Appendix B

Stochastic differential equations

B.1 Gaussian white noise

Although the description 'stochastic differential equation' (SDE) sounds rather general, it is usually taken to refer only to differential equations with a Gaussian white-noise term. In this appendix we begin by reviewing SDEs of this sort, which are also known as Langevin equations. In the final section, we generalize to other sorts of noise (in particular jumps). This review is intended not to be mathematically rigorous, but rather to build intuition about the physical assumptions behind the formalism. In particular, the concept of stochastic integration will not be introduced at all. A more formal treatment of SDEs and stochastic integrals, still aimed at physical scientists rather than mathematicians, can be found in Ref. [Gar85]. Another more elementary introduction may be found in Refs. [Gil93, Gil96].

Consider the one-dimensional case for simplicity. A SDE for the random variable X may then be written as

$$\dot{X} = \alpha(X) + \beta(X)\xi(t). \tag{B.1}$$

Here, the time argument of X has been omitted, α and β are arbitrary real functions, and $\xi(t)$ is a rapidly varying random process. This process, referred to as noise, is continuous in time, has zero mean and is a stationary process. The last descriptor means that all of its statistics, including in particular its correlation function

$$\mathrm{E}[\xi(t)\xi(t+\tau)], \tag{B.2}$$

are independent of t. Here E denotes an ensemble average or expectation value as usual. The noise is normalized such that its correlation function integrates to unity:

$$\int_{-\infty}^{\infty} \mathrm{d}\tau\, \mathrm{E}[\xi(t)\xi(t+\tau)] = 1. \tag{B.3}$$

Note that Eq. (B.3) implies that $[\alpha] = [X]T^{-1}$ and $[\beta] = [X]T^{-1/2}$, where here $[A]$ denotes the dimensionality of A and T is the time dimension.

We are interested in the case of Markovian SDEs, for which the correlation time of the noise must be zero. That is, we can replace Eq. (B.3) by

$$\int_{-\epsilon}^{\epsilon} \mathrm{d}\tau\, \mathrm{E}[\xi(t)\xi(t+\tau)] = 1, \tag{B.4}$$

https://doi.org/10.1017/CBO9780511813948.010 Published online by Cambridge University Press

for all $\epsilon > 0$. In this limit, $\xi(t)$ is called *Gaussian white noise*, which is completely characterized by the two moments

$$\mathrm{E}[\xi(t)\xi(t')] = \delta(t - t'), \tag{B.5}$$

$$\mathrm{E}[\xi(t)] = 0. \tag{B.6}$$

The correlation function contains a singularity at $t = t'$ because of the constraint of Eq. (B.4). Because of this singularity, one has to be very careful in finding the solutions of Eq. (B.1). The noise $\xi(t)$ is called *white* because the spectrum is flat in this limit, just like the spectrum of white light is flat (in the visible range of frequencies anyway). Recall that the spectrum of a noise process is the Fourier transform of the correlation function.

Physically, an equation like (B.1) could be obtained by deriving it for a physical (non-white) noise source $\xi(t)$, and then taking the idealized limit. In that case, Eq. (B.1) is known as a *Stratonovich* SDE. This result is known as the Wong–Zakai theorem [WZ65]. The Stratonovich SDE for some function f of X is found by using the standard rules of differential calculus, that is,

$$\dot{f}(X) = f'(X)[\alpha(X) + \beta(X)\xi(t)], \tag{B.7}$$

where the prime denotes differentiation with respect to X. As stated above, the differences from standard calculus arise when actually solving Eq. (B.1).

Let $X(t)$ be known, and equal to x. If one were to assume that the infinitesimally evolved variable X were given by

$$X(t + \mathrm{d}t) = x + [\alpha(x) + \beta(x)\xi(t)]\mathrm{d}t \tag{B.8}$$

and, further, that the stochastic term $\xi(t)$ were independent of the system at the same time, then one would derive the expected increment in X from t to $t + \mathrm{d}t$ to be

$$\mathrm{E}[\mathrm{d}X] = \alpha(x)\mathrm{d}t. \tag{B.9}$$

The second assumption here seems perfectly reasonable since the noise is not correlated with any of the noise which has interacted with the system in the past, and so would be expected to be uncorrelated with the system. Applying the same arguments to f yields

$$\mathrm{E}[\mathrm{d}f] = f'(x)\alpha(x)\mathrm{d}t. \tag{B.10}$$

That is, all expectation values are independent of β. In particular, if we consider $f(X) = X^2$, then the above imply that the infinitesimal increase in the variance of X is

$$\mathrm{E}[\mathrm{d}(X^2)] - \mathrm{d}(\mathrm{E}[X])^2 = 2x\alpha(x)\mathrm{d}t - 2x\alpha(x)\mathrm{d}t = 0. \tag{B.11}$$

That is to say, the stochastic term has not introduced any noise into the variable X.

