P)L_2(\gamma) - \gamma^{-1}(1-P)L_3]^{-1}$$
$$\times [L_2(\gamma) + \gamma^{-1}(1-P)L_3]\bar{g}(s) + g(0). \qquad (7.7.126)$$

Notice, however, since for any function of v

$$P\phi(v) = p_x(v)\int dv\,\phi(v) \qquad (7.7.127)$$

where $p_x(v)$ satisfies

$$L_1 p_x(v) = 0 , \qquad (7.7.128)$$

that in $PL_2(\gamma)$, all terms with $\partial/\partial v$ in them vanish. Thus, to highest order in γ,

$$PL_2(\gamma) = \gamma^{-1/2}\left(-2v\frac{\partial}{\partial x} + v\frac{\partial^2}{\partial x^2}\right). \qquad (7.7.129)$$

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.,

$$\gamma^{1/2}\frac{\partial}{\partial x}[(a-x)x^2]\frac{\partial}{\partial v}. \qquad (7.7.130)$$

Thus, the large γ limit of (7.7.126) is

$$s\tilde{g}(s) = \gamma^{-1}\left\{PL_3\tilde{g} - P\left[-2\frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2}\right]vL_1^{-1}\frac{\partial}{\partial x}\left[(a-x)x^2\frac{\partial}{\partial v}\right]p_x(v)\tilde{p}\right\} + g(0)$$
(7.7.131)

where we have written

$$g = p_x(v)p, \qquad \tilde{g} = p_x(v)\tilde{p}.$$
(7.7.132)

We are now lead to the central problem of the evaluation of

$$\int dv'v'L_1^{-1}\frac{\partial}{\partial x}(a-x)x^2\frac{\partial}{\partial v'}p_x(v')$$
(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

$$\left[L_1^{-1}, \frac{\partial}{\partial x}\right] = L_1^{-1}\left[\frac{\partial}{\partial x}, L_1\right]L_1^{-1}$$
(7.7.134)

and from (7.7.121),

$$= L_1^{-1}\left\{\frac{\partial^2}{\partial v^2}[8x^3 - 6ax^2]\right\}L_1^{-1};$$
(7.7.135)

$$(7.7.133) = \frac{\partial}{\partial x}\int dv'v'L_1^{-1}(a-x)x^2\frac{\partial}{\partial v'}p_x(v')$$

$$+ \int dv'v'L_1^{-1}\frac{\partial^2}{\partial v'^2}L_1^{-1}[(8x^3 - 6ax^2)(a-x)x^2]\frac{\partial}{\partial v'}p_x(v').$$
(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

$$L_1^{-1}\frac{\partial^2}{\partial v^2}L_1^{-1}\frac{\partial}{\partial v}p_x(v)$$
(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

$$(7.7.133) = \frac{\partial}{\partial x}(a-x)x^2.$$
(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

$$p(v, t \,|\, v_0, t_0) = (1/2\pi) \int_{-\infty}^{\infty} dq \exp \{i[q(v - v_0) + \tfrac{1}{6} q^3(t - t_0)]\} . \tag{7.7.143}$$

The moments of V can be calculated, after a partial integration, to be

$$\langle [V(t) - \bar{V}_0]^n \rangle = 0, \qquad n \text{ not a multiple of } 3$$

$$\langle [V(t) - V_0]^{3m} \rangle = [(t - t_0)/6]^m (3m)!/m! . \tag{7.7.144}$$

Further, we assume the process (7.7.141) is some kind of generalized Markov process, for which the joint probability distribution is given by

$$p(v_2 t_2 : v_1 t_1) = p(v_2 t_2 | v_1 t_1) p(v_1, t_1) \tag{7.7.145}$$

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

$$\left.\begin{aligned}
&\lim_{t \to t_0} [\langle y(t) - y(t_0) \rangle / (t - t_0)] = a[y(t_0)] \\
&\lim_{t \to t_0} [\langle [y(t) - y(t_0)]^2 \rangle / (t - t_0)] = b[y(t_0)]^2 \\
&\lim_{t \to t_0} [\langle [y(t) - y(t_0)]^3 \rangle / (t - t_0)] = c[y(t_0)]^3
\end{aligned}\right\} \tag{7.7.151}$$

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

$$A + 2X \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3X \tag{7.7.156}$$

$$A \underset{k_4}{\overset{k_3}{\rightleftharpoons}} X$$
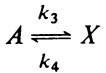
whose Poisson representation FPE is

$$\frac{\partial f(\alpha, t)}{\partial t} = -\frac{\partial}{\partial \alpha}[\kappa_1 V^{-1}\alpha^2 - \kappa_2 V^{-2}\alpha^3 + \kappa_3 V - \kappa_4\alpha)f(\alpha, t)]$$

$$+ \frac{1}{2}\frac{\partial^2}{\partial\alpha^2}[4(\kappa_1 V^{-1}\alpha^2 - \kappa_2 V^{-2}\alpha^3)f(\alpha, t)]$$

$$- \frac{1}{3!}\frac{\partial^3}{\partial\alpha^3}[6(\kappa_1 V^{-1}\alpha^2 - \kappa_2 V^{-2}\alpha^3)f(\alpha, t)], \tag{7.7.157}$$

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

$$
\begin{aligned}
d\eta(t)/dt = {} & \kappa_1\eta(t)^2 - \kappa_2\eta(t)^3 + \kappa_3 - \kappa_4\eta(t) \\
& + \mu^3\{4[\kappa_1\eta(t)^2 - \kappa_2\eta(t)^3]\}^{1/2}\xi(t) \\
& + \mu^4\{6[\kappa_1\eta(t)^2 - \kappa_2\eta(t)^3]\}^{1/3}\zeta(t) ,
\end{aligned}
\tag{7.7.158}
$$

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

$$
\langle x \rangle = V\eta_0 + \langle \eta_6 \rangle + \cdots = V\eta_0 + \frac{2ab}{c^2} + \ldots \tag{7.7.160a}
$$

$$
\begin{aligned}
\langle x^2 \rangle - \langle x \rangle^2 &= V\langle \eta_3^2 \rangle + [2\langle \eta_9\eta_3 \rangle + 2\langle \eta_8\eta_4 \rangle + \langle \eta_6^2 \rangle - \langle \eta_6 \rangle^2 + \langle \eta_6 \rangle] + \ldots \\
&= V\left[\frac{2a}{c}\right] + \left[\frac{28}{3}\frac{a^2b^2}{c^4} + \frac{8ab^2\eta_0}{c^3} - \frac{36\kappa_2 a^2}{c^3} + \frac{8ab}{c^2}\right] + \ldots \tag{7.7.160b}
\end{aligned}
$$

$$
\begin{aligned}
\langle (x - \langle x \rangle)^3 \rangle &= V[\langle \eta_4^3 \rangle - 3\langle \eta_3^2 \rangle\langle \eta_6 \rangle + 3\langle \eta_3^2 \rangle + \langle \eta_0 \rangle + \ldots \\
&= V\left[\frac{8a}{c} - \frac{12a^2b}{c^3} + \eta_0\right] + \ldots , \tag{7.7.160c}
\end{aligned}
$$

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* $j(r, t)$ exists such that

$$j(r, t) = -D\nabla \rho(r, t) \tag{8.1.1}$$

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 S, of V. Thus, if N is the total amount in V,

$$\frac{dN}{dt} = \frac{d}{dt} \int_V d^3r \; \rho(r, t) = - \int_S dS \cdot j(r, t) \tag{8.1.2}$$
$$= - \int_V d^3r \; \nabla \cdot j(r, t) \,.$$

Hence, since V is arbitrary,

$$\partial_t \rho(r, t) + \nabla \cdot j(r, t) = 0 \,. \tag{8.1.3}$$

Substituting Fick's law (8.1.1) into the conservation equation (8.1.3) we get

$$\partial_t \rho(r, t) = D\nabla^2 \rho(r, t) \,, \tag{8.1.4}$$

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

$$j(r, t) = -D\nabla \rho(r, t) + f_d(r, t) \,. \tag{8.1.5}$$

Here, $f_d(r, t)$ is a vector Langevin source. The simplest assumption to make concerning its stochastic properties is

$$\langle f_d(r, t) \rangle = 0$$

and

$$\langle f_{d,i}(r, t) f_{d,j}(r', t') \rangle = K_d(r, t) \delta_{ij} \delta(r - r') \delta(t - t') \,, \tag{8.1.6}$$

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

$$\partial_t \rho(r, t) = D\nabla^2 \rho(r, t) - \nabla \cdot f_d(r, t) \,. \tag{8.1.7}$$

Notice that

$$\langle \nabla \cdot f_d(r, t) \nabla' \cdot f_d(r', t') \rangle = \nabla \cdot \nabla'[K_d(r, t) \delta(r - r')] \delta(t - t') \,. \tag{8.1.8}$$

Now consider including a chemical reaction. Fick's law still applies, but instead of the conservation equation we need an equation of the form

$$\frac{dN}{dt} = \frac{d}{dt} \int_V d^3r \; \rho(r, t) = - \int_S dS \cdot j(r, t) + \int_V d^3r \; F[\rho(r, t)] \,, \tag{8.1.9}$$

where $F[\rho(r, t)]$ 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,

$$\partial_t \rho(r, t) + \nabla \cdot j(r, t) = F[\rho(r, t)] \,. \tag{8.1.10}$$

The production of the chemical by a chemical reaction does generate fluctuations, so we can add to (8.1.10) a term $f_c(r, t)$ which satisfies

$$\langle f_c(r, t) \rangle = 0$$
$$\langle f_c(r, t) f_c(r', t') \rangle = K_c(r, t) \delta(r - r') \delta(t - t') \tag{8.1.11}$$

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

$$\boxed{\partial_t \rho(r, t) = D\nabla^2 \rho(r, t) + F[\rho(r, t)] + g(r, t)} \tag{8.1.12}$$

where

$$g(r, t) = -\nabla \cdot f_d(r, t) + f_c(r, t) \tag{8.1.13}$$

and

$$\boxed{\langle g(r, t) g(r', t') \rangle = \{K_c(r-r', t) \delta(r-r') + \nabla \cdot \nabla'[K_d(r, t) \delta(r-r')]\} \delta(t-t').} \tag{8.1.14}$$

The simplest procedure for turning a classical reaction diffusion equation into a Langevin equation yields a rather complex expression Further, we know nothing about $K_c(r)$ or $K_d(r)$, 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 $\rho(r)$, where 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 r_i, and introduce the variables

$$x_i = l^3 \rho(r_i) \tag{8.1.15}$$

and consider functions of the variables $x = \{x_i\}$.

We now consider calculus of functions $F(x)$ of all these cell variables. Partial derivatives are easily defined in the usual way and we formally introduce the functional derivative by

$$\frac{\delta F(\rho)}{\delta \rho(r_i)} = \lim_{l \to 0} l^{-3} \frac{\partial F(x)}{\partial x_i}. \tag{8.1.16}$$

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

$$F[\rho(\mathbf{r}_i, t)] = \lim_{l \to 0} \tilde{F}(x_i)l^{-3} \tag{8.1.18}$$

$$g(\mathbf{r}_i, t) = \lim_{l \to 0} l^{-3} \sum_j \tilde{g}_{ij} dW_j(t) \,. \tag{8.1.19}$$

More precisely, we assume a more general correlation formula than (8.1.14), i.e.,

$$\langle g(\mathbf{r}, t)g(\mathbf{r}', t) \rangle = G(\mathbf{r}, \mathbf{r}')\delta(t - t') \,, \tag{8.1.20}$$

and require

$$G(\mathbf{r}_i, \mathbf{r}_j) = \lim_{l \to 0} l^{-6} \sum_k \tilde{g}_{ik}\tilde{g}_{jk} \,. \tag{8.1.21}$$

In this case, the FPE for x_i is

$$\partial_t P(x) = -\sum_{ij} \frac{\partial}{\partial x_i} \{[D_{ij}x_j + \delta_{ij}\tilde{F}(x_i)]P(x)\} + \sum_{ijk} \frac{1}{2} \frac{\partial^2}{\partial x_i \partial x_j} \tilde{g}_{ik}\tilde{g}_{jk}P(x) \,. \tag{8.1.22}$$

Now consider the limit $l^3 \to 0$. Some manipulation gives

$$\partial_t P(\rho) = -\int d^3r \frac{\delta}{\delta\rho(\mathbf{r})} \{[D\nabla^2(\mathbf{r}) + F[\rho(\mathbf{r})]]P(\rho)\}$$

$$+ \tfrac{1}{2} \iint d^3r \, d^3r' \left[\frac{\delta^2}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} G(\mathbf{r}, \mathbf{r}')P(\rho)\right]. \tag{8.1.23}$$

$P(\rho)$ is now a kind of functional probability and the definition of its normalisation requires a careful statement of the probability measure on $\rho(\mathbf{r})$. 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:

$$
\begin{aligned}
&\qquad\qquad\qquad\quad i \qquad\quad j \\
\boldsymbol{N}^{(i,j)} &= (0 \ldots 0, \quad 1, 0, \ldots 0, 0, 0, \ldots) \\
\boldsymbol{M}^{(i,j)} &= (0 \ldots 0, \quad 0, 0, \ldots 0, 1, 0, \ldots) \\
\boldsymbol{r}^{(i,j)} &= (0 \ldots 0, -1, 0, \ldots 0, 1, 0, \ldots) \\
k^+_{(i,j)} &= d_{ij} \\
k^-_{(i,j)} &= 0 .
\end{aligned}
\tag{8.2.2}
$$

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

$$d_t \langle x_i(t) \rangle = \sum_{j,k} r_i^{(j,k)} [\langle t_{jk}^+(\mathbf{x}) \rangle - t_{jk}^-(\mathbf{x}) \rangle]$$

$$= \sum_{j,k} (-\delta_{ij} + \delta_{ik}) \, d_{jk} \langle x_j \rangle .$$

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 $\mathbf{r}_i$ and we make the replacement

$$x_i(t) = l^3 \rho(\mathbf{r}_i, t) \tag{8.2.9}$$

and assume that $d_{ij} = 0$ (i, j not nearest neighbours)

$\qquad\qquad\qquad = d$ (i, j adjacent).

