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

$$= \frac{a\mu+b\lambda}{\mu+\lambda} + \mathrm{e}^{-(\lambda+\mu)(t-t_0)}\left(x_0 - \frac{a\mu+b\lambda}{\mu+\lambda}\right). \qquad (3.8.91)$$

The variance can also be computed but is a very messy expression.

b) Stationary Solutions: This process has the stationary solution obtained by letting $t_0 \to -\infty$:

$$P_s(a) = \frac{\mu}{\lambda+\mu}\,, \qquad P_s(b) = \frac{\lambda}{\lambda+\mu}\,, \qquad (3.8.92)$$

which is obvious from the master equation.

The stationary mean and variance are

$$\langle X\rangle_s = \frac{a\mu+b\lambda}{\mu+\lambda}\,, \qquad (3.8.93)$$

$$\mathrm{var}\,\{X\}_s = \frac{(a-b)^2\mu\lambda}{(\lambda+\mu)^2}\,. \qquad (3.8.94)$$

c) Stationary Correlation Functions: To compute the stationary time correlation function, let $t \geqslant s$, and write

$$\langle X(t)X(s)\rangle_s = \sum_{xx'} xx'P(x,t\,|\,x',s)P_s(x')\,, \qquad (3.8.95)$$

$$= \sum_{x'} x'\langle X(t)\,|\,[x',s]\rangle P_s(x')\,. \qquad (3.8.96)$$

Now use (3.8.91–3.8.94) to obtain

$$\langle X(t)X(s)\rangle_s = \langle X\rangle_s^2 + \exp[-(\lambda+\mu)(t-s)](\langle X^2\rangle_s - \langle X\rangle_s^2)\,, \qquad (3.8.97)$$

$$= \left(\frac{a\mu+b\lambda}{\mu+\lambda}\right)^2 + \exp[-(\lambda+\mu)(t-s)]\frac{(a-b)^2\mu\lambda}{(\lambda+\mu)^2}\,. \qquad (3.8.98)$$

Hence,

$$\langle X(t),X(s)\rangle_s = \langle X(t)X(s)\rangle_s - \langle X\rangle_s^2 = \frac{(a-b)^2\mu\lambda}{(\lambda+\mu)^2}\,\mathrm{e}^{(\lambda+\mu)|t-s|}\,. \qquad (3.8.99)$$

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

$$\frac{dx}{dt} = a(x,t) + b(x,t)\xi(t)\,, \qquad (4.1.1)$$

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 nonzero mean can be absorbed into the definition of $a(x,t)$, and thus require that

$$\langle\xi(t)\xi(t')\rangle = \delta(t-t')\,, \qquad (4.1.2)$$

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. 1.5.2 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

$$\langle X(t),X(t')\rangle = \frac{\gamma}{2}\,\mathrm{e}^{-\gamma|t-t'|}\,. \qquad (4.1.3)$$

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 $k \to \infty$ 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

$$u(t) = \int_0^t dt' \xi(t')\,, \tag{4.1.6}$$

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

$$u(t') = \int_0^t ds\, \xi(s) + \int_t^{t'} ds\, \xi(s)\,, \tag{4.1.7}$$

$$= \lim_{\varepsilon\to 0} \Big[\int_0^{t-\varepsilon} ds\, \xi(s) \Big] + \int_t^{t'} ds\, \xi(s)\,, \tag{4.1.8}$$

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

$$\langle u(t+\Delta t) - u_0 \,|\, [u_0, t] \rangle = \Big\langle \int_t^{t+\Delta t} \xi(s) ds \Big\rangle = 0\,, \tag{4.1.9}$$

and

$$\langle [u(t+\Delta t) - u_0]^2 \,|\, [u_0, t] \rangle = \int_t^{t+\Delta t} ds \int_t^{t+\Delta t} ds' \langle \xi(s)\xi(s') \rangle\,, \tag{4.1.10}$$

$$= \int_t^{t+\Delta t} ds \int_t^{t+\Delta t} ds' \delta(s-s') = \Delta t\,. \tag{4.1.11}$$

This means that the drift and diffusion coefficients are

$$A(u_0, t) = \lim_{\Delta t\to 0} \frac{\langle u(t+\Delta t) - u_0 \,|\, [u_0, t] \rangle}{\Delta t} = 0\,, \tag{4.1.12}$$

$$B(u_0, t) = \lim_{\Delta t\to 0} \frac{\langle [u(t+\Delta t) - u_0]^2 \,|\, [u_0, t] \rangle}{\Delta t} = 1\,. \tag{4.1.13}$$

The corresponding Fokker-Planck equation is that of the Wiener process and we can write

$$\int_0^t \xi(t') dt' = u(t) = W(t)\,. \tag{4.1.14}$$

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*

$$x(t) - x(0) = \int_0^t a[x(s), s]\, ds + \int_0^t b[x(s), s] \xi(s)\, ds\,, \tag{4.1.15}$$

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

$$dW(t) \equiv W(t+dt) - W(t) = \xi(t) dt \tag{4.1.16}$$

and thus write the second integral as

$$\int_0^t b[x(s), s] dW(s)\,, \tag{4.1.17}$$

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. 1.5.1. 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

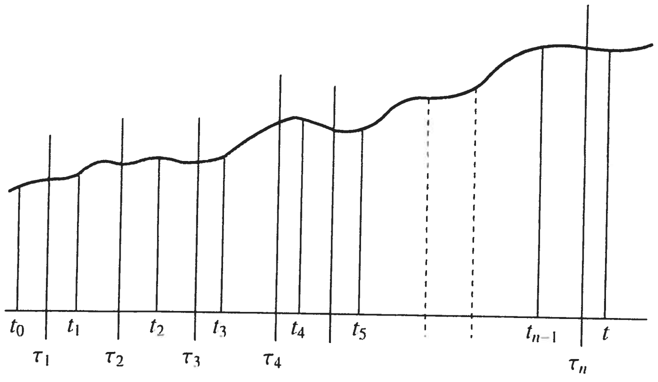

Fig. 4.1. Partitioning of the time interval used in the definition of stochastic integration

Namely, we divide the interval $[t_0, t]$ into n subintervals by means of partitioning points (as in Fig. 4.1)

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

and define intermediate points τ_i such that

$$t_{i-1} \leqslant \tau_i \leqslant t_i\,. \tag{4.2.2}$$

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

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

It is heuristically quite easy to see that, in general, the integral defined as the limit of S_n *depends on the particular choice of intermediate point* τ_i. For example, if we take the choice of $G(\tau_i) = W(\tau_i)$,

$$\langle S_n \rangle = \Big\langle \sum_{i=1}^{n} W(\tau_i)[W(t_i) - W(t_{i-1})] \Big\rangle\,, \tag{4.2.4}$$

$$= \sum_{i=1}^{n} [\min(\tau_i, t_i) - \min(\tau_i, t_{i-1})]\,, \tag{4.2.5}$$

$$= \sum_{i=1}^{n} (\tau_i - t_{i-1})\,. \tag{4.2.6}$$

If, for example, we choose for all i

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

$$\text{then } \langle S_n \rangle = \sum_{i=1}^{n} (t_i - t_{i-1})\,\alpha = (t - t_0)\alpha\,, \tag{4.2.8}$$

So that the mean value of the integral can be anything between zero and $(t - t_0)$, depending on the choice of intermediate points.

4.2.2 Ito Stochastic Integral

The choice of intermediate points characterised by $\alpha = 0$, that is the choice

$$\tau_i = t_{i-1} \tag{4.2.9}$$

defines the *Ito stochastic integral* of the function $G(t)$ by

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

By ms-lim we mean the *mean square limit*, as defined in Sect. 2.9.2.

4.2.3 Example $\int_{t_0}^{t} W(t')dW(t')$

An exact calculation is possible. We write [writing W_i for $W(t_i)$]

$$S_n = \sum_{i=1}^{n} W_{i-1}(W_t - W_{i-1}) \equiv \sum_{i=1}^{n} W_{i-1}\Delta W_i\,, \tag{4.2.12}$$

$$= \tfrac{1}{2} \sum_{i=1}^{n} [(W_{i-1} + \Delta W_i)^2 - (W_{i-1})^2 - (\Delta W_i)^2]\,, \tag{4.2.13}$$

$$= \tfrac{1}{2}[W(t)^2 - W(t_0)^2] - \tfrac{1}{2} \sum_{i=1}^{n} (\Delta W_i)^2\,. \tag{4.2.14}$$

We can calculate the mean square limit of the last term. Notice that

$$\Big\langle \sum \Delta W_i^{\,2} \Big\rangle = \sum_i \langle (W_i - W_{i-1})^2 \rangle = \sum_i (t_i - t_{i-1}) = t - t_0\,. \tag{4.2.15}$$

Because of this,

$$\Big\langle \Big[\sum_i (W_i - W_{i-1})^2 - (t - t_0)^2 \Big]^2 \Big\rangle = \Big\langle \sum_i (W_i - W_{i-1})^4 + 2\sum_{i<j} (W_i - W_{i-1})^2 (W_j - W_{j-1})^2 - 2(t - t_0) \sum_i (W_i - W_{i-1})^2 + (t - t_0)^2 \Big\rangle\,. \tag{4.2.16}$$

Notice the $W_i - W_{i-1}$ is a Gaussian variable and is independent of $W_j - W_{j-1}$. Hence, we can factorise. Thus,

$$\langle (W_i - W_{i-1})^2 (W_j - W_{j-1})^2 \rangle = (t_i - t_{i-1})(t_j - t_{j-1})\,, \tag{4.2.17}$$

and also, using the formula (2.8.7) for the fourth moment of a Gaussian variable

$$\langle (W_i - W_{i-1})^4 \rangle = 3\,\langle (W_i - W_{i-1})^2 \rangle^2 = 3(t_i - t_{i-1})^2\,, \tag{4.2.18}$$

which combined with (4.2.17) gives

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

$$= 2\sum_i (t_i - t_{i-i})^2 + \sum_{i,j} [(t_i - t_{i-1}) - (t - t_0)][(t_j - t_{j-1}) - (t - t_0)]\,,$$

$$= 2\sum_i (t_i - t_{i-1})^2 \to 0 \text{ as } n \to \infty\,. \tag{4.2.19}$$

Thus,

$$\underset{n\to\infty}{\text{ms-lim}} \sum_i (W_i - W_{i-1})^2 = t - t_0\,, \tag{4.2.20}$$

by definition of the mean square limit, so

$$\boxed{\int_{t_0}^{t} W(t')\,dW(t') = \tfrac{1}{2}[W(t)^2 - W(t_0)^2 - (t - t_0)]\,. \qquad (4.2.21)}$$

Comments

i) $$\Big\langle \int_{t_0}^{t} W(t)\,dW(t) \Big\rangle = \tfrac{1}{2}[\langle W(t)^2\rangle - \langle W(t_0)^2\rangle - (t - t_0)] = 0\,. \tag{4.2.22}$$

This is also obvious by definition, since the individual terms are $\langle W_{i-1}\Delta W_i\rangle$, 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.