Obviously this result is completely contrary to what one would wish from a stochastic equation. The lesson is that it is invalid to make simultaneously the following three assumptions.

1. The chain rule of standard calculus applies (Eq. (B.7)).
2. The infinitesimal increment of a quantity is equal to its rate of change multiplied by $\mathrm{d}t$ (Eq. (B.8)).
3. The noise and the system at the same time are independent.

With a Stratonovich SDE the first assumption is true, and the usual explanation [Gar85] is that the second is also true but that the third assumption is false. Alternatively (and this is the interpretation we adopt), one can characterize a Stratonovich SDE by saying that the second assumption is false (or true only in an implicit way) and that the third is still true.

In this way of looking at things, the fluxion $\dot{X}$ in a Stratonovich SDE is just a symbol that can be manipulated using the usual rules of calculus. It should not be turned into a ratio of differentials $\mathrm{d}X/\mathrm{d}t$. In particular, $\mathrm{E}[\dot{X}]$ is not equal to $\mathrm{d}\mathrm{E}[X]/\mathrm{d}t$ in general. This point of view is useful for later generalization to jump processes in Section B.6, where one can still consider starting with an SDE containing non-singular noise, and then taking the singular limit. In the jump case, the third assumption is inapplicable, so the problem must lie with the second assumption. Since the term Stratonovich is restricted to the case of Gaussian white noise, we will also use a more general terminology, referring to any SDE involving $\dot{X}$ as an *implicit* equation.

A different choice of which postulates to relax is that of the Itô stochastic calculus. With an Itô SDE, the first assumption above is false, the second is true in an explicit manner and the third is also true (for Gaussian white noise, but not for jumps). The Itô form has the advantage that it simply allows the increment in a quantity to be calculated, and also allows ensemble averages to be taken easily. It has the disadvantage that one cannot use the usual chain rule.

B.2 Itô stochastic differential calculus

Because different rules of calculus apply to the Itô and Stratonovich forms of a SDE, the equations will appear differently in general. The Itô form of the Stratonovich equation (B.1) is

$$\mathrm{d}X = [\alpha(X) + \tfrac{1}{2}\beta(X)\beta'(X)]\mathrm{d}t + \beta(X)\mathrm{d}W(t). \tag{B.12}$$

Here, the infinitesimal Wiener increment has been introduced, defined by

$$\mathrm{d}W(t) = \xi(t)\mathrm{d}t. \tag{B.13}$$

This is called a Wiener increment because if we define

$$W(t) = \int_{t_0}^{t} \xi(t')\mathrm{d}t' \tag{B.14}$$

then this has all of the properties of a Wiener process. That is, if we define $\Delta W(t) = W(t+\Delta t) - W(t)$, then this is independent of $W(s)$ for $s < t$, and has a Gaussian distribution with zero mean and variance Δt:

$$\Pr[\Delta W(t) \in (w, w+\mathrm{d}w)] = [2\pi\,\Delta t]^{-1/2} \exp\!\big[-w^2/(2\,\Delta t)\big]\mathrm{d}w. \tag{B.15}$$

It is actually quite easy to see these results. First the independence of $\Delta W(t)$ from $W(s)$ for $s < t$ follows simply from Eq. (B.5). Second, it is easy to show that

$$\mathrm{E}\big[\Delta W(t)^2\big] = \Delta t, \tag{B.16}$$

$$\mathrm{E}[\Delta W(t)] = 0. \tag{B.17}$$

Exercise B.1 *Verify these using Eqs. (B.5) and (B.6).*

To go from these moments to the Gaussian distribution (B.15), note that, for any finite time increment Δt, the Wiener increment ΔW is the sum of an infinite number of independent noises $\xi(t)\mathrm{d}t$, which are identically distributed. By the central limit theorem [Gil83], since the sum of the variances is finite, the resulting distribution is Gaussian. We thus see why $\xi(t)$ was called Gaussian white noise: because of the Gaussian probability distribution of the Wiener process. Note that the Wiener process is not differentiable, so

https://doi.org/10.1017/CBO9780511813948.010 Published online by Cambridge University Press

strictly $\xi(t)$ does not exist. This is another way of seeing why stochastic calculus is a tricky business and why we have to worry about the Itô versus Stratonovich definitions.