Then (8.2.8) becomes, in the limit $l \to 0$,

$$\partial_t \langle \rho(\mathbf{r}, t) \rangle = D \nabla^2 \langle \rho(\mathbf{r}, t) \rangle \quad \text{with} \tag{8.2.10}$$

$$D = l^2 d . \tag{8.2.11}$$

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,

$$t_{i,j}^+(\mathbf{x}) \to \infty \tag{8.2.12}$$

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

$$A_l(x) = \sum_j D_{jl} x_j \tag{8.2.13}$$

$$B_{lm}(x) = \delta_{lm} \sum_j (D_{lj} x_l + D_{jl} x_j) - D_{lm} x_l - D_{ml} x_m , \tag{8.2.14}$$

where

$$D_{jl} = d_{jl} - \delta_{jl} \sum_k d_{lk} \tag{8.2.15}$$

and thus, in this limit, $P(x, t)$ obeys the Fokker-Planck equation

$$\partial_t P = -\sum_l \partial_l A_l(x) P + \tfrac{1}{2} \sum_{l,m} \partial_l \partial_m B_{lm}(x) P . \tag{8.2.16}$$

b) Continuum Form of Kramers-Moyal Expansion

The continuum form is introduced by associating a point r with a cell i and writing

$$l^3 \sum_i \rightarrow \int d^3 r \tag{8.2.17}$$

$$D_{ji} \rightarrow D(r', r) \equiv \mathscr{D}(r', r - r') \tag{8.2.18}$$

$$l^{-3} \delta_{ij} \rightarrow \delta(r - r') . \tag{8.2.19}$$

At this stage we make no particular symmetry assumptions on D_{lj}, 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 r,

$$\int d^3\boldsymbol{\delta} \ \mathscr{D}(\boldsymbol{r} + \boldsymbol{\delta}, -\boldsymbol{\delta}) = 0 \qquad (8.2.24)$$

and from (8.2.23), we also have

$$\int d^3\boldsymbol{\delta} \ \mathscr{D}(\boldsymbol{r}, \boldsymbol{\delta}) = 0 \ . \qquad (8.2.25)$$

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.,

$$D_{ij} = D_{ji} \qquad (8.2.26)$$

which gives in the continuum form

$$\mathscr{D}(\boldsymbol{r} + \boldsymbol{\delta}, -\boldsymbol{\delta}) = \mathscr{D}(\boldsymbol{r}, \boldsymbol{\delta}) \ . \qquad (8.2.27)$$

The derivation of a continuum form now follows in a similar way to that of the Kramers-Moyal expansion.

We define the derivate moments

$$\boldsymbol{M}(\boldsymbol{r}) = \int d^3\boldsymbol{\delta} \ \boldsymbol{\delta}\mathscr{D}(\boldsymbol{r}, \boldsymbol{\delta}) \qquad (8.2.28)$$

$$\underline{\boldsymbol{D}}(\boldsymbol{r}) = \tfrac{1}{2} \int d^3\boldsymbol{\delta} \ \boldsymbol{\delta}\boldsymbol{\delta}\mathscr{D}(\boldsymbol{r}, \boldsymbol{\delta}) \qquad (8.2.29)$$

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

$$\begin{aligned}
\boldsymbol{M}(\boldsymbol{r}) &= \int d^3\boldsymbol{\delta} \ \boldsymbol{\delta}\mathscr{D}(\boldsymbol{r} + \boldsymbol{\delta}, -\boldsymbol{\delta}) \\
&= \int d^3\boldsymbol{\delta} \ \boldsymbol{\delta}[\mathscr{D}(\boldsymbol{r}, -\boldsymbol{\delta}) + \boldsymbol{\delta}\cdot\nabla\mathscr{D}(\boldsymbol{r}, -\boldsymbol{\delta}) + ...] \qquad (8.2.30) \\
&= -\boldsymbol{M}(\boldsymbol{r}) + 2\nabla\cdot\underline{\boldsymbol{D}}(\boldsymbol{r}) + ... \ .
\end{aligned}$$

Hence, detailed balance requires

$$\boldsymbol{M}(\boldsymbol{r}) = \nabla\cdot\underline{\boldsymbol{D}}(\boldsymbol{r}) \ . \qquad (8.2.31)$$

The weaker requirement (8.2.24) similarly requires the weaker condition

$$\nabla\cdot[\boldsymbol{M}(\boldsymbol{r}) - \nabla\cdot\underline{\boldsymbol{D}}(\boldsymbol{r})] = 0 \ . \qquad (8.2.32)$$

We now can make the continuum form of $A_i(\boldsymbol{x})$:

$$\begin{aligned}
A_i(\boldsymbol{x}) &\rightarrow \int d^3\boldsymbol{\delta} \ \mathscr{D}(\boldsymbol{r}, \boldsymbol{\delta})\rho(\boldsymbol{r} + \boldsymbol{\delta}) \\
&= \boldsymbol{M}(\boldsymbol{r})\cdot\nabla\rho(\boldsymbol{r}) + \underline{\boldsymbol{D}}(\boldsymbol{r}):\nabla\nabla\rho(\boldsymbol{r}) \qquad (8.2.33)
\end{aligned}$$

If detailed balance is true, we can rewrite, from (8.2.31)

$$\boxed{A_i(\boldsymbol{x}) \rightarrow \nabla\cdot[\underline{\boldsymbol{D}}(\boldsymbol{r})\cdot\nabla\rho(\boldsymbol{r})] \ .} \qquad (8.2.34)$$

The *general* form, without detailed balance, can be obtained by defining

$$J(r) = M(r) - \nabla \cdot \underline{D}(r) \qquad (8.2.35)$$

from (8.2.32)

$$\nabla \cdot J(r) = 0 \qquad (8.2.36)$$

so that we can write

$$J(r) = \nabla \cdot \underline{E}(r) , \qquad (8.2.37)$$

where $\underline{E}(r)$ is an antisymmetric tensor. Substituting, we find that by defining

$$\underline{H}(r) = \underline{D}(r) + \underline{E}(r) , \qquad (8.2.38)$$

we have defined a nonsymmetric diffusion tensor $\underline{H}(r)$ and that

$$A_l(x) \rightarrow \nabla \cdot [\underline{H}(r) \cdot \nabla \rho(r, t)] . \qquad (8.2.39)$$

This means that, deterministically,

$$\boxed{\partial_t \rho(r, t) = \nabla \cdot [\underline{H}(r) \cdot \nabla \rho(r, t)]} \qquad (8.2.40)$$

where $\underline{H}(r)$ is symmetric, if detailed balance holds.

We now come to the fluctuation term, given by (8.2.14). To compute $B_{lm}(x)$, we first consider the limit of

$$l^3 \sum_m B_{lm} \phi_m \rightarrow \int dr' B(r, r') \phi(r') \qquad (8.2.41)$$

where ϕ_m is an arbitrary function. By similar, but much more tedious computation, we eventually find

$$l^3 \sum_m B_{lm} \phi_m \rightarrow -2\nabla \cdot [\underline{D}(r)\rho(r) \cdot \nabla \phi(r)] \qquad (8.2.42)$$

so that

$$\boxed{B(r, r') = 2\nabla'\nabla : [\underline{D}(r)\rho(r)\delta(r - r')] .} \qquad (8.2.43)$$

The phenomenological theory of Sect. 8.1 now has a rational basis since (8.2.43) is what arises from assuming

$$j(r, t) = -\underline{H}(r)\nabla\rho(r, t) - \xi(r, t) \tag{8.2.44}$$

in which

$$\langle\xi(r, t)\xi(r', t')\rangle = 2\delta(t - t')\underline{D}(r)\delta(r - r')\rho(r) \tag{8.2.45}$$

and hence,

$$\partial_t\rho(r, t) = \nabla\cdot\underline{H}(r)\cdot\nabla\rho + \nabla\cdot\xi(r, t). \tag{8.2.46}$$

This corresponds to a theory of inhomogeneous anisotropic diffusion without detailed balance. This is usually simplified by setting

$$\underline{D}(r) = D\mathbf{1} \tag{8.2.47}$$

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 $\underline{D}$ 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

$$g(r, r') = \langle\rho(r), \rho(r')\rangle = \delta(r - r')\langle\rho(r)\rangle.$$

Since the theory is linear, we can apply (4.4.51) of Sect. 4.4.6. Here the matrices A and A^T become

$$A \rightarrow -\nabla\cdot\underline{H}(r)\cdot\nabla$$
$$A^T \rightarrow -\nabla'\cdot\underline{H}(r')\cdot\nabla'. \tag{8.2.49}$$

Thus,

$$B(r, r') \rightarrow BB^T$$
$$= A\sigma + \sigma A^T \rightarrow [-\nabla\cdot\underline{H}(r)\cdot\nabla - \nabla'\cdot\underline{H}(r')\cdot\nabla']g(r, r'). \tag{8.2.50}$$

We note that in the stationary state, $\langle\rho(r)\rangle = \langle\rho\rangle$, independent of r. Thus,

$$B(r, r') = [-\nabla\cdot\underline{H}(r)\cdot\nabla\langle\rho\rangle\delta(r - r') - \nabla'\cdot\underline{H}(r')\cdot\nabla'\langle\rho\rangle\delta(r - r')]$$
$$= \nabla\nabla': \{[\underline{H}(r) + \underline{H}^T(r)]\langle\rho\rangle\delta(r - r')\}$$
$$= 2\nabla\nabla': [\underline{D}(r)\langle\rho\rangle\delta(r - r')]. \tag{8.2.51}$$

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

$\rho(r)$. 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

$$X = (x_1, x_2, \ldots) \tag{8.2.52}$$

where 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

$$\partial_t P(X, t) = \partial_t P(\text{diffusion}) + \sum_i \{\sum_A t_A^+(x_i - r^A)P(x_i - r^A, \hat{X})$$
$$+ t_A^-(x_i + r^A)P(x_i + r^A, \hat{X}) - [t_A^+(x_i) + t_A^-(x_i)]P(X)\} \tag{8.2.53}$$

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 $W(r, t)$ as follows: we consider an isotropic constant diffusion tensor

$$\underline{D}(r) = D\mathbf{1} \tag{8.2.54}$$

$$d\rho_a = [D_a \nabla^2 \rho_a + \sum_A r_a^A (\kappa_A^+ \prod_a \rho_a^{N_a^A} - \kappa_A^- \prod_a \rho_a^{M_a^A})]dt + dW_a(r, t) \tag{8.2.55}$$

with

$$dW_a(r, t)dW_b(r', t) = \{2\delta_{a,b}\nabla' \cdot \nabla[D_a\rho_a(r)\delta(r - r')]$$
$$+ \delta(r - r') \sum_A r_a^A r_b^A (\kappa_A^+ \prod_a \rho_a^{N_a^A} + \kappa_A^- \prod_a \rho_a^{M_a^A})\} dt . \tag{8.2.56}$$

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^2d 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)

$$N = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad M = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad r = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$k^+ = k_1 = \kappa_1 \Omega^{-1+1}, \qquad k^- = k_2 = \kappa_1 \Omega^{-2+1} .$$

(8.2.57)

Thus, substituting

$$dp_1(r) = (D_1 \nabla^2 \rho_1 - \kappa_1 \rho_1 + \kappa_2 \rho_2^2) dt + dW_1(r, t)$$
$$dp_2(r) = (D_2 \nabla^2 \rho_2 + 2\kappa_1 \rho_1 - 2\kappa_2 \rho_2^2) dt + dW_2(r, t)$$

(8.2.58)

where $dW(r, t) dW^T(r', t) =$

$$\begin{bmatrix} 2\nabla \cdot \nabla'[D_1 \rho_1 \delta(r-r')] + \delta(r-r')[\kappa_1 \rho_1 + \kappa_2 \rho_2^2], & -2\delta(r-r')[\kappa_1 \rho_1 + \kappa_2 \rho_2^2] \\ -2\delta(r-r')[\kappa_1 \rho_1 + \kappa_2 \rho_2^2], & 2\nabla \cdot \nabla'[D_2 \rho_2 \delta(r-r')] + 4\delta(r-r')[\kappa_1 \rho_1 + \kappa_2 \rho_2] \end{bmatrix} dt .$$

(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 thé 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

$$\eta_a(r) = \alpha_a(r)/l^3$$

(8.2.60)

and we find

$$d\eta_a(r) = [D_a \nabla^2 \eta_a(r) + \sum_A r_a^A (\kappa_A^+ \prod_a \eta_a^{N_a^A} - \kappa_A^- \prod_a \eta_a^{M_a^A})] dt + dW_a(r, t)$$

(8.2.61)

$$dW_a(r, t) dW_b(r', t) = dt\, \delta(r-r') \sum_A (\kappa_A^+ \prod_a \eta_a^{N_a^A} - \kappa_A^- \prod_a \eta_a^{M_a^A})$$
$$\times (M_a^A M_b^A - N_a^A N_b^A - \delta_{ab} r_a^A) .$$

(8.2.62)

These equations are very similar to (8.2.55,56). When explicitly written out they are simpler. For example, considering again $X_1 \rightleftharpoons 2\,X_2$, we get

$$d\eta_1(r) = (D_1 \nabla^2 \eta_1 - \kappa_1 \eta_1 + \kappa_2 \eta_2^2)dt + dW_1(r, t) \tag{8.2.63}$$

$$d\eta_2(r) = (D_2 \nabla^2 \eta_2 + 2\kappa_1 \eta_1 - 2\kappa_2 \eta_2^2)dt + dW_2(r, t) \tag{8.2.64}$$

$$dW(r, t)dW^{\mathsf{T}}(r', t) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} (\kappa_1 \eta_1 - \kappa_2 \eta_2^2)\delta(r - r')dt . \tag{8.2.65}$$