4.2.4 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. 4.4—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

$$(S)\int_{t_0}^{t} W(t')\,dW(t') = \underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n} \frac{W(t_i) + W(t_{i-1})}{2}[W(t_i) - W(t_{i-1})]\,, \tag{4.2.23}$$

$$= \tfrac{1}{2}[W(t)^2 - W(t_0)^2]\,. \tag{4.2.24}$$

However, the integral as defined by Stratonovich [which we will always designate by a prefixed (S) as in (4.2.23)] 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. 4.4].

4.2.5 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

$$x(t) - x(t_0) = \int_{t_0}^{t} a[x(t'),t']dt' + \int_{t_0}^{t} b[x(t'),t']dW(t')\,. \tag{4.2.25}$$

A function $G(t)$ is called a *nonanticipating function of t* if $G(t)$ is statistically independent of $W(s) - W(t)$ for all s and t such that $t < s$. 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 (4.2.25) in which it seems heuristically obvious that $x(t)$ involves $W(t')$ only for $t' \leqslant t$.

For example, specific nonanticipating functions of t are:

i) $W(t)\,,$

ii) $\int_{t_0}^{t} F[W(t')]\,dt'\,,$

iii) $\int_{t_0}^{t} F[W(t')]\,dW(t'),$

$$\left.\begin{array}{ll}\text{iv)} & \int_{t_0}^{t} G(t')\,dt'\,, \\ \text{v)} & \int_{t_0}^{t} G(t')\,dW(t')\,,\end{array}\right\} \quad \text{when } G(t) \text{ is itself a nonanticipating function.} \tag{4.2.26}$$

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.

4.2.6 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 that for an arbitrary *nonanticipating function* $G(t)$

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

$$= \begin{cases} \int_{t_0}^{t} dt' G(t')\,, & \text{for } N = 0\,, \\ 0\,, & \text{for } N > 0\,. \end{cases} \tag{4.2.27}$$

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

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

$$= \lim_{n\to\infty} \left\{ \left\langle \sum_i \underbrace{(G_{i-1})^2} \, \underbrace{(\Delta W_i^2 - \Delta t_i)^2} + \sum_{i>j} \underbrace{2G_{i-1}G_{j-1}(\Delta W_j^2 - \Delta t_j)} \, \underbrace{(\Delta W_i^2 - \Delta t_i)} \right\rangle \right\}. \tag{4.2.29}$$

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_j 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

$$I = 2 \lim_{n\to\infty} \left[\sum_i \Delta t_i^2 \langle (G_{i-1})^2 \rangle \right]. \tag{4.2.30}$$

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

$$\underset{n\to\infty}{\text{ms-lim}} \left(\sum_i G_{i-1}\Delta W_i^2 - \sum_i G_{i-1}\Delta t_i \right) = 0\,, \tag{4.2.31}$$

and since

$$\underset{n\to\infty}{\text{ms-lim}} \sum_i G_{i-1}\Delta t_i = \int_{t_0}^{t} dt' G(t')\,, \tag{4.2.32}$$

we have

$$\boxed{\int_{t_0}^{t} [dW(t')]^2 G(t') = \int_{t_0}^{t} dt' G(t')\,.} \tag{4.2.33}$$

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 given in Sect. 2.8.1.

ii) $dW(t)$ only occurs in integrals so that when we restrict ourselves to nonanticipating functions, we can simply write

$$dW(t)^2 \equiv dt\,, \tag{4.2.34}$$

$$dW(t)^{2+N} \equiv 0\,, \quad (N > 0)\,. \tag{4.2.35}$$

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,

$$\Delta W_i = W(t_i) - W(t_{i-1})\,, \tag{4.2.36}$$

$$G_{i-1} = G\left(\tfrac{1}{2}(t_i + t_{i-1})\right)\,, \tag{4.2.37}$$

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

$$\int_{t_0}^{t} G(t')\,dt'\,dW(t') \equiv \underset{n\to\infty}{\text{ms-lim}} \sum G_{i-1}\Delta W_i \Delta t_i = 0\,, \tag{4.2.38}$$

and similarly for higher powers. The simplest way of characterising these results is to say that $dW(t)$ is an infinitesimal order of $\frac{1}{2}$ and that in calculating differentials, infinitesimals of order higher than 1 are discarded.

4.2.7 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. 4.2.6:

$$d[W(t)]^n = [W(t) + dW(t)]^n - W(t)^n = \sum_{r=1}^{n} \binom{n}{r} W(t)^{n-r}\,dW(t)^r\,, \tag{4.2.39}$$

and using the fact that $dW(t)^r \to 0$ for all $r > 2$,

$$= nW(t)^{n-1}\,dW(t) + \frac{n(n-1)}{2} W(t)^{n-2}\,dt\,, \tag{4.2.40}$$

so that

$$\boxed{\int_{t_0}^{t} W(t')^n\,dW(t') = \frac{1}{n+1}[W(t)^{n+1} - W(t_0)^{n+1}] - \frac{n}{2}\int_{t_0}^{t} W(t')^{n-1}\,dt'\,.} \tag{4.2.41}$$

c) Two Kinds of Integral: We note that for each $G(t)$ there are two kinds of integrals, namely,

$$\int_{t_0}^{t} G(t')dt' \quad \text{and} \quad \int_{t_0}^{t} G(t')dW(t')\,, \tag{4.2.42}$$

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,

$$d\{\exp[W(t)]\} = \exp[W(t) + dW(t)] - \exp[W(t)]\,, \tag{4.2.43}$$

$$= \exp[W(t)]\left[dW(t) + \tfrac{1}{2}dW(t)^2\right]\,, \tag{4.2.44}$$

$$= \exp[W(t)]\left[dW(t) + \tfrac{1}{2}dt\right]. \tag{4.2.45}$$

For an arbitrary function

$$df[W(t),t] = \frac{\partial f}{\partial t}dt + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial f}{\partial W}dW(t) + \frac{1}{2}\frac{\partial^2 f}{\partial W^2}[dW(t)]^2 + \frac{\partial^2 f}{\partial W \partial t}dt\,dW(t) + \ldots \qquad (4.2.46)$$

and we use

$$[dW(t)]^2 \to dt\,, \qquad (4.2.47)$$

$$dt\,dW(t) \to 0\,, \quad \text{[Sect. 4.2.6, comment (iv)]} \qquad (4.2.48)$$

$$(dt)^2 \to 0\,, \qquad (4.2.49)$$

and all higher powers vanish, to arrive at

$$\boxed{df[W(t),t] = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W^2}\right)dt + \frac{\partial f}{\partial W}dW(t)\,. \qquad (4.2.50)}$$

e) Mean Value Formula: For nonanticipating $G(t)$,

$$\boxed{\left\langle \int_{t_0}^{t} G(t')\,dW(t')\right\rangle = 0\,. \qquad (4.2.51)}$$

Proof: Since $G(t)$ is nonanticipating, in the definition of the stochastic integral,

$$\left\langle \sum_i G_{i-1}\Delta W_i \right\rangle = \sum_i \langle G_{i-1}\rangle\langle \Delta W_i\rangle = 0\,. \qquad (4.2.52)$$

We know from Sect. 2.9.5 that operations of ms-lim and $\langle\ \rangle$ may be interchanged. Hence, taking the limit of (4.2.52), 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 may be correlated with ΔW_i.

f) Correlation Formula: If $G(t)$ and $H(t)$ are arbitrary continuous nonanticipating functions,

$$\boxed{\left\langle \int_{t_0}^{t} G(t')\,dW(t') \int_{t_0}^{t} H(t')\,dW(t')\right\rangle = \int_{t_0}^{t} \langle G(t')H(t')\rangle\,dt'\,. \qquad (4.2.53)}$$

Proof: Notice that

$$\left\langle \sum_i G_{i-1}\Delta W_i \sum_j H_{j-1}\Delta W_j \right\rangle$$

$$= \left\langle \sum_i G_{i-1}H_{i-1}(\Delta W_i)^2 \right\rangle + \left\langle \sum_{i>j} (G_{i-1}H_{j-1} + G_{j-1}H_{i-1})\Delta W_j \Delta W_i \right\rangle. \qquad (4.2.54)$$

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

$$\langle \Delta W_i^2 \rangle = \Delta t_i\,, \qquad (4.2.55)$$

and interchanging mean and limit operations, the result follows.

g) Relation to Delta-Correlated White Noise: Formally, this is equivalent to the idea that Langevin terms $\xi(t)$ are delta correlated and uncorrelated with $F(t)$ and $G(t)$. For, rewriting

$$dW(t) \to \xi(t)dt\,, \qquad (4.2.56)$$

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

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

$$= \int_{t_0}^{t} dt' \langle G(t')H(t')\rangle\,, \qquad (4.2.57)$$

which implies

$$\langle \xi(t)\xi(s)\rangle = \delta(t-s)\,. \qquad (4.2.58)$$

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

$$I_1 = \int_{t_1}^{t_2} dt f(t)\delta(t-t_1)\,, \qquad (4.2.59)$$

or

$$I_2 = \int_{t_1}^{t_2} dt f(t)\delta(t-t_2). \qquad (4.2.60)$$

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

$$I_1 = f(t_1)\,, \qquad (4.2.61)$$

$$I_2 = 0\,, \qquad (4.2.62)$$

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

$$\left\langle \int_{t_0}^{t} G(t')dW(t')\left[\int_{t_0}^{t'} H(s')dW(s')\right]\right\rangle = 0\,. \qquad (4.2.63)$$

This follows, since the function defined by the integral inside the square bracket is, by Sect. 4.2.5 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. 4.2.7e.