In Eq. (B.12) we have introduced a convention of indicating Itô equations by an explicit representation of an infinitesimal increment (as on the left-hand side of Eq. (B.12)), whereas Stratonovich equations will be indicated by an implicit equation with a fluxion on the left-hand side (as in Eq. (B.1)). If an Itô (or *explicit*) equation is given as

$$\mathrm{d}X = a(X)\mathrm{d}t + b(X)\mathrm{d}W(t), \tag{B.18}$$

then the corresponding Stratonovich equation is

$$\dot{X} = a(X) - \tfrac{1}{2}b'(X)b(X) + b(X)\xi(t). \tag{B.19}$$

Here the prime indicates differentiation with respect to X. A simple non-rigorous derivation of this relation will be given in Section B.3.

In the Itô form, the noise is independent of the system, so the expected increment in X from Eq. (B.18) is simply

$$\mathrm{E}[\mathrm{d}X] = a(X)\mathrm{d}t. \tag{B.20}$$

However, the nonsense result (B.11) is avoided because the chain rule does not apply to calculating $\mathrm{d}f(X)$. The actual increment in $f(X)$ is simple to calculate by using a Taylor expansion for $f(X + \mathrm{d}X)$. The difference from the usual chain rule is that second-order infinitesimal terms cannot necessarily be ignored. This arises because the noise is so singular that second-order noise infinitesimals are as large as first-order deterministic infinitesimals. Specifically, the infinitesimal Wiener increment $\mathrm{d}W(t)$ can be assumed to be defined by the following Itô rules:

$$\mathrm{E}[\mathrm{d}W(t)^2] = \mathrm{d}t, \tag{B.21}$$

$$\mathrm{E}[\mathrm{d}W(t)] = 0. \tag{B.22}$$

These can be obtained from Eqs. (B.16) and (B.17) simply by taking the infinitesimal limit $\Delta \to d$.

Note that there is actually no restriction that $\mathrm{d}W(t)$ must have a Gaussian distribution. As long as the above moments are satisfied, the increment $\Delta W(t)$ over any finite time will be Gaussian from the central limit theorem. By a similar argument, it is actually possible to omit the expectation value in Eq. (B.21) because, over any finite time, a time average effects an ensemble average of what is primarily a deterministic rather than stochastic quantity. This can be seen as follows. Consider the variable

$$\Delta\tau = \sum_{j=0}^{N-1} [\delta W(t_j)]^2, \tag{B.23}$$

where $t_j = t_0 + j\,\delta t$, where $\delta t = \Delta t/N$. Then it follows that

$$\langle \Delta\tau \rangle = \Delta t, \tag{B.24}$$

$$\sqrt{\langle (\Delta\tau)\rangle^2 - \langle \Delta\tau \rangle^2} = \frac{\Delta t}{\sqrt{N/2}}. \tag{B.25}$$

Exercise B.2 *Show these results.*
Hint: *For the second of these, first show that* $\left\langle [\delta W(t_j)]^2 [\delta W(t_k)]^2 \right\rangle = (\delta t)^2(1 + 2\delta_{jk})$. *Remember that the* $\delta W(t_j)$ *and* $\delta W(t_k)$ *are independent for* $k \neq j$*, while, for* $j = k$*, use the fact that, for a Gaussian random variable* X *of mean* $0, \left\langle X^4 \right\rangle = 3\left\langle X^2 \right\rangle^2$.

In the limit $N \to \infty$, where $\delta t \to \mathrm{d}t$, the standard deviation in $\Delta\tau$ vanishes and $\Delta\tau$ converges to Δt in the mean-square sense. Since this is true for any finite time interval, we may as well replace $\mathrm{d}W^2$ by $\mathrm{d}t$.

Using this result and expanding the Taylor series to second order gives the modified chain rule

$$\mathrm{d}f(X) = f'(X)\mathrm{d}X + \tfrac{1}{2}f''(X)(\mathrm{d}X)^2. \tag{B.26}$$

Specifically, with $\mathrm{d}X$ given by Eq. (B.18), and using the rule $\mathrm{d}W(t)^2 \equiv \mathrm{d}t$,

$$\mathrm{d}f(X) = \left[f'(X)a(X) + \tfrac{1}{2}f''(X)b(X)^2\right]\mathrm{d}t + f'(X)b(X)\mathrm{d}W(t). \tag{B.27}$$

With this definition, and with $f(X) = X^2$, one finds that the expected increase in the variance of X in a time $\mathrm{d}t$ is

$$\mathrm{E}[\mathrm{d}X(t)^2] - \mathrm{d}(\mathrm{E}[X(t)])^2 = b(x)^2\,\mathrm{d}t. \tag{B.28}$$

That is to say, the effect of the noise is to increase the variance of X. Thus, the correct use of the stochastic calculus evades the absurd result of Eq. (B.11).