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.,

$$\partial_t \eta(r, t) = D\nabla^2 \eta - k_1 \eta + k_2 \mu$$
$$\partial_t \mu(r, t) = D\nabla^2 \mu + k_1 \eta - k_2 \mu . \tag{8.3.1}$$

a) Spatial Correlations
We now note that

$$\langle \eta(r, t) \rangle = \langle \rho_x(r, t) \rangle$$
$$\langle \mu(r, t) \rangle = \langle \rho_y(r, t) \rangle \tag{8.3.2}$$
$$\langle \eta(r,t),\, \eta(r', t) \rangle = \langle \rho_x(r,t),\, \rho_x(r', t) \rangle - \delta(r - r')\langle \rho_x(r, t) \rangle \equiv g(r, r', t)$$
$$\langle \eta(r,t),\, \mu(r', t) \rangle = \langle \rho_x(r,t),\, \rho_y(r', t) \rangle \equiv f(r, r', t)$$
$$\langle \mu(r,t),\, \mu(r', t) \rangle = \langle \rho_y(r,t),\, \rho_x(r', t) \rangle - \delta(r - r')\langle \rho_y(r, t) \rangle = h(r, r', t) ,$$

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

$$\langle \rho_x(\mathbf{r}, t) \rangle = \langle \rho_x(t) \rangle$$
$$\langle \rho_y(\mathbf{r}, t) \rangle = \langle \rho_y(t) \rangle$$

$$g(\mathbf{r}, \mathbf{r}', t) = g(\mathbf{r} - \mathbf{r}', t)$$
$$f(\mathbf{r}, \mathbf{r}', t) = f(\mathbf{r} - \mathbf{r}', t) \qquad\qquad (8.3.3)$$
$$h(\mathbf{r}, \mathbf{r}', t) = h(\mathbf{r} - \mathbf{r}', t)$$

and compute equations of motion for $\langle \eta(\mathbf{r}, t)\eta(0, t) \rangle$, etc, we quickly find

$$\frac{\partial g(\mathbf{r}, t)}{\partial t} = 2D\nabla^2 g(\mathbf{r}, t) - 2k_1 g(\mathbf{r}, t) + 2k_2 f(\mathbf{r}, t)$$

$$\frac{\partial f(\mathbf{r}, t)}{\partial t} = 2D\nabla^2 f(\mathbf{r}, t) - (k_1 + k_2)f(\mathbf{r}, t) + k_2 h(\mathbf{r}, t) + k_1 g(\mathbf{r}, t) \qquad (8.3.4)$$

$$\frac{\partial h(\mathbf{r}, t)}{\partial t} = 2D\nabla^2 h(\mathbf{r}, t) - 2k_2 h(\mathbf{r}, t) + 2k_1 f(\mathbf{r}, t) ,$$

The stationary solution of these equations has the form

$$g(\mathbf{r}) = \xi k_2^2, \qquad f(\mathbf{r}) = \xi k_1 k_2, \qquad h(\mathbf{r}) = \xi k_1^2 \qquad (8.3.5)$$

where ξ is an arbitrary parameter. The corresponding stationary solutions for the means are

$$\langle x(\mathbf{r}_i) \rangle = \lambda k_2, \qquad \langle y(\mathbf{r}_i) \rangle = \lambda k_1 \qquad (8.3.6)$$

where λ is another arbitrary parameter. If $\xi = 0$, we recover the Poissonian situation where

$$\langle \rho_x(\mathbf{r}), \rho_x(\mathbf{r}') \rangle = \langle \rho_x(\mathbf{r}) \rangle \, \delta(\mathbf{r} - \mathbf{r}')$$
$$\langle \rho_y(\mathbf{r}), \rho_y(\mathbf{r}') \rangle = \langle \rho_y(\mathbf{r}) \rangle \, \delta(\mathbf{r} - \mathbf{r}') \qquad (8.3.7)$$
$$\langle \rho_x(\mathbf{r}), \rho_y(\mathbf{r}') \rangle = 0 .$$

(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

$$g(\mathbf{r}, 0) = \alpha \, \delta(\mathbf{r}), \qquad f(\mathbf{r}, 0) = \beta \, \delta(\mathbf{r}), \qquad h(\mathbf{r}, 0) = \gamma \, \delta(\mathbf{r}). \qquad (8.3.8)$$

Time-dependent solutions are

$$
\begin{bmatrix} g(r,\,t) \\ f(r,\,t) \\ h(r,\,t) \end{bmatrix} = \frac{\exp(-r^2/8Dt)}{(8\pi Dt)^{3/2}} \begin{bmatrix} k_2^2\varepsilon_1 - 2k_2\varepsilon_2 e^{-(k_1+k_2)t} + \varepsilon_3 e^{-2(k_1+k_2)t} \\ k_1 k_2 \varepsilon_1 + (k_2 - k_1)\varepsilon_2 e^{-(k_1+2)t} - \varepsilon_3 e^{-2(k_1+k_2)t} \\ k_1^2 \varepsilon_1 + 2k_1\varepsilon_2 e^{-(k_1+k_1)t} + \varepsilon_3 e^{-2(k_1+k_2)t} \end{bmatrix} \tag{8.3.9}
$$

where

$$
\begin{aligned}
\varepsilon_1 &= (\alpha + 2\beta + \gamma)/(k_1 + k_2)^2 \\
\varepsilon_2 &= [k_2(\beta + \gamma) - k_1(\alpha + \beta)]/(k_1 + k_2)^2 \tag{8.3.10}\\
\varepsilon_3 &= [k_1^2 \alpha + k_2^2 \gamma - 2k_1 k_2 \beta]/(k_1 + k_2)^2 \,.
\end{aligned}
$$

Comments

i) The terms ε_1, ε_2, and ε_3 correspond, respectively, to deviations from an uncorrelated Poissonian of the quantities $\langle (x_i + y_i),\, (x_j + y_j) \rangle$, $\langle (x_i + y_i),\, (k_1 y_j - k_2 x_j) \rangle$, and $\langle (k_1 y_i - k_2 x_i),\, (k_1 y_j - k_2 x_j) \rangle$, 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

$$
\text{var}\{x[V,\,0]\} = \int_V d^3r \int_V d^3r' \langle \rho_x(r,\,0),\, \rho_x(r',\,0) \rangle, \qquad \text{i.e.,}
$$

$$
\text{var}\{x[V,\,0]\} = \langle x[V,\,0] \rangle + \alpha V \tag{8.3.11}
$$

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

$$
\left.
\begin{aligned}
\text{var}\{x[V,\,t]\} &\simeq \langle x[V,\,t] \rangle \\
&\quad + \frac{V^2}{(8\pi Dt)^{3/2}}(k_1^2\varepsilon_1 - 2k_2\varepsilon_2 e^{-(k_1+k_2)t} + \varepsilon_3 e^{-2(k_1+k_2)t}) \\
\text{var}\{y[V,\,t]\} &\simeq \langle y[V,\,t] \rangle \\
&\quad + \frac{V^2}{(8\pi Dt)^{3/2}}(k_1^2\varepsilon_1 + 2k_1\varepsilon_2 e^{-(k_1+k_2)t} + \varepsilon_3 e^{-2(k_1+k_2)t}) \\
\langle x[V,\,t],\, y[V,\,t] \rangle &\simeq \frac{V^2}{(8\pi Dt)^{3/2}} \\
&\quad \times [k_1 k_2 \varepsilon_1 + (k_2 - k_1)\varepsilon_2 e^{-(k_1+k_2)t} - \varepsilon_3 e^{-2(k_1+k_2)t}] \,.
\end{aligned}
\right\} \tag{8.3.12}
$$

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 $\underline{G}(r, t)$ by

$$\underline{G}(r, t) = \begin{bmatrix} \langle \rho_x(r, t), \rho_x(0, 0) \rangle_s & \langle \rho_y(r, t), \rho_x(0, 0) \rangle \\ \langle \rho_x(r, t), \rho_y(0, 0) \rangle_s & \langle \rho_y(r, t), \rho_y(0, 0) \rangle \end{bmatrix}. \tag{8.3.13}$$

Then the equation corresponding to (3.7.63) is

$$\partial_t \underline{G}(r, t) = \begin{bmatrix} D\nabla^2 - k_1 & k_2 \\ k_1 & D\nabla^2 - k_2 \end{bmatrix} \underline{G}(r, t). \tag{8.3.14}$$

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

$$\underline{G}(r, t) = \left(\frac{\langle \rho_x \rangle + \langle \rho_y \rangle}{(k_1 + k_2)^2} \right) \frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}}$$
$$\begin{bmatrix} k_2^2 + k_1 k_2 e^{-(k_1+k_2)t} & k_1 k_2(1 - e^{-(k_1+k_2)t}) \\ k_1 k_2(1 - e^{-(k_1+k_2)t}) & k_1^2 + k_1 k_2 e^{-(k_1+k_2)t} \end{bmatrix}. \tag{8.3.15}$$

If we define variables

$$n(r, t) = \rho_x(r, t) + \rho_y(r, t) \tag{8.3.16}$$

$$c(r, t) = [k_1 \rho_x(r, t) - k_2 \rho_y(r, t)]/(k_1 + k_2), \tag{8.3.17}$$

the solution (8.3.15) can be written as

$$\langle n(r, t), n(0, 0) \rangle_s = \langle n \rangle_s \frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}} \tag{8.3.18}$$

$$\langle n(r, t), c(0, 0) \rangle_s = \langle c(r, t), n(0, 0) \rangle_s = 0 \tag{8.3.19}$$

$$\langle c(r, t), c(0, 0) \rangle_s = \frac{k_1 k_2 \langle n \rangle_s}{(k_1 + k_2)^2} \frac{\exp(-r^2/4Dt)}{(4\pi Dt)^{3/2}} e^{-(k_1+k_2)t}. \tag{8.3.20}$$

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 $B + X \underset{k_3}{\overset{k_1}{\rightleftharpoons}} C, \quad A + X \xrightarrow{k_2} 2X$
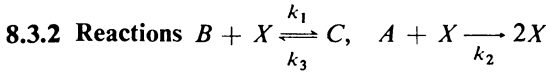
This reaction has already been treated in Sect. 7.6.4a without spatial dependence. We find the Poisson representation equation for the concentration variable $\eta(r, t)$ is [from (8.2.60–62), (7.6.55,56)]

$$d\eta(r, t) = [D\nabla^2\eta(r, t) + (K_2 - K_1)\eta(r, t) + K_3]dt + dW(r, t) \tag{8.3.21}$$

$$dW(r, t)dW(r', t) = 2dt\,\delta(r - r')K_2\eta(r, t), \tag{8.3.22}$$

where

$$\left.\begin{array}{l} k_3 C = K_3 l^3 \\ k_1 B = K_1 \\ k_2 A = K_2 . \end{array}\right\} \tag{8.2.23}$$

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 $\rho(r, t)$, but the $dW(r, t)$ would satisfy

$$\begin{aligned} dW(r, t)\,dW(r', t) = dt\,\{&2\nabla' \cdot \nabla[D\rho(r)\delta(r-r')] \\ &+ \delta(r-r')[(K_2 + K_1)\eta(r, t) + K_3]\}] . \end{aligned} \tag{8.3.24}$$

a) Spatial Correlations
Define now

$$\begin{aligned} g(r, r', t) &= \langle\rho(r, t), \rho(r', t)\rangle - \delta(r - r')\langle\rho(r, t)\rangle \\ &= \langle\eta(r, t), \eta(r', t)\rangle = \langle\eta(r, t)\eta(r', t)\rangle - \langle\eta(r, t)\rangle\langle\eta(r', t)\rangle \end{aligned} \tag{8.3.25}$$

We consider the stationary homogeneous situation in which clearly

$$\langle\eta(r, t)\rangle_s = \langle\rho(r, t)\rangle_s = \langle\rho\rangle_s . \tag{8.3.26}$$

Then,

$$\begin{aligned} dg(r, r', t) &= d\langle\eta(r, t)\eta(r', t)\rangle \\ &= \langle\eta(r, t)d\eta(r', t)\rangle + \langle d\eta(r, t)\eta(r' t)\rangle + \langle d\eta(r, t)d\eta(r', t)\rangle \end{aligned} \tag{8.3.27}$$

and using the usual Ito rules and (8.3.21,22),

$$\begin{aligned} = [D\nabla^2 &+ D\nabla'^2 + 2(K_2 - K_1)]\langle\eta(r, t)\eta(r', t)\rangle \\ &+ 2K_3\langle\rho\rangle_s dt + 2K_2\delta(r - r')\langle\rho\rangle_s dt . \end{aligned} \tag{8.3.28}$$

Note that

$$\langle \rho \rangle_s = K_3/(K_1 - K_2) \,, \tag{8.3.29}$$

and that in a spatially homogeneous situation, $g(r, r', t)$ can only be a function of $r - r'$, which we call $g(r, t)$.