Now using the formulation in terms of the Langevin source $\xi(t)$, we can rewrite (4.2.63) as

$$\int_{t_0}^{t} dt' \int_{t_0}^{t'} ds' \langle G(t')H(s')\rangle \delta(t'-s') = 0\,, \tag{4.2.64}$$

which corresponds to not counting the weight of the delta function at the upper limit. Consequently, the full weight must be counted at the lower limit.

This property is a direct consequence of the definition of the Ito integral as in (4.2.10), in which the increment points "towards the future". That is, we can interpret

$$dW(t) = W(t+dt) - W(t)\,. \tag{4.2.65}$$

In the case of the Stratonovich integral, we get quite a different formula, which is by no means as simple to prove as in the Ito case, but which amounts to choosing

$$\left.\begin{aligned} I_1 &= \tfrac{1}{2}f(t_1)\,, \\ I_2 &= \tfrac{1}{2}f(t_2). \end{aligned}\right\} \quad \text{(Stratonovich)} \tag{4.2.66}$$

This means that in both cases, the delta function occurring at the limit of an integral has half its weight counted. This formula, although intuitively more satisfying than the Ito form, is more complicated to use, especially in the perturbation theory of stochastic differential equations, where the Ito method makes very many terms vanish.

4.3 Stochastic Differential Equations (SDE)

We concluded in Sect. 4.1, that the most satisfactory interpretation of the Langevin equation

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

is a stochastic integral equation

$$x(t) - x(0) = \int_0^t dt' a[x(t'),t'] + \int_0^t dW(t')b[x(t'),t']\,. \tag{4.3.2}$$

Unfortunately, the kind of stochastic integral to be used is not given by the reasoning of Sect. 4.1. The Ito integral is mathematically and technically the most satisfactory, but 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.

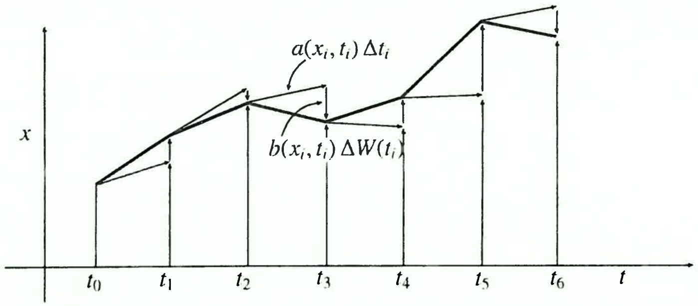

Fig. 4.2. Illustration of the Cauchy-Euler procedure for constructing an approximate solution of the stochastic differential equation $dx(t) = a[x(t),t]dt + b[x(t),t]dW(t)$

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. 8.1.

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,

$$x(t) = x(t_0) + \int_{t_0}^{t} a[x(t'),t']\,dt' + \int_{t_0}^{t} b[x(t'),t']\,dW(t')\,. \tag{4.3.4}$$

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

$$x_{i+1} = x_i + a(x_i,t_i)\Delta t_i + b(x_i,t_i)\Delta W_i\,. \tag{4.3.6}$$

Here,

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

a) Cauchy-Euler Construction of the Solution of an Ito SDE: 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 $W(t)$.

This method of constructing a solution is called the *Cauchy-Euler* method, and can be used to generate simulations. However, there are significantly better algorithms, as is explained in Chap. 10.

b) Existence and Uniqueness of Solutions of an Ito SDE: 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,

$$\frac{dx}{dt} = \tfrac{1}{2} a x^3 \tag{4.3.12}$$

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 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

$$dx = a(\lambda, x, t)\, dt + b(\lambda, x, t)\, dW(t) \,, \tag{4.3.14}$$

with initial condition

$$x(t_0) = c(\lambda) \,, \tag{4.3.15}$$

where λ is a parameter. Let the solution of (4.3.14) be $x(\lambda, t)$. Suppose

i) $\displaystyle \operatorname*{st-lim}_{\lambda\to\lambda_0} c(\lambda) = c(\lambda_0) \,,$ (4.3.16)

ii) For every $N > 0$

$$\lim_{\lambda\to\lambda_0} \left\{ \sup_{t\in[t_0,T],\ |x|<N} \left[|a(\lambda, x, t) - a(\lambda_0, x, t)| + |b(\lambda, x, t) - b(\lambda_0, x, t)| \right] \right\} = 0 \,, \tag{4.3.17}$$

iii) There exists a K independent of λ such that

$$|a(\lambda, x, t)|^2 + |b(\lambda, x, t)|^2 \leqslant K^2(1 + |x^2|) \,. \tag{4.3.18}$$

Then,

$$\operatorname*{st-lim}_{\lambda\to\lambda_0} \left\{ \sup_{t\in[t_0,T]} |x(\lambda, t) - x(\lambda_0, t)| \right\} = 0 \,. \tag{4.3.19}$$

For a proof see [4.1].

Check ii)

Comments

i) Recalling the definition of stochastic limit, the interpretation of the limit (4.3.19) 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 $\lim\limits_{\lambda\to\lambda_0}$ by $\operatorname*{st-lim}\limits_{\lambda\to\lambda_0}$.

4.3.3 Markov Property of the Solution of an Ito SDE

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].

4.3.4 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. 4.2.6 to expand $df[x(t)]$ to second order in $dW(t)$:

$$df[x(t)] = f[x(t) + dx(t)] - f[x(t)]\,, \tag{4.3.20}$$

$$= f'[x(t)]dx(t) + \tfrac{1}{2}f''[x(t)]dx(t)^2 + \dots\,, \tag{4.3.21}$$

$$= f'[x(t)]\Big\{a[x(t),t]dt + b[x(t),t]dW(t)\Big\} + \tfrac{1}{2}f''[x(t)]b[x(t),t]^2\,dW(t)^2. \tag{4.3.22}$$

where all other terms have been discarded since they are of higher order. Now use $dW(t)^2 = dt$ to obtain

$$\boxed{df[x(t)] = \Big\{a[x(t),t]f'[x(t)] + \tfrac{1}{2}b[x(t),t]^2 f''[x(t)]\Big\}\,dt + b[x(t),t]f'[x(t)]\,dW(t)\,. \qquad (4.3.23)}$$

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 infinitesimal of order $\frac{1}{2}$. By similar methods to those used in Sect. 4.2.6, we can show that for an n dimensional Wiener process $\boldsymbol{W}(t)$,

$$dW_i(t)\,dW_j(t) = \delta_{ij}\,dt\,, \tag{4.3.24a}$$

$$dW_i(t)^{N+2} = 0\,, \qquad (N > 0)\,, \tag{4.3.24b}$$

$$dW_i(t)\,dt = 0\,, \tag{4.3.24c}$$

$$dt^{1+N} = 0\,, \qquad (N > 0)\,. \tag{4.3.24d}$$

which imply that $dW_i(t)$ is an infinitesimal of order $\frac{1}{2}$. Note, however, that (4.3.24a) is a consequence of the independence of $dW_i(t)$ and $dW_j(t)$. To develop Ito's formula for functions of an n dimensional vector $\boldsymbol{x}(t)$ satisfying the stochastic differential equation

$$d\boldsymbol{x} = A(\boldsymbol{x},t)dt + \mathsf{B}(\boldsymbol{x},t)d\boldsymbol{W}(t)\,, \tag{4.3.25}$$

we simply follow this procedure. The result is

$$\boxed{\begin{aligned} df(\boldsymbol{x}) = &\Big\{\sum_i A_i(\boldsymbol{x},t)\partial_i f(\boldsymbol{x}) + \tfrac{1}{2}\sum_{i,j}[\mathsf{B}(\boldsymbol{x},t)\mathsf{B}^{\mathrm{T}}(\boldsymbol{x},t)]_{ij}\partial_i\partial_j f(\boldsymbol{x})\Big\}\,dt \\ &+ \sum_{i,j} B_{ij}(\boldsymbol{x},t)\partial_i f(\boldsymbol{x})\,dW_j(t)\,. \qquad (4.3.26)\end{aligned}}$$

4.3.5 Connection Between Fokker-Planck Equation and Stochastic Differential Equation

a) Forward Fokker-Planck Equation: We now consider the time development of an arbitrary $f(x(t))$. Using Ito's formula

$$\frac{\langle df[x(t)]\rangle}{dt} = \left\langle\frac{df[x(t)]}{dt}\right\rangle = \frac{d}{dt}\langle f[x(t)]\rangle\,,$$

$$= \Big\langle a[x(t),t]\,\partial_x f + \tfrac{1}{2}b[x(t),t]^2\partial_x^2 f\Big\rangle\,. \tag{4.3.27}$$

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

$$\frac{d}{dt}\langle f[x(t)]\rangle = \int dx\, f(x)\partial_t p(x,t\,|\,x_0,t_0)\,,$$

$$= \int dx\Big[a(x,t)\partial_x f + \tfrac{1}{2}b(x,t)^2\partial_x^2 f\Big]p(x,t\,|\,x_0,t_0)\,. \tag{4.3.28}$$

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

$$\int dx f(x)\partial_t p = \int dx f(x)\Big\{-\partial_x[a(x,t)p] + \tfrac{1}{2}\partial_x^2[b(x,t)^2 p]\Big\}\,, \tag{4.3.29}$$

and hence, since $f(x)$ is arbitrary,

$$\boxed{\partial_t p(x,t\,|\,x_0,t_0) = -\partial_x[a(x,t)p(x,t\,|\,x_0,t_0)] + \tfrac{1}{2}\partial_x^2[b(x,t)^2 p(x,t\,|\,x_0,t_0)]\,.} \tag{4.3.30}$$

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. 3.5.2, in which it was shown that the diffusion process could be locally approximated by an equation resembling an Ito stochastic differential equation.