B.3 The Itô–Stratonovich relation

Consider again the Stratonovich equation, with the (boring) deterministic term set to zero:

$$\dot{X} = \beta(X)\xi(t). \tag{B.29}$$

Assuming that the chain rule of standard calculus applies, and that the noise at time t is independent of the system at that time, we have shown that naively turning this from an equation for the rate of change of X into an equation for the increment of X,

$$X(t + \mathrm{d}t) = X(t) + \beta(X)\xi(t)\mathrm{d}t, \tag{B.30}$$

leads to absurd results in general. This is because, when the noise $\xi(t)\mathrm{d}t$ is as singular as we are assuming (scaling as $\sqrt{\mathrm{d}t}$, rather than $\mathrm{d}t$), even an infinitesimal time increment cannot be assumed to yield a change of size scaling as $\mathrm{d}t$.

Since Eq. (B.30) comes from a first-order Taylor expansion of $X(t + \mathrm{d}t)$ in $\mathrm{d}t$, it makes sense from the above arguments that we should use a higher-order expansion. The all-order expansion is

$$X(t + \mathrm{d}t) = \exp\left(\mathrm{d}t\,\frac{\partial}{\partial s}\right)X(s)\bigg|_{s=t}. \tag{B.31}$$

We can evaluate this by rewriting Eq. (B.29) as

$$\frac{\partial}{\partial s}X(s)\bigg|_{s=t} = \beta(X(s))\xi(t)|_{s=t}. \tag{B.32}$$

Note that $\xi(t)$ is assumed constant while $X(s)$ changes. This is an expression of the fact that the noise $\xi(t)$ cannot in reality be δ-correlated. As emphasized above, equations of the Stratonovich form arise naturally only when $\xi(t)$ is a physical (non-white) noise source, and the idealization to white noise is made later. Thus the physical noise will have some finite correlation time over which it remains relatively constant. This idealization is valid if the physical correlation time is much smaller than the characteristic evolution time of the system.

We now expand Eq. (B.31) to second order in $\mathrm{d}t$. As in the Itô chain rule, this is all that is necessary. The result, using Eq. (B.32), is

$$X(t+\mathrm{d}t) = X(t) + \mathrm{d}t\,\beta(X(t))\xi(t) + \frac{1}{2}(\mathrm{d}t)^2\left[\frac{\partial}{\partial s}\beta(X(s))\xi(t)\right]_{s=t}. \tag{B.33}$$

Now, using the usual chain rule to expand $\partial\beta(X(s))/\partial s$, again using Eq. (B.32), we get

$$\mathrm{d}X = \beta(X)\xi(t)\mathrm{d}t + \frac{1}{2}(\xi(t)\mathrm{d}t)^2\beta(X)\beta'(X). \tag{B.34}$$

Replacing $\xi(t)\mathrm{d}t$ by $\mathrm{d}W(t)$ and using Eq. (B.21) yields the correct Itô equation (B.12).

In cases for which β is linear in X, as in

$$\dot{X}(t) = \lambda X(t)\xi(t), \tag{B.35}$$

the Itô equation

$$\mathrm{d}X(t) = [\lambda\,\mathrm{d}W(t) + (\lambda^2/2)\mathrm{d}t]X(t) \tag{B.36}$$

can be found easily since Eq. (B.31) becomes

$$X(t+\mathrm{d}t) = \exp(\lambda\,\mathrm{d}W(t))X(t). \tag{B.37}$$

This case is particularly relevant in quantum systems.

B.4 Solutions to SDEs

B.4.1 The meaning of 'solution'

Because the equations we are considering are stochastic, they have no simple solution as for deterministic differential equations, as a single number that changes with time. Rather, there are infinitely many solutions, depending on which noise $\xi(t)$ actually occurs. It might seem that this is more of a *problem* than a *solution*, since it is not easy to characterize such an infinite ensemble in general. However, this infinite ensemble of solutions has definite *statistical* properties, because the noise $\xi(t)$ has definite statistical properties. For example, the *moments* $\mathrm{E}[X(t)]$ and $\mathrm{E}\left[X(t)^2\right]$ are *deterministic functions of time*, as is the correlation function $\mathrm{E}[X(t)X(t+\tau)]$.