Substitute using (8.3.25,26) to obtain

$$\partial_t g(r, t) = 2[D\nabla^2 + (K_2 - K_1)]g(r, t) + 2K_2 \langle \rho \rangle_s \delta(r) \,. \tag{8.3.30}$$

The stationary solution $g_s(r)$ is best obtained by representing it as a Fourier integral:

$$g_s(r) = \int d^3q \; e^{-i\mathbf{q}\cdot r} \tilde{g}_s(q) \tag{8.3.31}$$

which, on using

$$\delta(r) = (2\pi)^{-3} \int d^3q \; e^{-i\mathbf{q}\cdot r} \,, \tag{8.3.32}$$

gives

$$\tilde{g}_s(q) = \frac{K_2 \langle \rho \rangle_s}{(2\pi)^3} \frac{1}{Dq^2 + K_1 - K_2} \tag{8.3.33}$$

whose inverse is given by

$$g_s(r) = \frac{K_2 \langle \rho \rangle_s}{4\pi Dr} \exp\left[-r\left(\frac{K_1 - K_2}{D}\right)^{1/2}\right]. \tag{8.3.34}$$

Hence,

$$\langle \rho(r, t), \rho(r', t) \rangle_s = \delta(r - r') \langle \rho \rangle_s$$
$$+ \frac{K_2 \langle \rho \rangle_s}{4\pi D |r - r'|} \exp\left[-|r - r'|\left[\frac{K_1 - K_2}{D}\right]^{1/2}\right]. \tag{8.3.35}$$

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:

$$x[V, t] = \int_V d^3r \; \rho(r, t) \,. \tag{8.3.37}$$

Then we would like to know what is the variance of this number. We can easily see

$$\text{var}\{x(V)\}_s = \int_V d^3r \int_V d^3r' \langle \rho(\mathbf{r}, t), \rho(\mathbf{r}', t) \rangle_s \tag{8.3.38}$$

and using (8.3.35),

$$= \langle x(V) \rangle_s + \frac{K_2 \langle \rho \rangle_s}{4\pi D} \int_V d^3r \int_V d^3r' |\mathbf{r} - \mathbf{r}'|^{-1} \exp\left(-|\mathbf{r} - \mathbf{r}'|/l_c\right) \tag{8.3.39}$$

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:

$$0 = 2(K_2 - K_1) \int_V d^3r \int_V d^3r' g(\mathbf{r} - \mathbf{r}') + 2K_2 \langle \rho \rangle_s \int_V d^3r' . \tag{8.3.40}$$

Thus,

$$\iint_{V\,V} d^3r \, d^3r' \, g(\mathbf{r} - \mathbf{r}') \sim \frac{K_2 \langle x(V) \rangle_s}{K_1 - K_2} \tag{8.3.41}$$

so that

$$\text{var}\{x(V)\}_s \sim \frac{K_1 \langle x(V) \rangle_s}{K_1 - K_2} \qquad (V \gg l_c^3) . \tag{8.3.42}$$

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

$$\iint_{VV} d^3r \, d^3r' |\mathbf{r} - \mathbf{r}'|^{-1} = \int_V d^3r' \int_0^R 4\pi r^2 dr \, d(\cos\theta)(r^2 + r'^2 - 2rr'\cos\theta)^{-1/2}$$

$$= \int_V d^3r' \int_0^R \frac{4\pi r}{r'} dr(|r + r'| - |r - r'|) \tag{8.3.43}$$

$$= 2R^5(4\pi)^2/5$$

so that

$$\text{var}\{x(V)\}_s \sim \langle x(V) \rangle_s \left(1 + \frac{2R^2 K_2}{5D}\right) \qquad \left(V = \frac{4}{3}\pi R^3 \ll l_c^3\right). \tag{8.3.44}$$

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

$$\text{var}\{x(V)\} = \langle x(V) \rangle \left\{1 + \frac{3K_2 l_c^5}{2DR^3}\left[1 - \left(\frac{R}{l_c}\right)^2 + \frac{2}{3}\left(\frac{R}{l_c}\right)^3 \right.\right.$$

$$\left.\left. - e^{-2R/l_c}\left(1 + \frac{R}{l_c}\right)^2\right]\right\} \tag{8.3.45}$$

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 $\langle \rho(r, t), \rho(r', t) \rangle$ is given by adding those of $g_s(r)$ and $\delta(r)\langle \rho \rangle_s$ 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 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

$$\langle \rho \rangle_s \frac{K_1}{K_1 - K_2} \delta(r - r') \tag{8.3.49}$$

which corresponds to the same variance as the global Master equation.
For large q, i.e., small wavelengths, we fine

$$\langle \rho \rangle_s (2\pi)^{-3} \tag{8.3.50}$$

corresponding to

$$\langle \rho \rangle_s \delta(r - r') , \tag{8.3.51}$$

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

$$G(r, t) = \langle \rho(r, t), \rho(0, 0) \rangle_s . \tag{8.3.52}$$

Then the linear equation corresponding to (3.7.63) is

$$\partial_t G(r, t) = D\nabla^2 G(r, t) - (K_1 - K_2)G(r, t) \tag{8.3.53}$$

with an initial condition given by

$$G(r, 0) = \langle \rho(r), \rho(0) \rangle_s \tag{8.3.54}$$

which is given by (8.3.35). Representing $G(r, t)$ as a Fourier integral

$$\tilde{G}(r, t) = \int d^3q \, e^{-i q \cdot r} \tilde{G}(q, t) , \tag{8.3.55}$$

(8.3.53) becomes

$$\partial_t \tilde{G}(q, t) = -(Dq^2 + K_1 - K_2)\tilde{G}(q, t) \tag{8.3.56}$$

whose solution is, utilising the Fourier representation (8.3.47) of the initial condition (8.3.54),

$$\tilde{G}(q, t) = (2\pi)^{-3} \frac{Dq^2 + K_1}{Dq^2 + K_1 - K_2} \exp[-(Dq^2 + K_1 - K_1)t] . \tag{8.3.57}$$

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

$$
\begin{aligned}
G(r, t) = & \frac{\langle \rho \rangle \exp(-r^2/4Dt - Dt/l_c^2)}{(4\pi Dt)^{3/2}} \\
& + \frac{K_2 \langle \rho \rangle}{8\pi Dr} \left\{ [\exp(-r/l_c)] \operatorname{erfc} \left[\frac{(Dt)^{1/2}}{l_c} - \frac{r}{2(Dt)^{1/2}} \right] \right. \\
& \left. + [\exp(r/l_c)] \operatorname{erfc} \left[\frac{(Dt)^{1/2}}{l_c} + \frac{r}{2(Dt)^{1/2}} \right] \right\} .
\end{aligned}
\tag{8.3.58}
$$

For small t we find

$$G(r, t) \rightarrow \langle \rho \rangle \exp\left(\frac{-r^2}{4Dt}\right) \left[(4\pi Dt)^{-3/2} - \frac{4K_2(Dt)^{3/2}}{r^3 \sqrt{\pi}} \right] + \frac{K_2 \langle \rho \rangle e^{-r/l_c}}{4\pi Dr} , \tag{8.3.59}$$

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 $K_1 \rightarrow K_2$, 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 $l_c \rightarrow \infty$ (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 $l_c \rightarrow \infty$ 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, $\tau_c \rightarrow \infty$ 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

$$A + X \overset{k_2}{\underset{k_4}{\rightleftharpoons}} 2X$$

$$B + X \overset{k_1}{\underset{k_4}{\rightleftharpoons}} C \tag{8.3.62}$$

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

$$\partial_t \rho(\mathbf{r}, t) = D\nabla^2 \rho + \kappa_3 + (\kappa_2 - \kappa_1)\rho - \kappa_4 \rho^2 \tag{8.3.63}$$

whose stationary solution is given by

$$\rho(\mathbf{r}) = \rho_s = [\kappa_2 - \kappa_1 + \sqrt{(\kappa_2 - \kappa_1)^2 + 4\kappa_4\kappa_3}]/2\kappa_4 . \tag{8.3.64}$$

[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
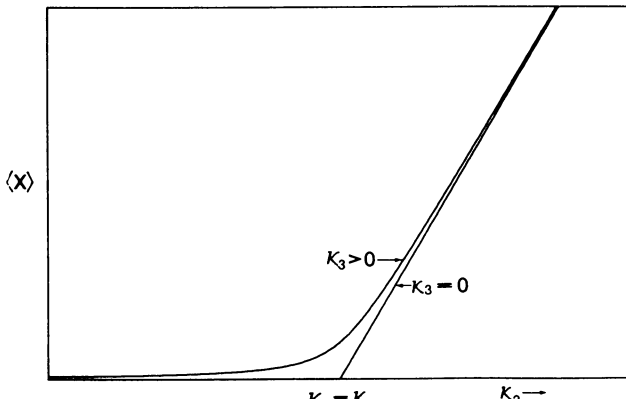
Fig. 8.1. Plot of $\langle x \rangle$ vs κ_2 for the second-order phase transition model

$$dη(\mathbf{r}, t) = [(\mathscr{D}η)(\mathbf{r}, t) + κ_3 + (κ_2 - κ_1)η(\mathbf{r}, t) - κ_4η^2(\mathbf{r}, t)]dt$$
$$+ \sqrt{2}[κ_2η(\mathbf{r}, t) - κ_4η^2(\mathbf{r}, t)]^{1/2}dW(\mathbf{r}, t), \tag{8.3.65}$$

where

$$dW(\mathbf{r}, t)dW(\mathbf{r}', t) = δ(\mathbf{r} - \mathbf{r}')dt. \tag{8.3.66}$$

Here we have *not* taken the continuum limit developed in Sect. 8.2.1 but have preserved a possible nonlocality of diffusion, i.e.,

$$(\mathscr{D}η)(\mathbf{r}, t) = -\int \mathscr{D}(|\mathbf{r} - \mathbf{r}'|)η(\mathbf{r}', t)d^3r' \tag{8.3.67}$$

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

$$dη(\mathbf{r}, t) = [(\mathscr{D}η)(\mathbf{r}, t) + κ_3 + (κ_2 - κ_1)η(\mathbf{r}, t) - κ_4η^2(\mathbf{r}, t)]dt$$
$$+ λ\sqrt{2}[κ_2η(\mathbf{r}, t) - κ_4η^2(\mathbf{r}, t)]^{1/2}dW(\mathbf{r}, t) \tag{8.3.68}$$

and expand $η(\mathbf{r}, t)$ in powers of $λ$ to

$$η(\mathbf{r}, t) = η_0(\mathbf{r}, t) + λη_1(\mathbf{r}, t) + λ^2η_2(\mathbf{r}, t) + \dots \tag{8.3.69}$$

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

$$\frac{dη_0(\mathbf{r}, t)}{dt} = (\mathscr{D}η_0)(\mathbf{r}, t) + κ_3 + (κ_2 - κ_1)η_0(\mathbf{r}, t) - κ_4η_0^2(\mathbf{r}, t) \tag{8.3.70}$$

$$dη_1(\mathbf{r}, t) = \{(\mathscr{D}η_1)(\mathbf{r}, t) + [κ_2 - κ_1 - 2κ_4η_0(\mathbf{r}, t)]η_1(\mathbf{r}, t)\}dt$$
$$+ \sqrt{2}[κ_2η_0(\mathbf{r}, t) - κ_4η_0^2(\mathbf{r}, t)]^{1/2}dW(\mathbf{r}, t) \tag{8.3.71}$$

$$dη_2(\mathbf{r}, t) = \{(\mathscr{D}η_2)(\mathbf{r}, t) + [κ_2 - κ_1 - 2κ_4η_0(\mathbf{r}, t)]η_2(\mathbf{r}, t) - κ_4η_1^2(\mathbf{r}, t)\}dt$$
$$+ \frac{[κ_2 - 2κ_4η_0(\mathbf{r}, t)]η_1(\mathbf{r}, t)}{\sqrt{2}[κ_2η_0(\mathbf{r}, t) - κ_4η_0^2(\mathbf{r}, t)]^{1/2}}dW(\mathbf{r}, t). \tag{8.3.72}$$

Equation (8.3.70) has a homogeneous steady-state solution (8.3.64) so that

$$κ_2 - κ_1 - 2κ_4η_0 \equiv κ = [(κ_1 - κ_2)^2 + 4κ_4κ_3]^{1/2}. \tag{8.3.73}$$

Substituting this in (8.3.70–72) and taking the Fourier transforms, we get

$$η_1(\mathbf{q}, t) = (2κ_1η_0)^{1/2}\int_0^t [\exp\{-[\mathscr{D}(q^2) + κ](t - t')\}]d\bar{W}(\mathbf{q}, t') \tag{8.3.74}$$

$$\bar{\eta}_2(\boldsymbol{q}, t) = -\kappa_4 \int d^3q_1 \int_0^t dt'[\exp\{-[\mathscr{D}(q^2) + \kappa](t - t')\}]\bar{\eta}_1(\boldsymbol{q} - \boldsymbol{q}_1, t')\bar{\eta}_1(\boldsymbol{q}_1, t')$$

(8.3.75)

$$+ \frac{\kappa_1 + \kappa}{[2(\kappa_1\eta_0 - \kappa_3]^{1/2}} \int d^3q_1 \int_0^t [\exp\{-[\mathscr{D}(q^2) + \kappa](t - t')\}]\bar{\eta}_1(\boldsymbol{q} - \boldsymbol{q}_1, t')d\tilde{W}(\boldsymbol{q}_1, t)$$

where

$$d\tilde{W}(\boldsymbol{q}, t) = \left(\frac{1}{2\pi}\right)^{3/2} \int \exp(-i\boldsymbol{q} \cdot \boldsymbol{r})dW(\boldsymbol{r}, t)d^3\boldsymbol{r}$$

(8.3.76)

and

$$\bar{\eta}_1(\boldsymbol{q}, t) = \left(\frac{1}{2\pi}\right)^{3/2} \int \exp(-i\boldsymbol{q} \cdot \boldsymbol{r})\eta_1(\boldsymbol{r}, t)d^3\boldsymbol{r} ,$$

(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)

$$\langle\rho(\boldsymbol{r}, t)\rangle = \eta_0$$

(8.3.78)

$$\langle\rho(\boldsymbol{r}, t)\rho(\boldsymbol{r}', t)\rangle - \langle\rho(\boldsymbol{r}, t)\rangle\langle\rho(\boldsymbol{r}', t)\rangle = \eta_0\delta(\boldsymbol{r} - \boldsymbol{r}')$$

$$+ \langle\eta_1(\boldsymbol{r}, t)\eta_1(\boldsymbol{r}', t)\rangle .$$

(8.3.79)

From (8.3.74) it follows that

$$\langle\bar{\eta}_1(\boldsymbol{q}, t)\bar{\eta}_1(\boldsymbol{q}', t)\rangle = \frac{\kappa_1\eta_0\delta(\boldsymbol{q} + \boldsymbol{q}')}{[\mathscr{D}(q^2) + \kappa]}[1 - \exp\{-2[\mathscr{D}(q^2) + \kappa]t\}] .$$

(8.3.80)

Hence, the lowest-order contributions to the correlation function in the steady state are given by

$$\langle\rho(\boldsymbol{r})\rho(\boldsymbol{r}') - \langle\rho(\boldsymbol{r})\rangle\langle\rho(\boldsymbol{r}')\rangle = \eta_0\delta(\boldsymbol{r} + \boldsymbol{r}')$$

$$+ \frac{\kappa_1\eta_0}{(2\pi)^3} \int d^3q \frac{\exp[i\boldsymbol{q} \cdot (\boldsymbol{r} - \boldsymbol{r}')]}{\mathscr{D}(q^2) + \kappa} .$$

(8.3.81)

If we assume that

$$(\mathscr{D}\eta)(\boldsymbol{r}) = D\nabla^2\eta(\boldsymbol{r}) ,$$

(8.3.82)

then

$$\mathscr{D}(q^2) = Dq^2 .$$

(8.3.83)