b) Backward Fokker-Planck Equation—the Feynman-Kac Formula: Suppose a function $g(x,t)$ obeys the backward Fokker-Planck equation

$$\partial_t g = -a(x,t)\partial_x g - \tfrac{1}{2}b(x,t)\partial_x^2 g\,, \tag{4.3.31}$$

with the final condition

$$g(x,T) = G(x)\,. \tag{4.3.32}$$

If $x(t)$ obeys the stochastic differential equation (4.3.3), then using Ito's rule (adapted appropriately to account for explicit time dependence), the function $g[x(t),t]$ obeys the stochastic differential equation

$$\begin{aligned} dg[x(t),t] = {} & \left\{\partial_t g + a[x(t),t]\,\partial_x g[x(t),t] + \tfrac{1}{2}b[x(t),t]^2\,\partial_x^2 g[x(t),t]\right\}dt \\ & + b[x(t),t]\partial_x g[x(t),t]\,dW(t)\,, \end{aligned} \tag{4.3.33}$$

and using (4.3.31) this becomes

$$dg[x(t),t] = b[x(t),t]\,\partial_x g[x(t),t]\,dW(t)\,. \tag{4.3.34}$$

Now integrate from t to T, and take the mean

$$\langle g[x(T),T]\rangle - \langle g[x(t),t]\rangle = \left\langle \int_t^T b[x(t'),t']\,\partial_x g[x(t'),t']\,dW(t')\right\rangle = 0. \tag{4.3.35}$$

Let the initial condition of the stochastic differential equation for $x(t')$ and $t' = t$ be

$$x(t) = x\,, \tag{4.3.36}$$

where x is a non-stochastic value, so that

$$\langle g[x(t),t]\rangle = g(x,t)\,. \tag{4.3.37}$$

At the other end of the interval, use the final condition (4.3.32) to write

$$\langle g[x(T),T]\rangle = \langle G[x(T)]\,|\,x(t) = x\rangle\,, \tag{4.3.38}$$

where the notation on the right hand side indicates the mean conditioned on the initial condition (4.3.36).

Putting these two together, the *Feynman-Kac* formula results:

$$\langle G[x(T)]\,|\,x(t) = x\rangle = g(x,t)\,, \tag{4.3.39}$$

where $g(x,t)$ is the solution of the backward Fokker-Planck equation (4.3.31) with initial condition (4.3.32).

This formula is essentially equivalent to the fact that $p(x,t\,|\,x_0,t_0)$ obeys the backward Fokker-Planck equation in the arguments x_0, t_0, as shown in Sect. 3.6, since

$$\langle G[x(T)]|x(t_0) = x_0\rangle = \int dx\, G(x)p(x,T\,|\,x_0,t_0)\,. \tag{4.3.40}$$

4.3.6 Multivariable Systems

In general, many variable systems of stochastic differential equations can be defined for n variables by

$$d\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x},t)\,dt + \mathsf{B}(\boldsymbol{x},t)\,d\boldsymbol{W}(t)\,, \tag{4.3.41}$$

where $d\boldsymbol{W}(t)$ is an n variable Wiener process, as defined in Sect. 3.8.1. The many variable version of the reasoning used in Sect. 4.3.5 shows that the Fokker-Planck equation for the conditional probability density $p(\boldsymbol{x},t\,|\,\boldsymbol{x}_0,t_0) \equiv p$ is

$$\partial_t p = -\sum_i \partial_i[A_i(\boldsymbol{x},t)p] + \tfrac{1}{2}\sum_{i,j}\partial_i\partial_j\{[\mathsf{B}(\boldsymbol{x},t)\mathsf{B}^{\mathrm{T}}(\boldsymbol{x},t)]_{ij}p\}\,. \tag{4.3.42}$$

Notice that the same Fokker-Planck equation arises from all matrices B such that $\mathsf{B}\mathsf{B}^{\mathrm{T}}$ is the same. This means that we can obtain the same Fokker-Planck equation by replacing B by $\mathsf{B}\mathsf{S}$ where S is orthogonal, i.e., $\mathsf{S}\mathsf{S}^{\mathrm{T}} = 1$. Notice that S may depend on $\boldsymbol{x}(t)$.

This can be seen directly from the stochastic differential equation. Suppose $\mathsf{S}(t)$ is an orthogonal matrix with an arbitrary *nonanticipating* dependence on t. Then define

$$d\boldsymbol{V}(t) = \mathsf{S}(t)\,d\boldsymbol{W}(t)\,. \tag{4.3.43}$$

Now the vector $d\boldsymbol{V}(t)$ is a linear combination of Gaussian variables $d\boldsymbol{W}(t)$ with coefficients $\mathsf{S}(t)$ which are independent of $d\boldsymbol{W}(t)$, since $\mathsf{S}(t)$ is nonanticipating. For any fixed value of $\mathsf{S}(t)$, the $d\boldsymbol{V}(t)$ are thus Gaussian and their correlation matrix is

$$\begin{aligned} \langle dV_i(t)\,dV_j(t)\rangle &= \sum_{l,m} S_{il}(t)\,S_{jm}(t)\,\langle dW_l(t)\,dW_m(t)\rangle\,, \\ &= \sum_l S_{il}(t)\,S_{jl}(t)\,dt = \delta_{ij}\,dt\,, \end{aligned} \tag{4.3.44}$$

since $\mathsf{S}(t)$ is orthogonal. Hence, all the moments are independent of $\mathsf{S}(t)$ and are the same as those of $d\boldsymbol{W}(t)$, so $d\boldsymbol{V}(t)$ is itself Gaussian with the same correlation matrix as $d\boldsymbol{W}(t)$. Finally, averages at different times factorise, for example, if $t > t'$ in

$$\sum_{i,k}\langle[dW_i(t)S_{ij}(t)]^m[dW_k(t')S_{kl}(t')]^n\rangle\,, \tag{4.3.45}$$

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 $\mathsf{S}(t)$ gives, after averaging over $dW_i(t)$, simply

$$\sum_k\langle[dW_j(t)]^m\rangle\,\langle[dW_k(t')S_{kl}(t')]^n\rangle\,, \tag{4.3.46}$$

which similarly gives $\langle[dW_j(t)]^m[dW_l(t')]^n\rangle$. Hence, the $d\boldsymbol{V}(t)$ 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 (4.3.41) we can write

$$\begin{aligned} d\boldsymbol{x} &= \boldsymbol{A}(\boldsymbol{x},t)dt + \mathsf{B}(\boldsymbol{x},t)\mathsf{S}^{\mathrm{T}}(t)\mathsf{S}(t)\,d\boldsymbol{W}(t)\,, && (4.3.47) \\ &= \boldsymbol{A}(\boldsymbol{x},t)dt + \mathsf{B}(\boldsymbol{x},t)\mathsf{S}^{\mathrm{T}}(t)\,d\boldsymbol{V}(t)\,, && (4.3.48) \end{aligned}$$

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

$$dx = A(x,t)dt + \mathsf{B}(x,t)\mathsf{S}^{\mathrm{T}}(t)\,dW(t)\,, \tag{4.3.49}$$

which has exactly the same Fokker-Planck equation (4.3.42).

We will return to some examples in which this identity is relevant in Sect. 4.5.5.

4.4 The Stratonovich Stochastic Integral

The Stratonovich stochastic integral is an alternative to the Ito definition, in which Ito's formula, developed in Sect. 4.3.4, is replaced by the ordinary chain rule for change of variables. This apparent advantage does not come without cost, since in Stratonovich's definition the independence of a non-anticipating integrand $G(t)$ and the increment $dW(t)$ in a stochastic integral no longer holds. This means that increment and the integrand are correlated, and therefore to give a full definition of the Stratonovich integral requires some way of specifying what this correlation is.

This correlation is implicitly specified in the situation of most interest, the case in which the integrand is a function whose stochastic nature arises from its dependence on a variable $x(t)$ which obeys a stochastic differential equation. Since the aim is to recover the chain rule for change of variables in a stochastic differential equation, this seems a reasonable restriction.

4.4.1 Definition of the Stratonovich Stochastic Integral

Stratonovich [4.2] defined a stochastic integral of an integrand which is a function of $x(t)$ and t by

$$(S)\int_{t_0}^{t} G[x(t'),t']dW(t') = \operatorname*{ms-lim}_{n\to\infty} \sum_{i=1}^{n} G\left\{\tfrac{1}{2}\big(x(t_i)+x(t_{i-1})\big),t_{i-1}\right\}[W(t_i)-W(t_{i-1})]\,. \tag{4.4.1}$$

The Stratonovich integral is clearly *related* to a mid-point choice of τ_i in the definition of stochastic integration as given in Sect. 4.2.1, but clearly is *not* necessarily equivalent to that definition. Rather, instead of evaluating x at the midpoint $\frac{1}{2}(t_i+t_{i-1})$, the average of the values at the two time points is taken. Furthermore it is only the dependence on $x(t)$ that is averaged in this way, and not the explicit dependence on t. However, if $G(z,t)$ is differentiable in t, the integral can be shown to be independent of the particular choice of value for t in the range $[t_{i-1},t_i]$.

4.4.2 Stratonovich Stochastic Differential Equation

It is possible to write a stochastic differential equation (SDE) using Stratonovich's integral

$$x(t) = x(t_0) + \int_{t_0}^{t} dt'\alpha[x(t'),t'] + (S)\int_{t_0}^{t} dW(t')\beta[x(t'),t']\,. \tag{4.4.2}$$

a) Change of Variables for the Stratonovich SDE: The definition of the Stratonovich integral is such as to make the ordinary rules of calculus valid for change of variables. This means, that for the Stratonovich integral, Ito's formula (4.3.23) is replaced by the simple calculus rule

$$\boxed{(S)\,df[x(t)] = f'[x(t)]\left\{a[x(t),t]\,dt + b[x(t),t]\,dW(t)\right\}. \qquad (4.4.3)}$$

This can be proved quite simply from the definition (4.4.1). The essence of the proof can be explained by using the simple SDE

$$(S)\,dx(t) = B[x(t)]\,dW(t). \tag{4.4.4}$$

In discretised form, this can be written

$$x_{i+1} = x_i + B[\tfrac{1}{2}(x_{i+1}+x_i)]\,(W_{i+1}-W_i). \tag{4.4.5}$$

To find the Stratonovich SDE for $f[x(t)]$, we need only use the Taylor series expansion of a function about a midpoint in the form

$$f(x+a) = f(x-a) + \sum_{n=0}^{\infty} \frac{f^{2n+1}(x)\,a^{2n+1}}{(2n+1)!}\,. \tag{4.4.6}$$

In expanding $f(x_{i+1})$ we only need to keep terms up to second order, so we drop all but the first two terms and write

$$f(x_{i+1}) = f(x_i) + f'[\tfrac{1}{2}(x_{i+1}+x_i)](x_{i+1}-x_i)\,, \tag{4.4.7}$$

$$= f'[\tfrac{1}{2}(x_{i+1}+x_i)]\,B[\tfrac{1}{2}(x_{i+1}+x_i)]\,(W_{i+1}-W_i). \tag{4.4.8}$$

This means that the Stratonovich SDE for $f[x(t)]$ is

$$(S)\,df[x(t)] = f'[x(t)]\,B[x(t)]\,dW(t)\,, \tag{4.4.9}$$

which is the ordinary calculus rule. The extension to the general case (4.4.3) is straightforward.

b) Equivalent Ito SDE: We shall show that the Stratonovich SDE is in fact equivalent to an appropriate Ito SDE. Let us assume that $x(t)$ is a solution of the Ito SDE

$$dx(t) = a[x(t),t]dt + b[x(t),t]\,dW(t)\,, \tag{4.4.10}$$

and deduce the α and β for a corresponding Stratonovich equation of the form (4.4.2). In both cases, the solution $x(t)$ is the same function.