To find averages such as these, in general a stochastic numerical solution is required. That is, the SDE is solved for one particular realization of the noise and the result $X(t)$ recorded. It is then solved again for a different (and independent) realization of the noise. Any given moment can then be approximated by the finite ensemble average F. For example, the one-time average $\mathrm{E}\left[f\left(X(t)\right)\right]$, at any particular time t, can be estimated from

$$\mathrm{F}\left[f\left(X(t)\right)\right] \equiv \frac{1}{M}\sum_{j=1}^{M} f\left(X_j(t)\right), \tag{B.38}$$

where $X_j(t)$ is the solution from the jth run and M is the total number of runs. The error in the estimate $\mathrm{F}\left[f\left(X(t)\right)\right]$ can be estimated by the usual statistical formula

$$\sigma\left\{\mathrm{F}\left[f\left(X(t)\right)\right]\right\} = \sqrt{\frac{\mathrm{F}\left[[f\left(X(t)\right)]^2\right] - \mathrm{F}\left[f\left(X(t)\right)\right]^2}{M}}. \tag{B.39}$$

https://doi.org/10.1017/CBO9780511813948.010 Published online by Cambridge University Press

Thus M has to be chosen large enough for this to be below some acceptable level. Two-time averages such as correlation functions, and the uncertainties in these estimates, may be determined in a similar way.

B.4.2 Itô versus Stratonovich

The existence of two forms of the same SDE, Itô and Stratonovich, may seem problematic at this point. Which one is actually used to solve SDEs? The answer depends on the method of solution.

Using a simple Euler step method, the Itô SDE giving an explicit increment is appropriate. That is, for the one-dimensional example, one has

$$X(t_{j+1}) = X(t_j) + a\big(X(t_j)\big)\delta t + b\big(X(t_j)\big)\sqrt{\delta t}\, S_j, \tag{B.40}$$

where δt is a very small increment, with $t_{j+1} - t_j = \delta t$, and S_j is a random number with a standard normal distribution[1] generated by the computer for this time step. The S_{j+1} for the next time step is a new number, and the numbers in one run should be independent of those in any other run. If one were to use a more sophisticated integration routine than the Euler one, then the Stratonovich equation may be the one needed. See Ref. [KP00] for a discussion.

In some cases, it is possible to obtain analytical solutions to a SDE. By this we mean a closed integral form. Of course, this integral will not evaluate to a number, because it will contain the noise term $\xi(t)$. However, it can be manipulated so as to give moments easily. Again, the question arises, which equation is actually integrated in these cases, the Itô one or the Stratonovich one? Here the answer is that in practice it does not matter. The only cases in which an analytical solution is possible are those in which the Itô equation has been (perhaps by an appropriate change of variable) put in the form

$$\mathrm{d}X = a(t)\mathrm{d}t + b(t)\mathrm{d}W, \tag{B.41}$$

that is, where a and b are not functions of X. In this case the Stratonovich equation is

$$\dot{X} = a(t) + b(t)\xi(t). \tag{B.42}$$

That is, it looks the same as the Itô equation, so one could naively integrate it instead, to obtain the solution

$$X(t) = X(0) + \int_0^t a(s)\mathrm{d}s + \int_0^t b(s)\xi(s)\mathrm{d}s. \tag{B.43}$$

B.5 The connection to the Fokker–Planck equation

An alternative to describing a stochastic process using a SDE for X is to use a Fokker–Planck equation (FPE). This is an evolution equation for the probability distribution $\wp(x)$ for the variable. In this section we show how the FPE corresponding to a SDE can very easily be derived. In the process we obtain other results that are used in the main text.

First note that the probability density for a continuous variable X is by definition

$$\wp(x) = \mathrm{E}[\delta(X - x)]. \tag{B.44}$$

[1] That is, a Gaussian distribution with mean zero and variance unity.

https://doi.org/10.1017/CBO9780511813948.010 Published online by Cambridge University Press

Now $\delta(X - x)$ is just a function of X, so we can consider the SDE it obeys. If X obeys

$$\mathrm{d}X = a(X)\mathrm{d}t + b(X)\mathrm{d}W(t) \tag{B.45}$$

then, using the Itô chain rule, one obtains

$$\begin{aligned} \mathrm{d}\delta(X - x) &= \left[\frac{\partial}{\partial X}\delta(X - x)\right][a(X)\mathrm{d}t + b(X)\mathrm{d}W(t)] \\ &\quad + \left[\frac{1}{2}\,\frac{\partial^2}{(\partial X)^2}\delta(X - x)\right] b(X)^2\,\mathrm{d}t, \qquad \text{(B.46)} \\ &= \left[-\frac{\partial}{\partial x}\delta(X - x)\right][a(X)\mathrm{d}t + b(X)\mathrm{d}W(t)] \\ &\quad + \left[\frac{1}{2}\,\frac{\partial^2}{(\partial x)^2}\delta(X - x)\right] b(X)^2\mathrm{d}t. \qquad \text{(B.47)} \end{aligned}$$