Equation (8.3.81) then has exactly the same form as derived previously in (8.3.35), namely,

$$\langle\rho(\boldsymbol{r}), \rho(\boldsymbol{r}')\rangle_s = \langle\rho\rangle_0\delta(\boldsymbol{r} - \boldsymbol{r}') + \frac{\kappa_1\langle\rho_0\rangle}{4\pi D|\boldsymbol{r} - \boldsymbol{r}'|}\exp(-|\boldsymbol{r}-\boldsymbol{r}'|/l_c)$$

(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 $\langle \eta_2(r, t) \rangle$. Now from (8.3.75)

$$\langle \tilde{\eta}_2(q, t) \rangle = - \int d^3q_1 \int_0^t dt' \exp\{-[\mathscr{D}(q^2) + \kappa](t - t')\}$$
$$\times \langle \eta_1(q - q_1, t')\eta_1(q_1, t') \rangle .\tag{8.3.86}$$

In the steady state (8.3.86) gives

$$\langle \tilde{\eta}_2(q) \rangle = -\kappa_4\kappa_1\eta_0\delta(q) \int \frac{d^3q_1}{\mathscr{D}(q^2) + \kappa}.\tag{8.3.87}$$

For the choice of $\mathscr{D}$ given by (8.3.83,87) gives

$$\langle \eta_2 \rangle = - \frac{\kappa_4\kappa_1\eta_0}{(2\pi)^{3/2}} \int \frac{d^3q_1}{Dq^2 + \kappa}\tag{8.3.88}$$

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) $\int d^3q$ is a sum over all allowed q which have a maximum value $|q| \sim 1/l$.
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

$$l|q| \ll 1,$$
$$\mathscr{D}(q^2) \to Dq^2 .$$

Nonlocal Reactions. The divergence, however, can also be avoided by preserving the continuum but modifying the $\eta^2(r)$ 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

$$\eta(r)^2 \rightarrow \int d^3r' d^3r'' \, m(r - r', r - r'')\eta(r')\eta(r'') , \tag{8.3.89}$$

one finds that instead of (8.3.88) one has

$$\langle \eta_2 \rangle = -\frac{\kappa_4 \kappa_1 \eta_0}{(2\pi)^{3/2}} \int \frac{d^3q \, \tilde{m}(q, -q)}{Dq^2 + \kappa} \tag{8.3.90}$$

where $\tilde{m}(q, q')$ is the fourier transform of $m(r, r')$. If m is sufficiently nonlocal, at high q, $\tilde{m}(q, -q)$ 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 $f_i(q)$ of D_{ij}, which satisfy

$$\sum_j D_{ij}f_j(q) = -D\lambda(q)f_i(q) .$$ (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

$$f_j(q) \propto \cos(qlj)$$ (8.4.3)

and the reflecting boundary condition requires that q has the form

$$q = \frac{n\pi}{Nl} \qquad (n = 0, 1, ..., N)$$ (8.4.4)

and

$$\lambda(q) = 4 \sin^2 \left(\frac{ql}{2}\right)\bigg/ l^2 .$$ (8.4.5)

Appropriate modifications must be made to take care of more dimensions, but the basic structure is the same, namely,

$$\lambda(0) = 0$$
$$\lambda(q) > 0 \qquad (q \neq 0) \qquad \text{and}$$ (8.4.6)

$$f_i(0) = \text{constant} = N^{-1/2} .$$ (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

$$\sum_j f_j(q)f_j(q') = \delta_{q,q'} ,$$
$$\sum_q f_i(q)f_j(q) = \delta_{i,j} .$$ (8.4.8)

We now introduce the variables

$$x(q) = \sum_j f_j(q)x_j .$$ (8.4.9)

The variable of interest is proportional to $x(0)$, since

$$x(0) = N^{-1/2} \sum_j x_j$$ (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

$$\frac{\partial P}{\partial t} = \left[\frac{\partial}{\partial x} a(x) + \frac{1}{N} \frac{\partial^2}{\partial x^2} b(x) \right] P \qquad (8.4.20)$$

which corresponds to the global Master equation, since x is by (8.4.12) a concentration. Notice that the condition for validity is that

$$D\lambda(1) \gg K \qquad (8.4.21)$$

where K represents a typical rate of the chemical reaction, or noticing that $\lambda(1) \simeq (\pi/Nl^2)$,

$$D \gg \frac{KN^2l^2}{\pi^2} \qquad (8.4.22)$$

or

$$\sqrt{\frac{D}{K}} \gg \frac{L}{\pi} . \qquad (8.4.23)$$

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 r and velocities 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 r_n and velocities v_n, which do not collide but move under the influence of an *external* force field $mA(r)$. The equations of motion are

$$\begin{aligned} \dot{r}_n &= v_n \\ \dot{v}_n &= A(r_n) . \end{aligned} \qquad (8.5.1)$$

Then a phase-space density can be defined as

$$f(r, v, t) = \sum_n \delta(r - r_n(t))\delta(v - v_n(t)) \tag{8.5.2}$$

so that

$$\partial_t f(r, v, t) = \sum_n [\dot{r}_n \cdot \nabla \delta(r - r_n(t))\delta(v - v_n(t)) + \delta(r - r_n(t))\dot{v}_n \cdot \nabla_v \delta(v - v_n(t))]$$

and using the properties of the delta function and the equations of motion (8.5.1), we get

$$\boxed{[\partial_t + v \cdot \nabla + A(r) \cdot \nabla_v] f(r, v, t) = 0.} \tag{8.5.3}$$

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 $f(r, v, t)$ is to be interpreted as the average of $f(r, v, t)$ as defined in (8.5.2) over the initial positions and velocities.

Equation (8.5.3) is exact. The variable $f(r, v, t)$ 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_t centred on the phase-space coordinate (r_i, v_i) is, of course,

$$X(r_i, v_i) = \int_{\Delta_t} d^3r \, d^3v \, f(r, v) . \tag{8.5.4}$$

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_t. 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_t, defined as

$$x_i(t) = \int_{\Delta_t} dr \, \rho(r, t) \simeq \lambda \rho(r_i, t) \tag{8.5.6}$$

where Δ_t is an interval of length λ around a cell point at r_t. 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:

$$\begin{aligned} t^+(x) &= \kappa x_i / \lambda \\ t^-(x) &= 0 \,. \end{aligned} \tag{8.5.9}$$

This is of the form studied in Sect. 7.5, whose notation (7.5.4, 5) takes the form here of

$$\left. \begin{aligned} A &\to i \\ N_j^i &\to \delta_{i,j} \\ M_j^i &\to \delta_{i+1,j} \\ r_j^i &\to \delta_{j,i+1} - \delta_{j,i} \,. \end{aligned} \right\} \tag{8.5.10}$$

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.,

$$\begin{aligned} A_a(x) &= \sum_i (\delta_{a,i+1} - \delta_{a,i}) \kappa x_i / \lambda \\ &= \frac{\kappa}{\lambda} (x_{a-1} - x_a) \end{aligned} \tag{8.5.11}$$

and the diffusion matrix

$$B_{a,b}(x) = \frac{\kappa}{\lambda} [(\delta_{a,b} - \delta_{a-1,b}) x_{a-1} + (\delta_{a,b} - \delta_{a,b-1}) x_a)] \,. \tag{8.5.12}$$

We now set $x_i = \rho(r_i)$ as in (8.5.6) and take a small λ limit; one finds

$$A_a(x) \to \frac{\kappa}{\lambda} [\lambda \rho(r_a - \lambda, t) - \lambda \rho(r_a, t)] \tag{8.5.13}$$

$$\to -\lambda \kappa \partial_r \rho(r, t) \tag{8.5.14}$$

and the limiting value of $B_{a,b}(x)$ is found similarly, but also using

$$\delta_{a,b} \to \lambda \, \delta(r_a - r_b) \tag{8.5.15}$$

$$B_{a,b}(x) \to \kappa \lambda^3 \partial_r \partial_{r'} [\delta(r - r') \rho(r)] \,. \tag{8.5.16}$$

Thus, in this limit, $\rho(r, t)$ obeys a stochastic differential equation

$$dp(r, t) = -\kappa \partial_r p(r, t)dt + \lambda^{1/2}dW(r, t) \tag{8.5.17}$$

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 $t^+(x)$ 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

$$f(r^q, v^q) = X(r^q, v^q)/\xi^3\lambda^3 . \tag{8.5.19}$$

We consider transitions of two types. For simplicity, consider these only in the x-direction and define

$$\lambda_x = (\lambda, 0, 0) . \tag{8.5.20}$$

Type 1: Flow in position space:

$$X(r^i, v^i) \to X(r^i, v^i) - 1$$

and either

$$X(r^i + \lambda_x, v^i) \to X(r^i + \lambda_x, v^i) + 1 \qquad (v_x > 0) \tag{8.5.21}$$

or

$$X(r^i - \lambda_x, v^i) \to X(r^i - \lambda_x, v^i) + 1 \qquad (v_x < 0) .$$

Then we make the labelling $A \to (i, x, 1)$:

$$\left.\begin{aligned}
N_a^{(i,x,1)} &= \delta(r^i, r^a)\delta(v^i, v^a) \\
M_a^{(i,x,1)} &= \delta(r^i - \lambda_x, r^a)\delta(v^i, v^a) \\
r_a^{(i,x,1)} &= M_a^{(i,x,1)} - N_a^{(i,x,1)}
\end{aligned}\right\} \tag{8.5.22}$$

$$t^+_{(i,x,1)}(X) = \lambda^2 |v_x| X(r^i, v^i)/\lambda^3$$

$$t^-_{(i,x,1)}(X) = 0 \,.$$

(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 $v_x > 0$,

$$A^{(x,1)}_a = v_x[X(r^a + \lambda_x, v^a) - X(r^a, v^a)]/\lambda$$

(8.5.24)

$$= \xi^3 \lambda^3 v_x[f(r^a + \lambda_x, v^a) - f(r^a, v^a)]/\lambda$$

(8.5.25)

$$\rightarrow \xi^3 \lambda^3 v_x \frac{\partial}{\partial x} f(r, v) \,.$$

(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 $\lambda \rightarrow 0$.

Type 2: Flow in velocity space:
Define

$$\boldsymbol{\xi}_x = (\xi, 0, 0) \,.$$

(8.5.27)

Then we have

$$X(r^i, v^i) \rightarrow X(r^i, v^i) - 1$$

and either

$$X(r^i, v^i + \boldsymbol{\xi}_x) \rightarrow X(r^i, v^i + \boldsymbol{\xi}_x) + 1 \qquad (A_x > 0)$$

(8.5.28)

or

$$X(r^i, v^i - \boldsymbol{\xi}_x) \rightarrow X(r^i, v^i - \boldsymbol{\xi}_x) + 1 \qquad (A_x < 0) \,.$$

The labelling is

$$\left.
\begin{aligned}
&A \rightarrow (i, x, 2) \\
&N^{(i,x,2)}_a = \delta(v^i, v^a)\delta(r^i, r^a) \\
&M^{(i,x,2)}_a = \delta(v^i - \boldsymbol{\xi}_x, v^a)\delta(r^i, r^a) \\
&r^{(i,x,2)}_a = [\delta(v^i - \boldsymbol{\xi}_x, v^a) - \delta(v^i, v^a)]\delta(r^i, r^a) \\
&t^+_{(i,x,2)}(X) = \xi^2 |A_x(r^i)| X(r^i, v^i)/\xi^3 \\
&t^-_{(i,x,2)}(X) = 0 \,.
\end{aligned}
\right\}$$

(8.5.29)

Consequently, assuming $A_x(r^a) > 0$, the drift coefficient is

$$A_a = [X(r^a, v^a + \xi_x)A_x(r^a) - X(r^a, v^a)A_x(r^a)]/\xi \tag{8.5.30}$$

$$= \xi^3\lambda^3 A_x(r^a)[f(r^a, v^a + \xi_x) - f(r^a, v^a)]/\xi \tag{8.5.31}$$

$$\rightarrow \xi^3\lambda^3 A_x(r)\frac{\partial}{\partial v_x}f(r, v). \tag{8.5.32}$$

Again, similar reasoning shows the diffusion coefficient vanishes as $\xi \rightarrow 0$.

Putting together (8.5.26, 32), one obtains the appropriate flow equation (8.5.3) in the λ, ξ, $\rightarrow 0$ 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",

$$X(v_i) + X(v_j) \rightarrow X(v_k) + X(v_l). \tag{8.5.33}$$

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

$$A_a(\boldsymbol{X}) = \sum_{i,j,k,l} (-\delta_{a,i} - \delta_{a,j} + \delta_{a,k} + \delta_{a,l}) X(v_i) X(v_j) R(ij, kl)\tag{8.5.40}$$

and utilising all available symmetries,

$$= 2 \sum_{j,k,l} R(aj, kl) [X(v_k)X(v_l) - X(v_a)X(v_j)] .\tag{8.5.41}$$

The diffusion coefficient can also be deduced;

$$B_{ab}(\boldsymbol{X}) = \sum_{i,j,k,l} r_a^{ij,kl} r_b^{ij,kl} R(ij, kl) X(v_i) X(v_j)$$

and again, utilising to the full all available symmetries,

$$\begin{aligned} B_{ab}(\boldsymbol{X}) = \ &2\delta_{a,b} \sum_{j,k,l} R(aj, kl)[X(v_a)X(v_j) + X(v_k)X(v_l)] \\ &+ 2 \sum_{i,j} R(ij, ab)[X(v_i)X(v_j) + X(v_a)X(v_b)] \\ &- 4 \sum_{j,l} R(aj, bl)[X(v_a)X(v_j) + X(v_b)X(v_l)] .\end{aligned}\tag{8.5.42}$$