We first compute the connection between the Ito integral $\int_{t_0}^{t} dW(t')b[x(t'),t']$ and the Stratonovich integral $(S)\int_{t_0}^{t} dW(t')\beta[x(t'),t']$:

$$(S)\int_{t_0}^{t} dW(t')\beta[x(t'),t'] \simeq \sum_i \beta\left[\tfrac{1}{2}\big(x(t_i)+x(t_{i-1})\big),t_{i-1}\right]\Delta W(t_{i-1})\,. \tag{4.4.11}$$

In (4.4.11) we write

$$x(t_i) = x(t_{i-1}) + \Delta x(t_{i-1})\,, \tag{4.4.12}$$

and use the Ito SDE (4.4.10) to write

$$\int_{t_0}^{t} b(t')\,dW(t')\,, \tag{4.5.3}$$

is simply a linear combination of infinitesimal Gaussian variables. Further,

$$\langle x(t)\rangle = \langle x_0\rangle + \int_{t_0}^{t} a(t')\,dt'\,, \tag{4.5.4}$$

(since the mean of the Ito integral vanishes) and

$$\langle [x(t)-\langle x(t)\rangle][x(s)-\langle x(s)\rangle]\rangle \equiv \langle x(t), x(s)\rangle\,, \tag{4.5.5}$$

$$= \mathrm{var}[x_0] + \left\langle \int_{t_0}^{t} b(t')\,dW(t') \int_{t_0}^{s} b(s')\,dW(s')\right\rangle, \tag{4.5.6}$$

$$= \mathrm{var}[x_0] + \int_{t_0}^{\min(t,s)} [b(t')]^2\,dt'\,, \tag{4.5.7}$$

where we have used the result (4.2.53) with, however,

$$G(t') = \begin{cases} b(t')\,, & t' < t\,,\\ 0\,, & t' \geqslant t\,, \end{cases} \tag{4.5.8}$$

$$H(t') = \begin{cases} b(t')\,, & t' < s\,,\\ 0\,, & t' \geqslant s\,. \end{cases} \tag{4.5.9}$$

The process is thus completely determined.

4.5.2 Multiplicative Linear White Noise Process—Geometric Brownian Motion

The equation

$$dx = cx\,dW(t)\,, \tag{4.5.10}$$

is known as a multiplicative white noise process because it is linear in x, but the "noise term" $dW(t)$ multiplies x. It is also commonly know as *geometric Brownian motion.*

We can solve this exactly by using Ito's formula. Let us define a new variable by

$$y = \log x\,, \tag{4.5.11}$$

so that

$$dy = \frac{1}{x}dx - \frac{1}{2x^2}(dx)^2 = c\,dW(t) - \tfrac{1}{2}c^2\,dt\,. \tag{4.5.12}$$

This equation can now be directly integrated, so we obtain

$$y(t) = y(t_0) + c[W(t) - W(t_0)] - \tfrac{1}{2}c^2(t-t_0)\,, \tag{4.5.13}$$

and hence,

$$x(t) = x(t_0)\exp\left\{c[W(t) - W(t_0)] - \tfrac{1}{2}c^2(t-t_0)\right\}. \tag{4.5.14}$$

a) Mean value: We can calculate the mean by using the formula for any Gaussian variable z with zero mean

$$\langle \exp z\rangle = \exp\left(\tfrac{1}{2}\langle z^2\rangle\right)\,, \tag{4.5.15}$$

so that

$$\langle x(t)\rangle = \langle x(t_0)\rangle \exp\left[\tfrac{1}{2}c^2(t-t_0) - \tfrac{1}{2}c^2(t-t_0)\right] = \langle x(t_0)\rangle\,. \tag{4.5.16}$$

This result is also obvious from definition, since

$$d\langle x(t)\rangle = \langle dx(t)\rangle = \langle cx(t)\,dW(t)\rangle = 0\,. \tag{4.5.17}$$

b) Autocorrelation Function: We can also calculate the autocorrelation function

$$\begin{aligned}\langle x(t)x(s)\rangle &= \langle x(t_0)^2\rangle \left\langle \exp\left\{c[W(t)+W(s)-2W(t_0)] - \tfrac{1}{2}c^2(t+s-2t_0)\right\}\right\rangle\,,\\ &= \langle x(t_0)^2\rangle \exp\left\{\tfrac{1}{2}c^2[\langle [W(t)+W(s)-2W(t_0)]^2\rangle - (t+s-2t_0)]\right\}\,,\\ &= \langle x(t_0)^2\rangle \exp\left\{\tfrac{1}{2}c^2[t+s-4t_0+2\min(t,s) - (t+s-2t_0)]\right\}\,,\\ &= \langle x(t_0)^2\rangle \exp\{c^2\min(t-t_0, s-t_0)\}\,. \end{aligned} \tag{4.5.18}$$

c) Stratonovich Interpretation: The solution of this equation interpreted as a Stratonovich equation can also be obtained, but ordinary calculus would then be valid. Thus, instead of (4.5.12) we would obtain

$$(S)\,dy = c\,dW(t)\,, \tag{4.5.19}$$

and hence,

$$x(t) = x(t_0)\exp\{c[W(t) - W(t_0)]\}\,. \tag{4.5.20}$$

In this case,

$$\langle x(t)\rangle = \langle x(t_0)\rangle \exp\left[\tfrac{1}{2}c^2(t-t_0)\right]\,, \tag{4.5.21}$$

and

$$\langle x(t)x(s)\rangle = \langle x(t_0)^2 \exp\left\{\tfrac{1}{2}c^2[t+s-2t_0+2\min(t-t_0, s-T_0)]\right\}\,. \tag{4.5.22}$$

One sees that there is a clear difference between these two answers.

4.5.3 Complex Oscillator with Noisy Frequency

This is a simplification of a model due to *Kubo* [4.4] and is a slight generalisation of the previous example for complex variables. We consider

$$\frac{dz}{dt} = \mathrm{i}\left(\omega + \sqrt{2\gamma}\,\xi(t)\right)z\,, \tag{4.5.23}$$

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

$$(S)\,dz = \mathrm{i}\left(\omega\,dt + \sqrt{2\gamma}\,dW(t)\right)z\,, \tag{4.5.24}$$

which is equivalent to the Ito equation (from Sect. 4.4)

$$\Delta x(t_i) = a[x(t_{i-1}), t_{i-1}]\,\Delta t_{i-1} + b[x(t_{i-1}), t_{i-1}]\,\Delta W(t_{i-1})\,. \tag{4.4.13}$$

Then, applying Ito's formula, we can write

$$\begin{aligned}\beta\Big[\tfrac{1}{2}\big(x(t_i) + x(t_{i-1})\big), t_{i-1}\Big] &= \beta\Big[x(t_{i-1}) + \tfrac{1}{2}\Delta x(t_{i-1}), t_{i-1}\Big],\\ &= \beta(t_{i-1}) + \Big[a(t_{i-1})\partial_x\beta(t_{i-1}) + \tfrac{1}{4}b^2(t_{i-1})\Big]\tfrac{1}{2}\Delta t_{i-1}\\ &\quad + \tfrac{1}{2}b(t_{i-1})\partial_x\beta(t_{i-1})\,\Delta W(t_{i-1})\,.\end{aligned} \tag{4.4.14}$$

(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 (4.4.10) and dropping as usual $dt^2, dt\,dW(t)$, and setting $dW(t)^2 = dt$, 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

$$(S)\int_{t_0}^{t}\beta[x(t'), t']\,dW(t') = \int_{t_0}^{t}\beta[x(t'), t']\,dW(t') + \tfrac{1}{2}\int_{t_0}^{t} b[x(t'), t']\partial_x\beta[x(t'), t']\,dt'\,. \tag{4.4.15}$$

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 (4.4.2). It does not give a general connection between the Ito and Stratonovich integrals of arbitrary functions.

If we now make the choice

$$\begin{aligned}\alpha(x,t) &= a(x,t) - \tfrac{1}{2}b(x,t)\partial_x b(x,t)\\ \beta(x,t) &= b(x,t)\end{aligned} \tag{4.4.16}$$

we see that:

The Ito SDE

$$dx = a\,dt + b\,dW(t)\,, \tag{4.4.17}$$

is the same as the Stratonovich SDE

$$(S)\,dx = \big(a - \tfrac{1}{2}b\partial_x b\big)\,dt + b\,dW(t)\,, \tag{4.4.18}$$

and conversely,

The Stratonovich SDE

$$(S)\,dx = \alpha\,dt + \beta\,dW(t)\,, \tag{4.4.19}$$

is the same as the Ito SDE

$$dx = \big(\alpha + \tfrac{1}{2}\beta\partial_x\beta\big)\,dt + \beta\,dW(t)\,. \tag{4.4.20}$$

c) Many Variables: If a many variable Ito equation is

$$d\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x}, t)\,dt + \mathsf{B}(\boldsymbol{x}, t)\,d\boldsymbol{W}(t)\,, \tag{4.4.21}$$

then the corresponding Stratonovich equation can be shown similarly to be given by replacing

$$\begin{aligned}A_i^{\rm s} &= A_i - \tfrac{1}{2}\sum_{j,k} B_{kj}\partial_k B_{ij}\\ \mathsf{B}^{\rm s} &= \mathsf{B}\,.\end{aligned} \tag{4.4.22}$$

d) Fokker-Planck Equation: Corresponding to the Stratonovich SDE,

$$(S)\,d\boldsymbol{x} = \boldsymbol{A}^{\rm s}(\boldsymbol{x}, t)dt + \mathsf{B}^{\rm s}(\boldsymbol{x}, t)\,d\boldsymbol{W}(t)\,, \tag{4.4.23}$$

we can, by use of (4.4.22) and the known correspondence (Sect. 4.3.6) between the Ito stochastic differential equation and Fokker-Planck equation, show that the equivalent Fokker-Planck equation is

$$\partial_t p = -\sum_i \partial_i\{A_i^{\rm s} p\} + \tfrac{1}{2}\sum_{i,j,k}\partial_i\{B_{ik}^{\rm s}\partial_j[B_{jk}^{\rm s} p]\}\,, \tag{4.4.24}$$

which is often known as the "Stratonovich form" of the Fokker-Planck equation. In contrast to the two forms of the stochastic differential equation, 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—see Sect. 8.3.