Exercise B.3 *Convince yourself that, for an arbitrary smooth function $f(X)$,*

$$\left[\frac{\partial}{\partial x}\delta(X - x)\right] f(X) = \frac{\partial}{\partial x}[\delta(X - x)f(x)]. \tag{B.48}$$

Hint: *Consider the first-principles definition of a differential.*

Using the result of this exercise and its generalization to second derivatives, and then taking the expectation value over X, gives

$$\mathrm{d}\wp(x) = \left\{-\frac{\partial}{\partial x}[a(x)\mathrm{d}t + b(x)\mathrm{d}W(t)] + \frac{1}{2}\,\frac{\partial^2}{(\partial x)^2}b(x)^2\,\mathrm{d}t\right\}\wp(x). \tag{B.49}$$

If $\wp(x;t) = \delta(X(t) - x)$ at some time, then by construction this will remain true for all times by virtue of the stochastic equation (B.49). However, this equation (which we call a stochastic FPE) is more general than the SDE (B.45), insofar as it allows for initial uncertainty about X. Moreover, it allows the usual FPE to be obtained by assuming that we do not know the particular noise process $\mathrm{d}W$ driving the stochastic evolution of X and $\wp(x)$. Replacing $\mathrm{d}W$ in Eq. (B.49) by its expectation value gives the (deterministic) FPE

$$\dot{\wp}(x) = \left\{-\frac{\partial}{\partial x}a(x) + \frac{1}{2}\,\frac{\partial^2}{(\partial x)^2}b(x)^2\right\}\wp(x). \tag{B.50}$$

Note that this $\wp(x)$ is *not* the same as that appearing in Eq. (B.49) because we are no longer conditioning the distribution upon knowledge of the noise process. In Eq. (B.50), the term involving first derivatives is called the drift term and that involving second derivatives the diffusion term.

B.6 More general noise

As we have noted, our characterization of the Itô–Stratonovich distinction as an explicit–implicit distinction is not standard. Its advantage becomes evident when one considers point-process noise. Recall that, when one starts with evolution driven by physical noise and then idealizes this as Gaussian white noise, one ends up with a Stratonovich equation, which has to be converted into an Itô equation in order to find an

https://doi.org/10.1017/CBO9780511813948.010 Published online by Cambridge University Press

explicit solution. Similarly, if one has an equation driven by physical noise that one then idealizes as a point process (that is, a time-series of δ-functions), one also ends up with an implicit equation that one has to make explicit. The implicit–explicit relation is more general than the Itô–Stratonovich one for two reasons. First, for point-process noise the defining characteristic of an Itô equation, namely that the stochastic increment is independent of the current values of the system variables, need not be true. Secondly, when feedback is considered, this Itô rule fails even for Gaussian white noise. That is because the noise which is fed back is necessarily correlated with the system at the time it is fed back, and cannot be decorrelated by invoking Itô calculus.

Although point-process noise may be non-white (that is, it need not have a flat noise spectrum), it must still have an infinite bandwidth. If the correlation function for the noise were a smooth function of time, then there would be no need to use any sort of stochastic calculus; the normal rules of calculus would apply. But, for any noise with a singular correlation function, it is appropriate to make the implicit–explicit distinction. We write a general explicit equation (in one dimension) as

$$\mathrm{d}X = k(X)\mathrm{d}M(t). \tag{B.51}$$

Here, deterministic evolution is being ignored, so $\mathrm{d}M(t)$ is some stochastic increment. If $\mathrm{d}M(t) = \mathrm{d}W(t)$ then Eq. (B.51) is an Itô SDE. More generally, $\mathrm{d}M(t)$ will have well-defined moments that may depend on the system $X(t)$. A stochastic calculus will be necessary if second- or higher-order moments of $\mathrm{d}M(t)$ are not of second or higher order in $\mathrm{d}t$. For Gaussian white noise, only the second-order moments fit this description, with $\mathrm{d}W(t)^2 = \mathrm{d}t$. In contrast, all moments must be considered for a point-process increment $\mathrm{d}M(t) = \mathrm{d}N(t)$.