These imply a stochastic differential equation

$$dX(v_a) = \{2 \sum_{j,k,l} R(aj, kl) [X(v_k)X(v_l) - X(v_a)X(v_j)]\} dt + dW(v_a, t)\tag{8.5.43}$$

where

$$dW(v_a, t)dW(v_b, t) = dt\, B_{ab}(\boldsymbol{X}) .\tag{8.5.44}$$

Neglecting the stochastic term, we recover the *Boltzmann Equation* for $X(v_a)$ 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 $\alpha(v_a)$ as usual, we can follow

through the results (7.7.6–9) to obtain a Poisson representation FPE with a drift term

$$A_a(\alpha) = 2 \sum_{j,k,l} R(aj, kl) \left[\alpha(v_k)\alpha(v_l) - \alpha(v_a)\alpha(v_j) \right] \tag{8.5.45}$$

and diffusion matrix

$$
\begin{aligned}
B_{ab}(\alpha) = {}& 2\delta_{ab} \sum_{i,j,l} R(ij, kl)[\alpha(v_i)\alpha(v_j) - \alpha(v_a)\alpha(v_l)] \\
& + 2 \sum_{i,j} R(ij, ab)[\alpha(v_i)\alpha(v_j) - \alpha(v_a)\alpha(v_b)] \\
& - \delta_{a,b} \sum_{j,l} R(aj, bl)[\alpha(v_a)\alpha(v_j) - \alpha(v_b)\alpha(v_l)]
\end{aligned} \tag{8.5.46}
$$

and the corresponding S DE is

$$d\alpha(v_a) = 2 \sum_{j,k,l} R(aj, kl) \left[\alpha(v_k)\alpha(v_l) - \alpha(v_a)\alpha(v_j) \right] + d\tilde{W}(v_a, t) \tag{8.5.47}$$

where

$$d\tilde{W}(v_a, t)d\tilde{W}(v_b, t) = dt\, B_{ab}(\alpha) . \tag{8.5.48}$$

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 $t^-_{ij,kl}(X)$ zero, but we can alternatively write

$$t^-_{ij,kl}(X) = t^+_{ij,kl}(X) \tag{8.5.49}$$

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 $\langle X \rangle$ 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,

$$\alpha(v) = \exp\left[-(v - U)^2/mkT \right] . \tag{8.5.52}$$

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

$$t^+_{ij,kl}(X) = \delta(r_i, r_j)\, \delta(r_k, r_l)\, \delta(r_i, r_k)\, R(ij, kl) X_i X_j \ . \tag{8.5.53}$$

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

$$f(r_i, v_i) = X(r_i, v_i)/\lambda^3 \xi^3 \ . \tag{8.5.54}$$

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

$$R(ij, kl) = 8\sigma[(v_i - v_j)^2, (v_i - v_j)\cdot(v_k - v_l)]$$
$$\times\ \delta(v_i^2 + v_j^2 - v_k^2 - v_l^2)\, \delta(v_i + v_j - v_k - v_l) \tag{8.5.56}$$

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 $(v_i - v_j)\cdot(v_k - v_i)$ 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

$$(\xi^3\lambda^3) \sum_j \rightarrow \int d^3r_j d^3v_j \tag{8.5.57}$$

and change the variables by

$$
\begin{aligned}
v_j &= v_i \\
v_k &= \tfrac{1}{2}(p + q) \\
v_l &= \tfrac{1}{2}(p - q).
\end{aligned}
\tag{8.5.58}
$$

After a certain amount of manipulation in the collision term, the deterministic equation comes out in the form [from (8.5.3.41)]

$$
\begin{aligned}
df(r, v) = \Big\{ &-v\cdot\nabla f(r, v) - A\cdot\nabla_v f(r, v) \\
&+ \int d^3v_1 \int \frac{d^3q}{|q|} \delta(|q| - |v - v_1|)\sigma[q^2, q\cdot(v - v_1)] \\
&\times [f[r, \tfrac{1}{2}(v + v_1 - q)]f[r, \tfrac{1}{2}(v + v_1 + q)] - f(r, v)f(r, v_1)]\Big\}dt.
\end{aligned}
\tag{8.5.59}
$$

The stochastic differential equation, arising from a Kramers-Moyal expansion, is obtained by adding a stochastic term $dW(r, v, t)$ satisfying

$$
\begin{aligned}
dW(r, v, t)\,dW(r', v', t) = \delta(r - r')dt \times \Big\{ &\delta(v - v')\int d^3v_1 \frac{d^3q}{|q|}\delta(|q| - |v - v_1|) \\
\times\sigma(q^2, q\cdot(v - v_1))]f[r, \tfrac{1}{2}(v + v_1 - q)]&f[r, \tfrac{1}{2}(v + v_1 + q)] + f(r, v)f(r, v_1)] \\
-2\int d^3k\,\delta[(v_a - v_b)\cdot k]\,\sigma[(v_a - v_b)^2 &+ \tfrac{1}{4}k^2, -(v_a - v_b)^2 + \tfrac{1}{4}k^2] \\
\times [f(v_a)f(v_b - k) + f(v_b)f(v_a - k)]&\Big\}.
\end{aligned}
\tag{8.5.60}
$$

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 $\phi(r, v)$ defined by

$$\phi(r, v) = \alpha(r, v)/(\lambda^3\xi^3) . \tag{8.5.61}$$

We find

$$
\begin{aligned}
d\phi(r, v) = dt \Bigg\{ &-v\cdot\nabla\phi(r, v) - A\cdot\nabla_v\phi(r, v) + \int d^3v_1 \frac{d^3q}{|q|} \delta(|q| - |v - v_1|) \\
&\times \sigma[q^2, q\cdot(v - v_1)] [f[r, \tfrac{1}{2}(v + v_1 - q)]f[r, \tfrac{1}{2}(v + v_1 + q)] \\
&\quad - f(r, v)f(r, v_1)] \Bigg\} + dW(r, v, t),
\end{aligned} \tag{8.5.62}
$$

where

$$
\begin{aligned}
dW(r, v, t)dW(r', v', t) = \delta(r - r')dt \times \Bigg(&\delta(v - v') \int d^3v_1 \frac{d^3q}{|q|} \delta(|q| - |v - r_1|) \\
&\times \sigma[q^2, q\cdot(v - v_1)] \{\phi[r, \tfrac{1}{2}(v + v_1 - q)] \phi[r, \tfrac{1}{2}(v + v_1 + q)] - \phi(r, v)\phi(r, v_1)\} \\
+ \int \frac{d^3q}{|q|} \delta(|q| - |v - v_1|) &\sigma[q^2, q\cdot(v - v_1)] \\
&\times \{\phi[r, \tfrac{1}{2}(r + r' - q)] \phi[r, \tfrac{1}{2}(v + v' + q)] - \phi(r, v)\phi(r, v')\} \Bigg).
\end{aligned} \tag{8.5.63}
$$

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 diffu-sion 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 diffu-sion 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
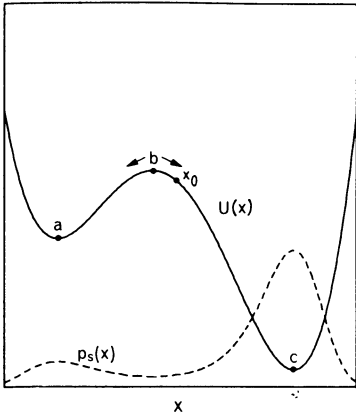
Fig. 9.1. Plot of $p_s(x)$ and $U(x)$ for a double well potential

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

$$\begin{aligned}
U(x) &\simeq U(a) + \tfrac{1}{2}U''(a)\,(x - a)^2 & |x - a| \text{ small} \\
&\simeq U(c) + \tfrac{1}{2}U''(c)\,(x - c)^2 & |x - c| \text{ small} \,.
\end{aligned} \tag{9.1.4}$$

If D is very small, then we may approximate

$$p_s(x) \simeq \mathcal{N}\exp\left[-U(a)/D - \tfrac{1}{2}U''(a)\,(x-a)^2/D\right] \qquad |x-a| \text{ small}$$
$$\simeq \mathcal{N}\exp\left[-U(c)/D - \tfrac{1}{2}U''(c)\,(x-c)^2/D\right] \qquad |x-c| \text{ small} \qquad (9.1.5)$$
$$\simeq \quad 0 \quad \text{(elsewhere)}$$

so that

$$\mathcal{N}^{-1} \simeq e^{-U(a)/D}\sqrt{2\pi D/U''(a)} + e^{-U(c)/D}\sqrt{2\pi D/U''(c)}\,. \qquad (9.1.6)$$

Suppose, as drawn in the figure,

$$U(a) > U(c)\,. \qquad (9.1.7)$$

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

$$p_s(x) = \sqrt{\frac{U''(c)}{2\pi D}}\exp\left[-\tfrac{1}{2}U''(c)\,(x-c)^2/D\right] \qquad |x-c| \sim\sqrt{D}$$
$$= 0 \quad \text{(otherwise)}\,. \qquad (9.1.8)$$

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

$$T(a \rightarrow b) = \pi[|U''(b)||U''(a)]^{-1/2} \exp \{[U(b) - U(a)]/D\} \qquad (9.1.9)$$

(half as long as the time for the particle to reach a point well to the right of b). For $D \rightarrow 0$, 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):

$$\begin{aligned}
\pi_a(x_0) &= \left[\int_{x_0}^{c} dx\, p_s(x)^{-1} \right] \Big/ \left[\int_{a}^{c} dx\, p_s(x)^{-1} \right] \\
\pi_c(x_0) &= \left[\int_{a}^{x_0} dx\, p_s(x)^{-1} \right] \Big/ \left[\int_{a}^{c} dx\, p_s(x^{-1}) \right].
\end{aligned} \qquad (9.1.10)$$

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 $D \rightarrow 0$.

a) x_0 a Finite Distance from b
We first evaluate

$$\int_{a}^{c} dx\, p_s(x)^{-1} = \int_{a}^{c} dx\, \mathcal{N}^{-1} \exp [U(x)/D]. \qquad (9.1.11)$$

This is dominated by the behaviour at $x \sim b$. An asymptotic evaluation is correctly obtained by setting

$$U(x) \simeq U(b) - \tfrac{1}{2}|U''(b)|/(b - x)^2. \qquad (9.1.12)$$

As $D \rightarrow 0$, 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

$$\mathcal{N}^{-1} \int_{-\infty}^{U(x_0)} e^{y/D} \, W'(y)dy \sim \mathcal{N}^{-1} \, D \, e^{U(x_0)/D} W'[U(x_0)]$$

$$= \mathcal{N}^{-1} \frac{D \, e^{U(x_0)/D}}{U'(x_0)}. \tag{9.1.15}$$

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]

$$\mathrm{erf}(x) = \sqrt{\frac{\pi}{2}} \int_0^x dt \, e^{-t^2}, \tag{9.1.19}$$

we find

$$\pi_c = 1 - \pi_a \sim \tfrac{1}{2}\{1 - \mathrm{erf}[y_0\sqrt{|U''(b)|}]\} \tag{9.1.20}$$

$$= \tfrac{1}{2}\left\{1 - \mathrm{erf}\left[(b - x_0)\sqrt{\frac{|U''(b)|}{D}}\right]\right\}. \tag{9.1.21}$$

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

$$T(a, x) = \frac{\pi_c(x)\int\limits_x^c dx'p_s(x')^{-1}\int\limits_a^{x'}\pi_a(z)p_s(z)dz - \pi_a(x)\int\limits_a^x dx'p_s(x')^{-1}\int\limits_a^{x'}\pi_a(z)p_s(z)dz}{D\pi_a(x)} \tag{9.1.24}$$

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

$$
\begin{aligned}
\pi_a(x) &\simeq 1 & (x < b)\\
&\simeq \tfrac{1}{2} & (x = b)\\
&\simeq 0 & (x > b)
\end{aligned}
\tag{9.1.26}
$$

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)

$$\int_a^{x'} \pi_a(z)p_s(z)dz \simeq n_a/2 ,$$ (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

$$\int_a^{x'} \pi_a(z)p_s(z)dz \simeq \int_a^{x'} p_s(z)dz = \frac{n_a}{2} - \int_{x'}^b p_s(z)dz .$$ (9.1.30)

Substituting these estimates into (9.1.24) we obtain

$$T(a, b) \simeq D^{-1} \int_a^b dx' p_s(x)^{-1} \int_{x'}^b p_s(z)dz .$$ (9.1.31)

which is the exact mean exit time from b to a in which there is a reflecting barrier at b. Similarly,

$$T(c, b) \simeq D^{-1} \int_b^c dx' p_s(x)^{-1} \int_b^{x'} p_s(z)dz .$$ (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

$$p(x, 0) = \delta(x - x_i).$$ (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 $\tag{9.2.3}$

$$N_0(t) = (c - a)\, p(x_0, t) .$$

Further, define the corresponding stationary quantities by

$$n_a = 1 - n_c = \int_{-\infty}^{c} p_s(x')dx'$$
$$\tag{9.2.4}$$
$$n_0 = (c-a)\, p_s(x_0)$$

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

$$p(x, t) = p_s(x)N_a(t)/n_a \qquad x < b$$
$$\tag{9.2.7}$$
$$= p_s(x)N_c(t)/n_c \qquad x > b .$$

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 $N_0(t)$

$$\dot{N}_0(t) = D\{N_a(t)/[n_a \kappa(x_0)] + N_c(t)/[n_c \mu(x_0)]\}$$
$$- N_0(t)\{[n_0 \kappa(x_0)]^{-1} + [n_0 \mu(x_0)]^{-1}\} . \qquad (9.2.14)$$

Since

$$n_0 = p_s(x_0)(c - a) = \mathscr{N}\exp[-U(x_0)/D](c - a) . \qquad (9.2.15)$$

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

$$\dot{N}_a(t) = -\dot{N}_c(t) = r_a N_a(t) + r_c N_c(t) \qquad (9.2.16)$$

with

$$r_a = D[n_a \int_a^c dx\, p_s(x)^{-1}]^{-1} \qquad r_c = D[n_c \int_a^c dx\, p_s(x)^{-1}]^{-1} , \qquad (9.2.17)$$

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 given by

$$\tau_r^{-1} = r_a + r_c = \frac{D}{n_a n_c \int_a^c dx\, p_s(x)^{-1}} . \qquad (9.2.18)$$

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

$$\dot{N}_a(t) = -DN_a(t)/n_a \kappa(x_0) \qquad (9.2.19)$$

so that the mean first passage time is given by

$$\tau_a = n_a D^{-1} \int_a^{x_0} dx\, p_s(x)^{-1} . \qquad (9.2.20)$$

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

$$\sqrt{U_0''(a)}\,\exp\left[U_0(a)/kT\right] = \sqrt{U_0''(c)}\,\exp\left[U_0(c)/kT\right]. \tag{9.2.29}$$