4.5 Some Examples and Solutions

4.5.1 Coefficients without *x* Dependence

The simple equation

$$dx = a(t)\,dt + b(t)\,dW(t)\,, \tag{4.5.1}$$

with $a(t)$ and $b(t)$ nonrandom functions of time, is solved simply by integrating

$$x(t) = x_0 + \int_{t_0}^{t} a(t')\,dt' + \int_{t_0}^{t} b(t')\,dW(t')\,. \tag{4.5.2}$$

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

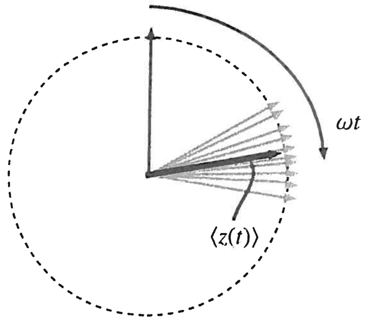

Fig. 4.3. Illustration of the decay of the mean amplitude of a complex oscillator of as a result of dephasing.

$$dz = [(\mathrm{i}\omega - \gamma)\,dt + \mathrm{i}\sqrt{2\gamma}\,dW(t)]z\,. \tag{4.5.25}$$

Taking the mean value, we see immediately that

$$\frac{d\langle z\rangle}{dt} = (\mathrm{i}\omega - \gamma)\langle z\rangle\,, \tag{4.5.26}$$

with the *damped* oscillatory solution

$$\langle z(t)\rangle = \exp[(i\omega - \gamma)t]\langle z(0)\rangle\,. \tag{4.5.27}$$

We shall show fully in Sect. 8.3, why the Stratonovich model is more appropriate. The most obvious way to see this is to note that $\xi(t)$ would, in practice, be somewhat smoother than a white noise and ordinary calculus would apply, as is the case in the Stratonovich interpretation.

Now in this case, the correlation function obtained from solving the original Stratonovich equation is

$$\langle z(t)z(s)\rangle = \langle z(0)^2\rangle \exp[(\mathrm{i}\omega - \gamma)(t+s) - 2\gamma \min(t,s)]\,. \tag{4.5.28}$$

In the limit $t, s \to \infty$, with $t + \tau = s$,

$$\lim_{t\to\infty}\langle z(t+\tau)z(t)\rangle = 0\,. \tag{4.5.29}$$

However, the correlation function of physical interest is the complex correlation

$$\begin{aligned}\langle z(t)z^*(s)\rangle &= \langle |z(0)|^2\rangle\langle \exp\{\mathrm{i}\omega(t-s) + \mathrm{i}\sqrt{2\gamma}[W(t) - W(s)]\}\rangle\,,\\ &= \langle |z(0)|^2\rangle \exp\{\mathrm{i}\omega(t-s) - \gamma[t + s - 2\min(t,s)]\}\,,\\ &= \langle |z(0)|^2\rangle \exp[\mathrm{i}\omega(t-s) - \gamma|t-s|]\,.\end{aligned} \tag{4.5.30}$$

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 the phases of an ensemble of initial states with identical phases diffuse away from the value ωt arising from the deterministic motion, as illustrated in Fig. 4.3. The mean of the ensemble consequently decays, although the amplitude $|z(t)|$ of any member of the ensemble is unchanged. 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 similar, and this model may be regarded as a prototype model of oscillators with noisy frequency.

4.5.4 Ornstein-Uhlenbeck Process

Taking the Fokker-Planck equation given for the Ornstein-Uhlenbeck process in Sect. 3.8.4, we can immediately write down the SDE using the result of Sect. 4.3.5:

$$dx = -kx\,dt + \sqrt{D}\,dW(t)\,, \tag{4.5.31}$$

and solve this directly. Putting

$$y = x\,\mathrm{e}^{kt}\,, \tag{4.5.32}$$

then

$$\begin{aligned}dy &= (dx)d(\mathrm{e}^{kt}) + (dx)\,\mathrm{e}^{kt} + xd(\mathrm{e}^{kt})\\ &= \left[-kx\,dt + \sqrt{D}\,dW(t)\right]k\,\mathrm{e}^{kt}\,dt + \left[-kx\,dt + \sqrt{D}\,dW(t)\right]\mathrm{e}^{kt} + kx\,\mathrm{e}^{kt}\,dt.\end{aligned} \tag{4.5.33}$$

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

$$dy = \sqrt{D}\,\mathrm{e}^{kt}\,dW(t)\,, \tag{4.5.34}$$

so that integrating and resubstituting for y, we get

$$x(t) = x(0)\,\mathrm{e}^{-kt} + \sqrt{D}\int_0^t \mathrm{e}^{-k(t-t')}\,dW(t')\,. \tag{4.5.35}$$

If the initial condition is deterministic or Gaussian distributed, then $x(t)$ is clearly Gaussian, with mean and variance

$$\langle x(t)\rangle = \langle x(0)\rangle\,\mathrm{e}^{-kt}\,, \tag{4.5.36}$$

$$\mathrm{var}[x(t)] = \left\langle\left\{[x(0) - \langle x(0)\rangle]\,\mathrm{e}^{-kt} + \sqrt{D}\int_0^t \mathrm{e}^{-k(t-t')}\,dW(t')\right\}^2\right\rangle. \tag{4.5.37}$$

Taking the initial condition to be nonanticipating, that is, independent of $dW(t)$ for $t > 0$, we can write using the result of Sect. 4.4f

$$\begin{aligned}\mathrm{var}[x(t)] &= \mathrm{var}[x(0)]\,\mathrm{e}^{-2kt} + D\int_0^t \mathrm{e}^{-2k(t-t')}\,dt'\,,\\ &= (\mathrm{var}[x(0)] - D/2k)\,\mathrm{e}^{-2kt} + D/2k\,.\end{aligned} \tag{4.5.38}$$

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,

$$\langle x(t), x(s)\rangle = \mathrm{var}[x(0)]\,\mathrm{e}^{-k(t+s)} + D\Big\langle \int_0^t \mathrm{e}^{-k(t-t')}\,dW(t') \int_0^s \mathrm{e}^{-k(s-s')}\,dW(s')\Big\rangle ,$$

$$= \mathrm{var}\{x(0)\}\,\mathrm{e}^{-k(t+s)} + D \int_0^{\min(t,s)} \mathrm{e}^{-k(t+s-2t')}\,dt' ,$$

$$= \Big[\mathrm{var}\{x(0)\} - \frac{D}{2k}\Big]\mathrm{e}^{-k(t+s)} + \frac{D}{2k}\,\mathrm{e}^{-k|t-s|} . \qquad (4.5.39)$$

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 (4.5.35) becomes

$$x(t) = \sqrt{D} \int_{-\infty}^{t} \mathrm{e}^{-k(t-t')}\,dW(t') . \qquad (4.5.40)$$

in which the correlation function and the mean obviously assume their stationary values. Since the process is Gaussian, this makes it stationary.

4.5.5 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.,

$$dE_1(t) = -\gamma E_1(t)\,dt + \varepsilon\,dW_1(t) , \qquad (4.5.41a)$$

$$dE_2(t) = -\gamma E_2(t)\,dt + \varepsilon\,dW_2(t) . \qquad (4.5.41b)$$

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

$$E_1(t) = a(t)\cos\phi(t) , \qquad (4.5.42)$$

$$E_2(t) = a(t)\sin\phi(t) , \qquad (4.5.43)$$

and for simplicity, also define

$$\mu(t) = \log a(t) , \qquad (4.5.44)$$

so that

$$\mu(t) + \mathrm{i}\phi(t) = \log[E_1(t) + \mathrm{i}E_2(t)] . \qquad (4.5.45)$$

We then use the Ito calculus to derive

$$d(\mu + \mathrm{i}\phi) = \frac{d(E_1 + \mathrm{i}E_2)}{E_1 + \mathrm{i}E_2} - \frac{[d(E_1 + \mathrm{i}E_2)]^2}{2(E_1 + \mathrm{i}E_2)^2} ,$$

$$= -\frac{\gamma(E_1 + \mathrm{i}E_2)}{E_1 + \mathrm{i}E_2}\,dt + \frac{\varepsilon[dW_1(t) + \mathrm{i}dW_2(t)]}{(E_1 + \mathrm{i}E_2)} - \frac{\varepsilon^2[dW_1(t) + \mathrm{i}dW_2(t)]^2}{2(E_1 + \mathrm{i}E_2)^2} , \qquad (4.5.46)$$

and noting $dW_1(t)\,dW_2(t) = 0$, and $dW_1(t)^2 = dW_2(t)^2 = dt$, it can be seen that the last term vanishes, so we find

$$d[\mu(t) + \mathrm{i}\phi(t)] = -\gamma\,dt + \varepsilon\exp[-\mu(t) - \mathrm{i}\phi(t)]\{dW_1(t) + \mathrm{i}\,dW_2(t)\} . \qquad (4.5.47)$$