The point-process increment can be defined by

$$\mathrm{E}[\mathrm{d}N(t)] = \lambda(X)\mathrm{d}t, \tag{B.52}$$

$$\mathrm{d}N(t)^2 = \mathrm{d}N(t). \tag{B.53}$$

Here $\lambda(X)$ is a positive function of the random variable X (here assumed known at time t). Equation (B.52) indicates that the mean of $\mathrm{d}N(t)$ is of order $\mathrm{d}t$ and may depend on the system. Equation (B.53) simply states that $\mathrm{d}N(t)$ equals either zero or one, which is why it is called a point process. From the stochastic evolution it generates it is also known as a jump process. Because $\mathrm{d}N$ is infinitesimal (at least in its mean), we can say that all second- and higher-order products containing $\mathrm{d}t$ are $o(\mathrm{d}t)$. This notation means that such products (like $\mathrm{d}N\,\mathrm{d}t$, but not $\mathrm{d}N^2$) are negligible compared with $\mathrm{d}t$. Obviously all moments of $\mathrm{d}N(t)$ are of the same order as $\mathrm{d}t$, so the chain rule for $f(X)$ will completely fail.

Unlike $\mathrm{d}W$, which is independent of the system at the same time, $\mathrm{d}N$ does depend on the system, at least statistically, through Eq. (B.52). In fact, we can use the above equations to show that

$$\mathrm{E}[\mathrm{d}N(t)f(X)] = \mathrm{E}[\lambda(X)f(X)]\mathrm{d}t, \tag{B.54}$$

for some arbitrary function f.

Exercise B.4 *Convince yourself of this.*

It turns out [Gar85] that for Markovian processes it is sufficient to consider only the above two cases, $\mathrm{d}M = \mathrm{d}W$ and $\mathrm{d}M = \mathrm{d}N$.

Now consider a SDE that, like the Stratonovich equation for Gaussian white noise, arises from a physical process in which the singularity of the noise is an idealization. Such

https://doi.org/10.1017/CBO9780511813948.010 Published online by Cambridge University Press

an equation would be written, using our convention, as

$$\dot{X} = \chi(X)\mu(t), \tag{B.55}$$

where $\mu(t)$ is a noisy function of time that is idealized by

$$\mu(t) = \mathrm{d}M(t)/\mathrm{d}t. \tag{B.56}$$

Equation (B.55) is an *implicit* equation in that it gives the increment in X only implicitly. It has the advantage that $f(X)$ would obey an implicit equation as given by the usual chain rule,

$$\dot{f}(X) = f'(X)\chi(X)\mu(t). \tag{B.57}$$

Notice that the third distinction between Itô and Stratonovich calculus, namely that based on the independence of the noise term and the system at the same time, has not entered this discussion. This is because, even in the explicit equation (B.51), the noise may depend on the system. The independence condition is simply a peculiarity of Gaussian white noise. The implicit–explicit distinction is more general than the Stratonovich–Itô distinction. As we will show below, the relationship between the Stratonovich and Itô SDEs can be easily derived within this more general framework.

The general problem is to find the explicit form of an implicit SDE with arbitrary noise. For implicit equations, the usual chain rule (B.57) applies, and can be rewritten

$$\dot{f} = f'(X)\chi(X)\mu(t) \equiv \phi\big(f(X)\big)\mu(t), \tag{B.58}$$

where this equation defines $\phi(f)$. Now, in order to solve Eq. (B.55), it is necessary to find an explicit expression for the increment in X. The correct answer may be found by expanding the Taylor series to all orders in $\mathrm{d}M$. This can be written formally as

$$X(t+\mathrm{d}t) = \exp\left(\mathrm{d}t\,\frac{\partial}{\partial s}\right)X(s)|_{s=t} \tag{B.59}$$

$$= \exp\left[\chi(x)\mathrm{d}M(t)\frac{\partial}{\partial x}\right]x|_{x=X(t)}. \tag{B.60}$$

Here we have used the relation

$$\left[\frac{\mathrm{d}}{\mathrm{d}s}X(s) = \chi\left(X(s)\right)\frac{\mathrm{d}M(t)}{\mathrm{d}t}\right]_{s=t}, \tag{B.61}$$

which is the explicit meaning of the implicit Eq. (B.55). Note that $\mu(t)$ is assumed to be constant, while $X(s)$ is evolved, for the same reasons as explained following Eq. (B.32). If the noise $\mu(t)$ is the limit of a physical process (which is the limit for which Eq. (B.55) is intended to apply), then it must have some finite correlation time over which it remains relatively constant. The noise can be considered δ-correlated if that time can be considered to be infinitesimal compared with the characteristic evolution time of the system X.

The explicit SDE is thus defined to be

$$\mathrm{d}X(t) = \left(\exp\left[\chi(X)\mathrm{d}M(t)\frac{\partial}{\partial X}\right] - 1\right)X(t), \tag{B.62}$$

which means

$$\mathrm{d}X(t) = \left(\exp\left[\chi(x)\mathrm{d}M(t)\frac{\partial}{\partial x}\right] - 1\right)x|_{x=X(t)}. \tag{B.63}$$

https://doi.org/10.1017/CBO9780511813948.010 Published online by Cambridge University Press

This expression will converge for all $\chi(X)$ for $\mathrm{d}M = \mathrm{d}N$ or $\mathrm{d}M = \mathrm{d}W$, and is compatible with the chain-rule requirement (B.58) for the implicit form. This can be seen from calculating the increment in $f(X)$ using the explicit form:

$$\begin{aligned}
\mathrm{d}f &= f\big(X(t) + \mathrm{d}X(t)\big) - f\big(X(t)\big) \\
&= f\left(\exp\left[\chi(x)\mathrm{d}M(t)\frac{\partial}{\partial x}\right]x|_{x=X(t)}\right) - f\big(X(t)\big) \\
&= \exp\left[\chi(x)\mathrm{d}M(t)\frac{\partial}{\partial x}\right]f(x)|_{x=X(t)} - f\big(X(t)\big) \\
&= \left(\exp\left[\phi(f)\mathrm{d}M(t)\frac{\partial}{\partial f}\right] - 1\right)f|_{f=f(X(t))},
\end{aligned} \tag{B.64}$$

as expected from Eq. (B.58). This completes the justification for Eq. (B.62) as the correct explicit form of the implicit Eq. (B.55).

For deterministic processes ($\chi = 0$), there is no distinction between the explicit and implicit forms, since only the first-order expansion of the exponential remains with $\mathrm{d}t$ infinitesimal. There is also no distinction if $\chi(x)$ is a constant. For Gaussian white noise, the formula (B.62) is the rule given in Section B.3 for converting from Stratonovich to Itô form. That is, if the Stratonovich SDE is Eq. (B.55) with $\mathrm{d}M(t) = \mathrm{d}W(t)$, then the Itô SDE is

$$\mathrm{d}X(t) = \chi(X)\mathrm{d}W(t) + \tfrac{1}{2}\chi(X)\chi'(X)\mathrm{d}t. \tag{B.65}$$

Exercise B.5 *Show this, using the Itô rule* $\mathrm{d}W(t)^2 = \mathrm{d}t$.

This rule implies that it is necessary to expand the exponential only to second order. This fact makes the inverse transformation (Itô to Stratonovich) easy. For the jump process, the rule $\mathrm{d}N(t)^2 = \mathrm{d}N(t)$ means that the exponential must be expanded to all orders. This gives

$$\mathrm{d}X(t) = \mathrm{d}N(t)\left(\exp\left[\chi(x)\frac{\partial}{\partial X}\right] - 1\right)x(t). \tag{B.66}$$

In this case, the inverse transformation would not be easy to find in general, but there seems no physical motivation for requiring it.

B.6.1 Multi-dimensional generalization

The multi-dimensional generalization of the above formulae is obvious. Writing X_i for the componenents of the vector $\vec{X}$ and using the Einstein summation convention, if the implicit form is

$$\dot{X}_i(t) = \chi_{ij}\big(\vec{X}(t)\big)\mu_j(t), \tag{B.67}$$

then the explicit form is

$$\mathrm{d}X_i(t) = \left(\exp\left[\chi_{kj}(\vec{X})\mathrm{d}M_j(t)\frac{\partial}{\partial X_k}\right] - 1\right)X_i(t). \tag{B.68}$$

https://doi.org/10.1017/CBO9780511813948.010 Published online by Cambridge University Press

This is quite complicated in general. Fortunately, when considering quantum feedback processes, the equations for the state are linear. Thus, if one has the implicit equation

$$\dot{\rho}(t) = \mu(t)\mathcal{K}\rho(t), \tag{B.69}$$

where $\mathcal{K}$ is a Liouville superoperator, then the explicit SDE is simply

$$\mathrm{d}\rho(t) = (\exp[\mathcal{K}\,\mathrm{d}M(t)] - 1)\rho(t). \tag{B.70}$$

Exercise B.6 *Convince yourself that this is consistent with Eq. (B.68).*