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).
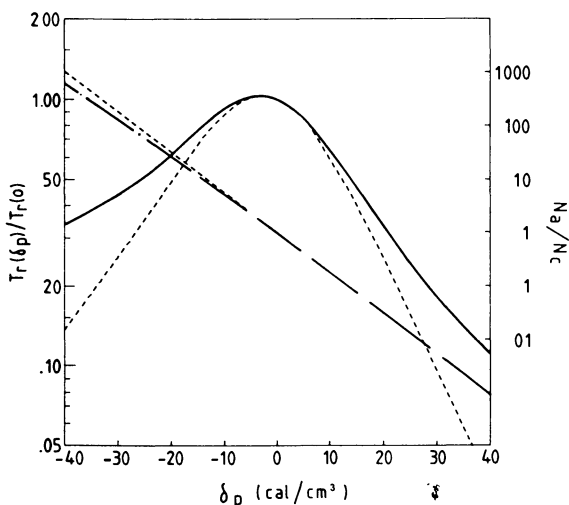
Fig. 9.2. Relaxation rate $\tau_r(\delta p)$ and ratio of concentrations of natural and denatured forms of chymotrypsinogen according to (9.2.33). Dashed line omits compressibility corrections

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

$$\frac{d}{dt} \sum_{z=a+1}^{x_0} \left[\frac{M(z, t)}{P_s(z)t^-(z)} \right] = \beta(x_0, t) - \beta(a, t)$$

$$= \frac{P(x_0, t)}{P_s(x_0)} - \frac{P(a, t)}{P_s(a)} \, . \tag{9.2.45}$$

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

$$P(x, t) = P_s(x)N_a(t)/n_a \qquad x < b$$
$$\qquad = P_s(x)N_c(t)/n_c \qquad x > b \tag{9.2.46}$$

and obtain relaxation equations exactly like those in (9.2.8)

$$\kappa(x_0)\dot{N}_a(t) = N_0(t)/n_0 - N_a(t)/n_a$$
$$\mu(x_0)\dot{N}_c(t) = N_0(t)/n_0 - N_c(t)/n_c . \tag{9.2.47}$$

where

$$\kappa(x_0) = \sum_{z=a+1}^{x_0} [P_s(z)t^-(z)]^{-1}[1 - \psi(z)]$$
$$\mu(x_0) = \sum_{z=x_0+1}^{c} [P_s(z)t^-(z)]^{-1}[1 - \psi(z)] \tag{9.2.48}$$

with

$$\psi(z) = n_a^{-1} \sum_{y=z}^{b} P_s(y) \qquad z < b$$
$$\qquad = n_c^{-1} \sum_{y=b+1}^{z+1} P_s(y) \qquad z > b . \tag{9.2.49}$$

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

$$\pi_a = \{ \sum_{z=x_0+1}^{c} [P_s(z)t^-(z)]^{-1}\}/\{ \sum_{z=a+1}^{c} [P_s(z)t^-(z)]^{-1}\}$$
$$\pi_c = \{ \sum_{z=a+1}^{x_0} [P_s(z)t^-(z)]^{-1}\}/\{ \sum_{z=a+1}^{c} [P_s(z)t^-(z)]^{-1}\} . \tag{9.2.50}$$

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

$$\partial_t p = \nabla \cdot [-v(x)p + \varepsilon \underset{\sim}{D}(x) \cdot \nabla p] \tag{9.3.1}$$

whose stationary solution is called $p_s(x)$ 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 $\underset{\sim}{D}(x)$ is the identity:

$$\underset{\sim}{D}(x) = 1 . \tag{9.3.2}$$

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 S, and that the velocity field $v(x)$ points inwards to a stationary point a. The problem is the asymptotic estimate of the distribution of points b on S at which the point escapes from R. We use (5.4.49) for $\pi(b, x)$ (the distribution of escape points on S, starting from the point x), which in this case takes the form

$$v(x) \cdot \nabla \pi(b, x) + \varepsilon \nabla^2 \pi(b, x) = 0 \tag{9.3.3}$$

with boundary condition

$$\pi(\boldsymbol{b}, \boldsymbol{u}) = \delta_s(\boldsymbol{b} - \boldsymbol{u}) \qquad (\boldsymbol{u} \in S) . \tag{9.3.4}$$

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

$$v(x) \cdot \nabla \pi(\boldsymbol{b}, x) = 0 \tag{9.3.5}$$

which implies that $\pi(\boldsymbol{b}, x)$ is constant along the flow lines of $v(x)$, since it simply states that the derivative of $\pi(\boldsymbol{b}, x)$ along these lines is zero. Since we assume all the flow lines pass through $\boldsymbol{a}$, we have

$$\pi(\boldsymbol{b}, x) = \pi(\boldsymbol{b}, \boldsymbol{a}) \tag{9.3.6}$$

for any x inside. However, the argument is flawed by the fact that $v(\boldsymbol{a}) = 0$ 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 $z(x)$ and $y_r(x)$ so that

$$
\begin{aligned}
v(x) \cdot \nabla z(x) &= -z(x) \\
v(x) \cdot \nabla y_r(x) &= 0 \\
z(\boldsymbol{a}) &= 0 .
\end{aligned}
\tag{9.3.7}
$$

The negative sign in the first of these equations takes account of the fact that $\boldsymbol{a}$ is assumed stable, so that $v(x)$ points towards $\boldsymbol{a}$. Thus, $z(x)$ *increases* as x travels further from $\boldsymbol{a}$.

Thus, we find, for any function f,

$$\nabla f = \nabla z(x) \frac{\partial f}{\partial z} + \sum_r \nabla y_r(x) \frac{\partial f}{\partial y_r} \tag{9.3.8}$$

and hence,

$$v(x) \cdot \nabla \pi = -z \frac{\partial \pi}{\partial z} \tag{9.3.9}$$

and

$$
\begin{aligned}
\nabla^2 \pi = {} & \nabla z(x) \cdot \nabla z(x) \frac{\partial^2 \pi}{\partial z^2} + 2 \sum_r \nabla z(x) \cdot \nabla y_r(x) \frac{\partial^2 \pi}{\partial z \, \partial y_r} \\
& + \sum_{r,s} \nabla y_r(x) \cdot \nabla y_s(x) \frac{\partial^2 \pi}{\partial y_r \, \partial y_s} + \nabla^2 z(x) \frac{\partial \pi}{\partial z} + \sum_r \nabla^2 y_r(x) \frac{\partial \pi}{\partial y_r} .
\end{aligned}
\tag{9.3.10}
$$

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)

$$H = \nabla z(a)\cdot\nabla z(a) \ .$$

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 $v(\boldsymbol{x})$ 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 $H \rightarrow -H$, and hence in a distance of order $\sqrt{\varepsilon}$ of 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

$$v(\boldsymbol{x})\cdot\nabla f(\boldsymbol{x}) + \varepsilon\nabla^2 f(\boldsymbol{x}) = 0$$

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) $\boldsymbol{v}\cdot v(\boldsymbol{x}) \neq 0$ *on 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 S we can write

$$v(\boldsymbol{u})\cdot\nabla f(\boldsymbol{x}) + \varepsilon\nabla^2 f(\boldsymbol{x}) = 0 \ . \tag{9.3.16}$$

Near S it is most convenient to introduce $v(\boldsymbol{u})$, the normal (pointing out) at $\boldsymbol{u}$ to S, and to define a variable ρ by

$$x = u - \varepsilon\rho v(u) \tag{9.3.17}$$

and other variables y_r parallel to S.

Then to lowest order in ε, (9.3.16) reduces to (at a point near u on S)

$$[v \cdot v(u)]\frac{\partial f}{\partial \rho} + H(u)\frac{\partial^2 f}{\partial \rho^2} = 0 \tag{9.3.18}$$

with $H(u) = v^2 = 1$.

The solution is then

$$f(x) = g(u) + C_1(u)\{1 - \exp[-v \cdot v(u)\rho]\} . \tag{9.3.19}$$

As $\rho \to \infty$, we approach a finite distance into the interior of R and thus

$$f(x) \to g(u) + C_1(u) = C_0 \tag{9.3.20}$$

from the analysis in (a), so

$$C_1(u) = C_0 - g(u) . \tag{9.3.22}$$

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 $p_s(x)$ be the usual stationary solution of the forward Fokker-Planck equation. We know it can be written

$$p_s(x) = \exp\left\{-\frac{1}{\varepsilon}[\phi(x) + O(\varepsilon)]\right\} \tag{9.3.23}$$

as has been shown in Sect. 6.3.3. We take (9.3.16), multiply by $p_s(x)$ and integrate over R. Using the fact that $p_s(x)$ satisfies the forward Fokker-Planck equation, this can be reduced to a surface integral

$$0 = \int_R dx \, p_s(x)[v(x) \cdot \nabla f(x) + \varepsilon \nabla^2 f(x)] \tag{9.3.24}$$

$$= \int_S dS \, \{p_s(x)v \cdot v(x)f(x) + \varepsilon[p_s(x)v \cdot \nabla f(x) - f(x)v \cdot \nabla p_s(x)]\} . \tag{9.3.25}$$

Noting that, to lowest order in ε.

$$v \cdot \nabla f(x) = -\frac{1}{\varepsilon}\frac{\partial f}{\partial \rho} = -v \cdot v(x)[C_0 - g(x)] \tag{9.3.26}$$

and

$$v \cdot \nabla p_s(x) = -\frac{1}{\varepsilon}v \cdot \nabla\phi(x) \exp[-\phi(x)/\varepsilon] , \tag{9.3.27}$$

we deduce that

$$C_0 = \frac{\int_S dS \, e^{-\phi(x)/\varepsilon} \, [2\mathbf{v} \cdot v(x) + \mathbf{v} \cdot \nabla \phi(x)] g(x)}{\int_S dS \, e^{-\phi(x)/\varepsilon} \, \mathbf{v} \cdot v(x)} . \tag{9.3.28}$$

Recalling that for this problem, $g(\mathbf{u}) = \delta_s(\mathbf{u} - \mathbf{b})$, we find that, if x is well in the interior of R,

$$\pi(x, \mathbf{b}) = C_0 = \frac{e^{-\phi(\mathbf{b})/\varepsilon} \, [2\mathbf{v} \cdot v(\mathbf{b}) + \mathbf{v} \cdot \nabla \phi(\mathbf{b})]}{\int_S dS \, e^{-\phi(x)/\varepsilon} \, \mathbf{v} \cdot v(x)} . \tag{9.3.29}$$

We see here that the exit distribution is essentially $\exp[-\phi(\mathbf{b})/\varepsilon]$, i.e., approximately the stationary distribution. If the Fokker-Planck equation has a potential solution, then

$$v(\mathbf{b}) = -\nabla \phi(\mathbf{b}) \tag{9.3.30}$$

and

$$\pi(x, \mathbf{b}) = e^{-\phi(\mathbf{b})/\varepsilon} \, \mathbf{v} \cdot v(\mathbf{b}) / [\int_S dS \, e^{-\phi(x)/\varepsilon} \, \mathbf{v} \cdot v(x)]$$

and we simply have a kind of average flow result.

ii) $\mathbf{v} \cdot v(x) = 0$ *on* S: this problem is more directly related to bistability, since midway between two stable points a and c, a curve $\mathbf{v} \cdot v(x) = 0$ which separates the two regions and is known as *a separatrix* is expected.

The method is much the same except that near $\mathbf{u}$ on S, we expect

$$\mathbf{v} \cdot v(x) \sim \mathbf{v} \cdot (x - \mathbf{u}) \kappa(\mathbf{u}) . \tag{9.3.31}$$

where $\kappa(\mathbf{u})$ is a coefficient which depends on $v(x)$ and is assumed to be nonzero.

The situation is now like that at $x = a$ and it is appropriate to substitute

$$x = \mathbf{u} - \sqrt{\varepsilon} \, \rho \mathbf{v}(\mathbf{u}) \tag{9.3.32}$$

and to lowest order in ε (9.3.16) reduces to (at a point near $\mathbf{u}$ on S)

$$\kappa(\mathbf{u}) \, \rho \frac{\partial f}{\partial \rho} + \frac{\partial^2 f}{\partial \rho^2} = 0 \tag{9.3.33}$$

so that

$$f(x) = g(\mathbf{u}) + C_1 \int_0^\rho d\rho \, \exp[-\tfrac{1}{2}\kappa(\mathbf{u})\rho^2] \tag{9.3.34}$$

and letting $\rho \to \infty$, we find

$$C_1 = [C_0 - g(\mathbf{u})]\sqrt{\frac{2\kappa(\mathbf{u})}{\pi}}. \tag{9.3.35}$$

The result for $\pi(\mathbf{x}, \mathbf{b})$ is now completed as before: one gets

$$\pi(\mathbf{x}, \mathbf{b}) = \frac{e^{-\phi(\mathbf{b})/\varepsilon}\left[\left(1 + \sqrt{\frac{2\kappa(\mathbf{b})}{\pi}}\right)\mathbf{v}\cdot\mathbf{v}(\mathbf{b}) + \mathbf{v}\cdot\nabla\phi(\mathbf{b})\right]}{\int dS\, e^{-\phi(\mathbf{x})/\varepsilon}\sqrt{\frac{2\kappa(\mathbf{x})}{\pi}}\,\mathbf{v}\cdot\mathbf{v}(\mathbf{x})}. \tag{9.3.36}$$

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

$$\tau(\mathbf{x}) = \exp(-K/\varepsilon)T(\mathbf{x}) \tag{9.3.37}$$

where $T(\mathbf{x})$ is the mean escape time from R starting at $\mathbf{x}$ and $\tau(\mathbf{x})$ satisfies (from Sect. 5.4)

$$\begin{aligned}
\mathbf{v}(\mathbf{x})\cdot\nabla\tau(\mathbf{x}) + \varepsilon\nabla^2\tau(\mathbf{x}) &= -e^{-K/\varepsilon}\\
\tau(\mathbf{u}) &= 0 \qquad \mathbf{u} \in S.
\end{aligned} \tag{9.3.38}$$

If this scaling is correct, then any expansion of $\tau(\mathbf{x})$ 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 $\tau(\mathbf{x})$ is essentially constant in the interior of R and can be written as [in the case $\mathbf{v}\cdot\mathbf{v}(\mathbf{x}) \neq 0$ on S]

$$\tau(\mathbf{x}) \sim C_0\{1 - \exp[-\mathbf{v}\cdot\mathbf{v}(\mathbf{u})\rho]\} \tag{9.3.39}$$

near S.