We now take the real part, set $a(t) = \exp[\mu(t)]$ and using the Ito calculus find

$$da(t) = \left(-\gamma a(t) + \frac{\varepsilon^2}{2a(t)}\right) dt + \varepsilon\Big(dW_1(t)\cos\phi(t) + dW_2(t)\sin\phi(t)]\Big) . \qquad (4.5.48)$$

The imaginary part yields

$$d\phi(t) = \frac{\varepsilon}{a(t)}\Big(- dW_1(t)\sin\phi(t) + dW_2\cos\phi(t)\Big) . \qquad (4.5.49)$$

We now define

$$\left.\begin{aligned} dW_a(t) &= dW_1(t)\cos\phi(t) + dW_2(t)\sin\phi(t) , \\ dW_\phi(t) &= -dW_1(t)\sin\phi(t) + dW_2(t)\cos\phi(t) . \end{aligned}\right\} \qquad (4.5.50)$$

We note that this is an orthogonal transformation of the kind mentioned in Sect. 4.3.6, 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

$$d\phi(t) = \frac{\varepsilon}{a(t)}dW_\phi(t) , \qquad (4.5.51a)$$

$$da(t) = \left(- \gamma a(t) + \frac{\varepsilon^2}{2a(t)}\right)dt + \varepsilon\,dW_a(t) . \qquad (4.5.51b)$$

Comment. Using the rules given in Sect. 4.4 (ii), it is possible to convert both the Cartesian equation (4.5.41a, 4.5.41b) and the polar equations (4.5.51a, 4.5.51b) 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 (4.5.47) where the term $[\varepsilon^2/2a(t)]\,dt$ 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 (4.5.49). We see the advantage of the Ito method which retains the statistical independence of $d\boldsymbol{W}(t)$ 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 (4.5.51b).

4.5.6 Multivariate Ornstein-Uhlenbeck Process

we define the process by the stochastic differential equation

$$d\boldsymbol{x}(t) = -\boldsymbol{A}\boldsymbol{x}(t)\,dt + \boldsymbol{B}\,d\boldsymbol{W}(t), \qquad (4.5.52)$$

($\boldsymbol{A}$ and $\boldsymbol{B}$ are constant matrices) for which the solution is easily obtained (as in Sect. 4.5.4):

$$x(t) = \exp(-At)x(0) + \int_0^t \exp[-A(t-t')]B\,dW(t')\,. \tag{4.5.53}$$

The mean is

$$\langle x(t)\rangle = \exp(-At)\langle x(0)\rangle. \tag{4.5.54}$$

The correlation function follows similarly

$$\begin{aligned}\langle x(t), x^{\mathrm{T}}(s)\rangle &\equiv \langle [x(t) - \langle x(t)\rangle][x(s) - \langle x(s)\rangle]^{\mathrm{T}}\rangle\,,\\ &= \exp(-At)\langle x(0), x^{\mathrm{T}}(0)\rangle \exp(-As)\\ &+ \int_0^{\min(t,s)} \exp[-A(t-t')]BB^T \exp[-A^{\mathrm{T}}(s-t')]\,dt\,.\end{aligned} \tag{4.5.55}$$

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 $\langle x(0), x^{\mathrm{T}}(0) = 0\rangle$, corresponding to a deterministic initial condition, and evaluate a few special cases.

a) The Case $AA^{\mathrm{T}} = A^{\mathrm{T}}A$: In this case (for real A) we can find a unitary matrix S such that

$$\begin{aligned}SS^\dagger &= 1\,,\\ SAS^\dagger &= SA^{\mathrm{T}}S^\dagger = \mathrm{diag}(\lambda_1, \lambda_2, \ldots \lambda_n)\,.\end{aligned} \tag{4.5.56}$$

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

$$\langle x(t), x^{\mathrm{T}}(s)\rangle = S^\dagger G(t,s)S\,, \tag{4.5.57}$$

where

$$[G(t,s)]_{ij} = \frac{(BB)^{\mathrm{T}}_{ij}}{\lambda_i + \lambda_j}[\exp(-\lambda_i|t-s|) - \exp(-\lambda_i t - \lambda_j s)]\,. \tag{4.5.58}$$

b) Stationary Variance: If A has only eigenvalues with positive real part, a stationary solution exists of the form

$$x_{\mathrm{s}}(t) = \int_{-\infty}^t \exp[-A(t-t')]B\,dW(t')\,. \tag{4.5.59}$$

We have of course

$$\langle x_{\mathrm{s}}(t)\rangle = 0\,, \tag{4.5.60}$$

$$\langle x_{\mathrm{s}}(t), x_{\mathrm{s}}^{\mathrm{T}}(s)\rangle = \int_{-\infty}^{\min(t,s)} \exp[-A(t-t')]BB^{\mathrm{T}} \exp[-A^{\mathrm{T}}(s-t')]\,dt'\,. \tag{4.5.61}$$

Let us define the stationary covariance matrix σ by

$$\sigma = \langle x_{\mathrm{s}}(t), x_{\mathrm{s}}^{\mathrm{T}}(t)\rangle\,. \tag{4.5.62}$$

This can be evaluated by means of an algebraic equation thus:

$$\begin{aligned}A\sigma + \sigma A^{\mathrm{T}} &= \int_{-\infty}^t A\exp[-A(t-t')]BB^{\mathrm{T}}\exp[-A^{\mathrm{T}}(t-t')]\,dt'\,,\\ &+ \int_{-\infty}^t \exp[-A(t-t')]BB^{\mathrm{T}}\exp[-A^{\mathrm{T}}(t-t')]A^{\mathrm{T}}\,dt'\,,\\ &= \int_{-\infty}^t \frac{d}{dt'}\{\exp[-A(t-t')]BB^{\mathrm{T}}\exp[-A^{\mathrm{T}}(t-t')]\}\,dt'\,.\end{aligned} \tag{4.5.63}$$

Carrying out the integral, we find that the lower limit vanishes by the assumed positivity of the eigenvalues of A and hence only the upper limit remains, giving

$$\boxed{A\sigma + \sigma A^{\mathrm{T}} = BB^{\mathrm{T}}\,,} \tag{4.5.64}$$

as an algebraic equation for the stationary covariance matrix.

c) Stationary Variance for Two Dimensions: We note that if A is a 2×2 matrix, it satisfies the characteristic equation

$$A^2 - (\mathrm{Tr}\,A)A + (\mathrm{Det}\,A) = 0\,, \tag{4.5.65}$$

and from (4.5.60) and the fact that (4.5.65) implies $\exp(-At)$ is a polynomial of degree 1 in A, we must be able to write

$$\sigma = \alpha BB^{\mathrm{T}} + \beta(ABB^{\mathrm{T}} + BB^{\mathrm{T}}A^{\mathrm{T}}) + \gamma ABB^{\mathrm{T}}A^{\mathrm{T}}\,. \tag{4.5.66}$$

Using (4.5.65), we find (4.5.64) is satisfied if

$$\alpha + (\mathrm{Tr}\,A)\beta - (\mathrm{Det}\,A)\gamma = 0\,, \tag{4.5.67}$$

$$2\beta(\mathrm{Det}\,A) + 1 = 0\,, \tag{4.5.68}$$

$$\beta + (\mathrm{Tr}\,A)\gamma = 0\,. \tag{4.5.69}$$

From which we have

$$\boxed{\sigma = \frac{(\mathrm{Det}\,A)BB^{\mathrm{T}} + [A - (\mathrm{Tr}\,A)]BB^{\mathrm{T}}[A - (\mathrm{Tr}\,A)]^{\mathrm{T}}}{2(\mathrm{Tr}\,A)(\mathrm{Det}\,A)}\,.} \tag{4.5.70}$$

d) Time Correlation Matrix in the Stationary State: From the solution of (4.5.60), we see that if $t > s$,

$$\begin{aligned}\langle x_{\mathrm{s}}(t), x_{\mathrm{s}}^{\mathrm{T}}(s)\rangle &= \exp[-A(t-s)] \int_{-\infty}^s \exp[-A(s-t')]BB^{\mathrm{T}}\exp[-A^{\mathrm{T}}(s-t')]\,dt'\,,\\ &= \exp[-A(t-s)]\sigma\,, \qquad t > s\,,\end{aligned} \tag{4.5.71a}$$

and similarly,

$$= \sigma\exp[-A^{\mathrm{T}}(s-t)]\,, \qquad t < s\,. \tag{4.5.71b}$$

This depends only on $s - t$, as expected of a stationary solution. Defining then

$$G_{\mathrm{s}}(t-s) = \langle x_{\mathrm{s}}(t), x_{\mathrm{s}}^{\mathrm{T}}(s)\rangle\,, \tag{4.5.72}$$

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

$$G_s(t-s) = [G_s(s-t)]^{\mathrm{T}} . \tag{4.5.73}$$

e) Spectrum Matrix in Stationary State: The spectrum matrix turns out to be rather simple. We define similarly to Sect. 1.5.2:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-i\omega\tau} G_s(\tau) d\tau , \tag{4.5.74}$$

$$= \frac{1}{2\pi} \left\{ \int_0^{\infty} \exp[-(i\omega + A)\tau]\sigma \, d\tau + \int_{-\infty}^{0} \sigma \exp[(-\mathrm{i}\omega + A^{\mathrm{T}})\tau] \, d\tau \right\} , \tag{4.5.75}$$

$$= \frac{1}{2\pi} [(A + \mathrm{i}\omega)^{-1}\sigma + \sigma(A^{\mathrm{T}} - \mathrm{i}\omega)^{-1}] . \tag{4.5.76}$$

Hence,

$$(A + \mathrm{i}\omega) S(\omega) (A^T - \mathrm{i}\omega) = \frac{1}{2\pi} (\sigma A^{\mathrm{T}} + A\sigma) , \tag{4.5.77}$$

and using (4.5.64), we get

$$\boxed{S(\omega) = \frac{1}{2\pi} (A + \mathrm{i}\omega)^{-1} BB^{\mathrm{T}} (A^{\mathrm{T}} - \mathrm{i}\omega)^{-1} .} \tag{4.5.78}$$

f) Regression Theorem: The result (4.5.71a) 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, as in (4.5.54). It is a consequence of the Markovian linear nature of the problem. The time derivative of the stationary correlation function is

$$\frac{d}{d\tau}[G_s(\tau)] \, d\tau = \frac{d}{d\tau} \langle \boldsymbol{x}_s(\tau), \boldsymbol{x}_s^{\mathrm{T}}(0) \rangle \, d\tau ,$$

$$= \langle [-A\boldsymbol{x}_s(\tau) d\tau + B \, d\boldsymbol{W}(\tau)], \boldsymbol{x}_s^{\mathrm{T}}(0) \rangle . \tag{4.5.79}$$

Since $\tau > 0$, the increment $d\boldsymbol{W}(\tau)$ is uncorrelated with $\boldsymbol{x}_s^{\mathrm{T}}(0)$, this means that

$$\boxed{\frac{d}{d\tau}[G_s(\tau)] = -A \, G_s(\tau) .} \tag{4.5.80}$$

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.

4.5.7 The General Single Variable Linear Equation

a) Homogeneous Case: We consider firstly the homogeneous case

$$dx = [b(t) \, dt + g(t) \, dW(t)] x , \tag{4.5.81}$$

and using the usual Ito rules, write

$$y = \log x , \tag{4.5.82}$$

so that

$$dy = \frac{dx}{x} - \frac{dx^2}{2x^2} = b(t) \, dt + g(t) \, dW(t) - \tfrac{1}{2} g(t^2) \, dt , \tag{4.5.83}$$

and integrating and inverting (4.5.82), we get

$$x(t) = x(0) \exp\left\{ \int_0^t \left[b(t') - \tfrac{1}{2} g(t')^2 \right] dt' + \int_0^t g(t') \, dW(t') \right\} , \tag{4.5.84}$$

$$\equiv x(0) \, \phi(t) , \tag{4.5.85}$$

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

We note that [using (4.5.15)]

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

$$= \langle [x(0)]^n \rangle \exp\left\{ n \int_0^t b(t') \, dt' + \tfrac{1}{2} n(n-1) \int_0^t g(t')^2 \, dt' \right\} . \tag{4.5.86}$$

b) Inhomogeneous Case: Now consider

$$dx = [a(t) + b(t)x] \, dt + [f(t) + g(t)x] \, dW(t) , \tag{4.5.87}$$

and write

$$z(t) = x(t)[\phi(t)]^{-1} , \tag{4.5.88}$$

with $\phi(t)$ as defined in (4.5.85) and a solution of the homogeneous equation (4.5.81). Then we write

$$dz = dx[\phi(t)]^{-1} + x \, d[\phi(t)^{-1}] + dx \, d[\phi(t)^{-1}] . \tag{4.5.89}$$

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

$$dz = \{[a(t) - f(t)g(t)] \, dt + f(t) \, dW(t)\} \phi(t)^{-1} \tag{4.5.90}$$

which is directly integrable. Hence, the solution is

$$x(t) = \phi(t) \left\{ x(0) + \int_0^t \phi(t')^{-1} \{[a(t') - f(t')g(t')] \, dt' + f(t') \, dW(t')\} \right\} . \tag{4.5.91}$$

c) Moments and Autocorrelation: It is better to derive equations for the moments from (4.5.87) rather than calculate moments and autocorrelation directly from the solution (4.5.91).

For we have

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

$$= n x(t)^{n-1} dx(t) + \tfrac{1}{2} n(n-1) x(t)^{n-2} [f(t) + g(t)x(t)]^2 \, dt . \tag{4.5.92}$$

Hence,

$$\frac{d}{dt}\langle x(t)^n\rangle = \langle x(t)^n\rangle[nb(t) + \tfrac{1}{2}n(n-1)g(t)^2]\,,$$
$$+ \langle x(t)^{n-1}\rangle[na(t) + n(n-1)f(t)g(t)]\,,$$
$$+ \langle x(t)^{n-2}\rangle\tfrac{1}{2}n(n-1)f(t)^2\,. \qquad (4.5.93)$$

These equations from a hierarchy in which the nth equation involves the solutions of the previous two, and can be integrated successively.

4.5.8 Multivariable Linear Equations

a) Homogeneous Case: The equation is

$$d\boldsymbol{x}(t) = \Big[B(t)\,dt + \sum_i G_i(t)dW_i(t)\Big]\boldsymbol{x}(t)\,, \qquad (4.5.94)$$

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.

$$\left.\begin{aligned} G_i(t)G_j(t') &= G_j(t')G_i(t)\,,\\ B(t)G_i(t') &= G_i(t')B(t)\,,\\ B(t)B(t') &= B(t')B(t)\,. \end{aligned}\right\} \qquad (4.5.95)$$

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

$$\boldsymbol{x}(t) = \Phi(t)\boldsymbol{x}(0)\,, \qquad (4.5.96)$$

with

$$\Phi(t) = \exp\Big\{\int_0^t \Big[B(t) - \tfrac{1}{2}\sum_i G_i(t)^2\Big]dt + \int_0^t\sum_i G_i(t)dW_i(t)\Big\}\,. \qquad (4.5.97)$$

b) Inhomogeneous Case: We can reduce the inhomogeneous case to the homogeneous case in exactly the same way as in one dimension. Thus, we consider

$$d\boldsymbol{x}(t) = [\boldsymbol{A}(t) + B(t)\boldsymbol{x}]\,dt\,, + \sum_i[\boldsymbol{F}_i(t) + G_i(t)\boldsymbol{x}]\,dW(t)\,, \qquad (4.5.98)$$

and write

$$\boldsymbol{y}(t) = \psi(t)^{-1}\boldsymbol{x}(t)\,, \qquad (4.5.99)$$

where $\psi(t)$ is a matrix solution of the homogeneous equation (4.5.94). We first have to evaluate $d[\psi^{-1}]$. For any matrix M we have $MM^{-1} = 1$, so, expanding to second order, $Md[M^{-1}] + dMM^{-1} + dMd[M^{-1}] = 0$.

Hence, $d[M^{-1}] = -[M + dM]^{-1}dM\,M^{-1}$ and again to second order

$$d[M^{-1}] = -M^{-1}dM\,M^{-1} + M^{-1}dM\,M^{-1}dM\,M^{-1}\,, \qquad (4.5.100)$$

and thus, since $\psi(t)$ satisfies the homogeneous equation,

$$d[\psi(t)^{-1}] = \psi(t)^{-1}\Big\{\Big[-B(t) + \sum_i G_i(t)^2\Big]dt - \sum_i G_i(t)dW_i(t)\Big\}\,, \qquad (4.5.101)$$

and, again taking differentials

$$d\boldsymbol{y}(t) = \psi(t)^{-1}\Big\{\Big[\boldsymbol{A}(t) - \sum_i G_i(t)\boldsymbol{F}_i(t)\Big]dt + \sum_i \boldsymbol{F}_i(t)dW_i(t)\Big\}\,. \qquad (4.5.102)$$

Hence,

$$\boldsymbol{x}(t) = \psi(t)\Big\{\boldsymbol{x}(0) + \int_0^t \psi(t')^{-1}\Big\{[A(t') - \sum_i G_i(t')\boldsymbol{F}_i(t')]\,dt' + \sum \boldsymbol{F}_i(t')dW_i(t')\Big\}\Big\}\,. \qquad (4.5.103)$$

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.

4.5.9 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. 4.5.6) to include time-dependent parameters, namely,

$$d\boldsymbol{x}(t) = -A(t)\boldsymbol{x}(t)\,dt + B(t)d\boldsymbol{W}(t)\,. \qquad (4.5.104)$$

This is clearly of the same form as (4.5.98) with the replacements

$$\left.\begin{array}{lcl} A(t) & \rightarrow & 0\,,\\ B(t) & \rightarrow & -A(t)\,,\\ \sum_i F_i(t)dW_i(t) & \rightarrow & B(t)\,d\boldsymbol{W}(t)\,,\\ G_i(t) & \rightarrow & 0\,. \end{array}\right\} \qquad (4.5.105)$$

The corresponding homogeneous equation is simply the deterministic equation

$$d\boldsymbol{x}(t) = -A(t)\boldsymbol{x}(t)\,dt\,, \qquad (4.5.106)$$

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

$$\boldsymbol{x}(t) = \psi(t)\boldsymbol{x}(0)\,, \qquad (4.5.107)$$

with

$$\psi(t) = \exp\Big[-\int_0^t A(t')\,dt'\Big]\,. \qquad (4.5.108)$$

Thus, applying (4.5.103),

$$\boldsymbol{x}(t) = \exp\Big[-\int_0^t A(t')\,dt'\Big]\boldsymbol{x}(0) + \int_0^t\Big\{\exp\Big[-\int_{t'}^t A(s)ds\Big]\Big\}B(t')d\boldsymbol{W}(t')\,. \qquad (4.5.109)$$

This is very similar to the solution of the time-independent Ornstein-Uhlenbeck process, as derived in Sect. 4.5.6, equation (4.5.53).

From this we have

$$\langle \boldsymbol{x}(t)\rangle = \exp\Big[-\int_0^t A(t')\,dt'\Big]\langle \boldsymbol{x}(0)\rangle\,, \tag{4.5.110}$$

$$\begin{aligned}\langle \boldsymbol{x}(t), \boldsymbol{x}^{\mathrm{T}}(t)\rangle &= \exp\Big[-\int_0^t A(t')\,dt'\Big]\langle \boldsymbol{x}(0), \boldsymbol{x}(0)^{\mathrm{T}}\rangle \exp\Big[-\int_0^t A^{\mathrm{T}}(t')\,dt'\Big] \\ &\quad + \int_0^t dt' \exp\Big[-\int_{t'}^t A(s)ds\Big]B(t')B^{\mathrm{T}}(t')\exp\Big[-\int_t^t A^{\mathrm{T}}(s)ds\Big].\end{aligned} \tag{4.5.111}$$

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 the next two chapters, 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. This chapter is devoted mainly to single variable systems, since there are a large number of exact results for single variable systems, which makes the separate treatment of such systems appropriate. The next chapter deals with the more general multivariable aspects of many of the same issues treated one-dimensionally in this chapter.

The construction of appropriate boundary conditions is of fundamental importance, and is carried out in Sect. 5.1 in a form applicable to both one-variable and many-variable systems. A corresponding treatment for the boundary conditions on the backward Fokker-Planck equation is given in Sect. 5.1.2. The remaining of the chapter is 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.

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. 4.3.5). 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, 5.2] 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 whom this book is addressed.