We multiply (9.3.38) by $p_s(\mathbf{x}) = \exp[-\phi(\mathbf{x})/\varepsilon]$ and use Green's theorem to obtain [in much the same way as (9.3.25) but with $\tau(\mathbf{x}) = 0$ on S]

$$-e^{-K/\varepsilon}\int_R d\mathbf{x}\, e^{-\phi(\mathbf{x})/\varepsilon} = -\int_S dS\, e^{-\phi(\mathbf{x})/\varepsilon}\,[C_0\mathbf{v}\cdot\mathbf{v}(\mathbf{x})], \tag{9.3.40}$$

i.e.,

$$C_0 = \int_R d\mathbf{x}\, e^{-[K+\phi(\mathbf{x})]/\varepsilon}\bigg/\int_S dS\, e^{-\phi(\mathbf{x})/\varepsilon}\,\mathbf{v}\cdot\mathbf{v}(\mathbf{x}). \tag{9.3.41}$$

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 $\phi(\mathbf{x})$ which occurs at the point $\mathbf{a}$, whereas in the denominator, it occurs at the point on S where $\phi(\mathbf{x})$ is a minimum, which we shall call $\mathbf{x}_0$. Thus, the ratio behaves like

$$\exp\{[\phi(a) - \phi(x_0) - K]/\varepsilon\} \ .$$

and hence for C_0 to be asymptotically constant,

$$K = \phi(a) - \phi(x_0) \tag{9.3.42}$$

and, for x well into the interior of R, we have

$$\tau(x) = \int_R dx \, e^{\phi(x)/\varepsilon}/\int_S dS \, [e^{-\phi(x)/\varepsilon} \, \mathbf{v} \cdot v(x)] \ . \tag{9.3.43}$$

In the case where $\mathbf{v} \cdot v(x) = 0$ on all of S, we now have

$$\tau(x) \sim C_0 \int_0^\rho d\rho \, \exp[-\tfrac{1}{2}\kappa(\mathbf{u})\rho^2] \tag{9.3.44}$$

and hence in the interior,

$$\tau(x) \sim C_0 \sqrt{\frac{\pi}{2\kappa(\mathbf{u})}} \ . \tag{9.3.45}$$

The analysis proceeds similarly and we find, for x well in the interior of R,

$$\tau(x) \sim \sqrt{\frac{\pi}{2\kappa(\mathbf{u})}} \int_R dx \, e^{\phi(x)/\varepsilon}/\int_S dS \, e^{-\phi(x)/\varepsilon} \ . \tag{9.3.46}$$

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 $P(x)$ for the probability density for notational ease]

$$\partial_t P = \nabla \cdot [-v(x)P + \varepsilon \underline{D}(x) \cdot \nabla P] \tag{9.3.47}$$

whose stationary solution is to be called $P_s(x)$ and can only be exhibited explicitly if (9.3.47) satisfies potential conditions. We assume that $P_s(x)$ has two well-defined maxima at a and c and well-defined saddlepoint at b (Fig. 9.3). We assume that the value at the saddlepoint is very much smaller than the values at a and c. We introduce a family of $(l-1)$ dimensional planes $S(w)$, where w is a parameter which labels the planes. We choose $S(a)$ to pass through a, $S(b)$ through b and $S(c)$ through c. The planes $S(w)$ are assumed to be oriented in such a way that $P_s(x)$ has a unique maximum when restricted to any one of them. We define, similarly to Sect. 9.2.1

$$M[S(w)] = \int_{L(w)} dx \, P(x) \ , \tag{9.3.48}$$
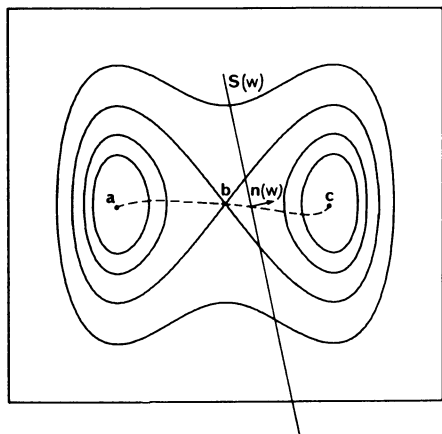
Fig. 9.3. Contours of the stationary distribution function $P_s(x)$. The plane $S(w)$ is oriented so that $P_s(x)$ has a unique maximum there, and the curve $x = u(w)$ (dashed line) is the locus of these maxima

where $L(w)$ is the region of space to the left of the plane $S(w)$; then

$$\dot{M}[S(w)] = \int_{S(w)} dS \cdot [-v(x)P + \varepsilon \underline{D}(x) \cdot \nabla P] . \tag{9.3.49}$$

Current in stationary state is defined by

$$J_s = -v(x)P_s + \varepsilon \underline{D}(x) \cdot \nabla P_s . \tag{9.3.50}$$

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

$$J_s = -\varepsilon \nabla \cdot (\underline{A}P_s) \tag{9.3.51}$$

where $\underline{A}$ is an antisymmetric tensor. *We require that $\underline{A}$ be of the some order of magnitude as $\underline{D}(x)$, or smaller.*

Relaxation equations are derived in two stages. Define a quantity $\beta(x)$ by

$$\beta(x) = P(x, t)/P_s(x) \simeq N_a(t)/n_a \qquad (x \text{ near } a) \tag{9.3.52}$$
$$\simeq N_c(t)/n_c \qquad (x \text{ near } c) .$$

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

$$\mathscr{D}(x) = \underline{D}(x) + \underline{A}(x) . \tag{9.3.54}$$

Assumption II: $P_s(x)$ is sharply singly peaked on $S(w)$ so we may make the approximate evaluation

$$\dot{M}[S(w)] = \{\varepsilon[\boldsymbol{n}(w)\cdot\mathscr{D}(\boldsymbol{x})\cdot\nabla\beta]_{\boldsymbol{u}(w)} + \delta(w)\}\,|\int\limits_{S(w)} dS\,P_s(\boldsymbol{x})|\,, \tag{9.3.55}$$

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 $S(w)$, and $\boldsymbol{n}(w)$ is the normal to $S(w)$.

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

$$p(w) = |\int\limits_{S(w)} dS\,P_s(\boldsymbol{x})|\,, \tag{9.3.57}$$

which is (up to a slowly varying factor) the probability density for the particle to be on the plane $S(w)$ 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

$$\partial_w\boldsymbol{u}(w)\cdot\nabla\beta[u(w)] = \partial_w\beta[\boldsymbol{u}(w)]\,, \tag{9.3.58}$$

we find

$$\frac{1}{\varepsilon}\int\limits_a^{w_0} dw\,\{\dot{M}[S(w)]/[p(w)d(w)]\} = \beta(w_0) - \beta(a)\,. \tag{9.3.59}$$

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

$$N(a,\,t) = M[S(b),\,t] \tag{9.3.60}$$

as well as defining

$$\kappa(w_0) = \int\limits_a^{w_0} [p(w)]^{-1}dw \tag{9.3.61}$$

$$\mu(w_0) = \int\limits_{w_0}^b [p(w)]^{-1}dw\,, \tag{9.3.62}$$

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

$$
\left.
\begin{aligned}
&x = (x, p) \\
&v(x) = (p, -U'(x) - \gamma p) \\
&\varepsilon = 1 \\
&\underset{\sim}{D}(x) = \begin{bmatrix} 0 & 0 \\ 0 & \gamma \end{bmatrix} \\
&P_s(x) = \mathcal{N}_2 \exp\left[-\tfrac{1}{2}p^2 - U(x)\right] \\
&\mathcal{N}_2 = (2\pi)^{-1/2}\mathcal{N}_1 \\
&\mathcal{N}_1 = \left\{\int\limits_{-\infty}^{\infty} dx\, \exp\left[-U(x)\right]\right\}^{-1}.
\end{aligned}
\right\}
\tag{9.3.66}
$$

Hence, we can write

$$v(x) = \begin{bmatrix} 0 & -1 \\ 1 & \gamma \end{bmatrix} \cdot \nabla(\log P_s) \tag{9.3.67}$$

and the current in the stationary state is

$$J_s = -vP_s + \underset{\sim}{D}\cdot\nabla P_s = -\nabla\cdot\left\{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}P_s\right\} \tag{9.3.68}$$

so that $\underset{\sim}{A}$ exists, and

$$\underset{\sim}{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{9.3.69}$$

Thus, Assumption I is satisfied.

The plane $S(w)$ can be written in the form

$$\lambda x + p = w. \tag{9.3.70}$$

Assumption II requires us to maximise $P_s(x)$ on this plane, i.e., to maximise $-\tfrac{1}{2}p^2 - U(x)$ on this plane. Using standard methods, we find that maxima must lie along the curve $u(w)$ 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,

$$\mathscr{D}^\mathsf{T} \boldsymbol{n} = (1 + \lambda^2)^{-1/2} \begin{bmatrix} 0 & 1 \\ -1 & \gamma \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = (1 + \lambda^2)^{-1/2} \begin{bmatrix} 1 \\ \gamma - \lambda \end{bmatrix} \tag{9.3.75}$$

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 $w = 0 \iff x = 0$. 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 $P_s(x)$ on it.

We now integrate (9.3.57) and determine $d(w)$. Notice that $d(w)$ must be defined with $n(w)$ 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,

$$\begin{aligned}
p(w) &= \int_{S(w)} |dS\,P_s(x)| = \int_{S(w)} \sqrt{dx^2 + dp^2}\,P_s(x, p) \\
&= \mathcal{N}_2 \frac{[1 + \lambda^2]^{1/2}}{\lambda} \int dp\,\exp\left[-\frac{p^2}{2} - U\left(\frac{w - p}{\lambda}\right)\right].
\end{aligned} \tag{9.3.85}$$

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

$$p(w) = \frac{(1 + \lambda^2)^{1/2}}{\lambda}\,\mathcal{N}_2\,e^{-U_0}\exp\left[\frac{U_2 w^2}{2(\lambda^2 - U_2)}\right] \tag{9.3.87}$$

and thus

$$\kappa(0) = \int_{-\infty}^{0} p(w)^{-1}dw = \tfrac{1}{2}\mathcal{N}_2^{-1}\frac{\lambda\gamma}{(1 + \lambda^2)^{1/2}}\frac{e^{U_0}}{\sqrt{U_2}} = \mu(0). \tag{9.3.88}$$

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 $x = 0$,

$$\tau_0 = \kappa(0)\,d(0)^{-1} = \frac{\lambda}{2}\,e^{U_0}\,\mathcal{N}_1^{-1}\sqrt{\frac{2\pi}{U_2}}, \tag{9.3.89}$$

i.e.,

$$\boxed{\tau_0 = \frac{1}{2}\left(\frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} + U_2}\right)e^{U_0}\,\mathcal{N}_1^{-1}\sqrt{\frac{2\pi}{U_2}}.} \tag{9.3.90}$$

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

$$\tau_1 = \gamma \int\limits_a^0 dx\, \exp\left[U(x)\right] \int\limits_{-\infty}^x dz\, \exp\left[-U(z)\right]. \tag{9.3.93}$$

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

$$\tau_2 = \tfrac{1}{2}\gamma\, e^{U_0}\,\mathcal{N}_1^{-1}\sqrt{\frac{2\pi}{U_2}} \tag{9.3.94}$$

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);

$$\frac{\partial \hat{P}}{\partial t} = \frac{1}{\gamma}\frac{\partial}{\partial x}\left\{\left[1 + \gamma^{-2}U''(x)\right]\left[U'(x)\,\hat{P} + \frac{\partial \hat{P}}{\partial x}\right]\right\}. \tag{9.3.96}$$

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

$$\begin{aligned} p &\to p T^{1/2} \\ x &\to x \, T^{1/2} , \end{aligned} \tag{9.3.106}$$

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

$$V(x, T) = U(xT^{1/2}) .$$ (9.3.108)

The simulations were performed with

$$U(x) = \tfrac{1}{4}(x^2 - 1)^2$$ (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.
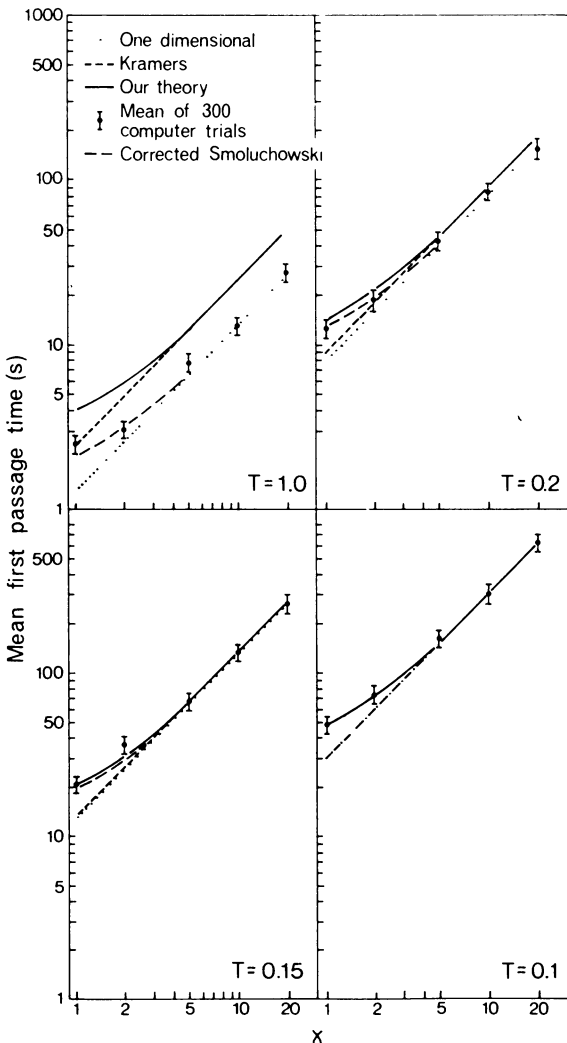
Fig. 9.4. Comparison of various estimates of the mean exit time from the double well potential of Sect. 9.3.4

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. **24**C, 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

Greek Symbols

Mathematical Symbols

Author Index

Subject Index

Springer Series in Synergetics

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: