4. Fokker-Planck Equation

As shown in Sects. 3.1, 2 we can immediately obtain expectation values for processes described by the linear Langevin equations (3.1, 31). For nonlinear Langevin equations (3.67, 110) expectation values are much more difficult to obtain, so here we first try to derive an equation for the distribution function. As mentioned already in the introduction, a differential equation for the distribution function describing Brownian motion was first derived by *Fokker* [1.1] and *Planck* [1.2]: many review articles and books on the Fokker-Planck equation now exist [1.5 – 15].

Our derivation starts with an expansion of the distribution function, known as Kramers-Moyal expansion [1.17, 19]. In this equation, only the Kramers-Moyal coefficients (3.95, 118 – 120) will enter. As seen in Sect. 3.3, these Kramers-Moyal coefficients can also be calculated for the nonlinear Langevin equations. As it turned out, these coefficients vanish for $n \geqq 3$ for the Langevin equations (3.67, 110) with δ-correlated Gaussian-distributed Langevin forces, and only the drift and diffusion coefficients (3.107, 108, 118, 119) enter in the distribution function equation. Hence the Kramers-Moyal expansion with an infinite number of terms stops after the second term. This equation is then the Fokker-Planck equation or the forward Kolmogorov equation.

The problem of obtaining averages is thus reduced to the problem of solving this Fokker-Planck equation. For pedagogic reasons we first treat the one-variable case and then the more complicated case of N variables.

4.1 Kramers-Moyal Forward Expansion

It follows from the definition of the transition probability (2.69) that the probability density $W(x, t+\tau)$ at time $t+\tau$ and the probability density $W(x,t)$ at time t are connected by ($\tau \geqq 0$)

$$W(x, t+\tau) = \int P(x, t+\tau|x', t)\, W(x', t)\, \mathrm{d}x' \,. \tag{4.1}$$

To derive an expression for the differential $\partial W(x,t)/\partial t$, we must know the transition probability $P(x, t+\tau\,|\,x', t)$ for small τ. We first assume that we know all the moments ($n \geqq 1$)

$$M_n(x', t, \tau) = \langle [\xi(t+\tau) - \xi(t)]^n \rangle|_{\xi(t)=x'} = \int (x-x')^n P(x, t+\tau|x', t)\, \mathrm{d}x \,, \tag{4.2}$$

where $|_{\xi(t)=x'}$ means that at time t the random variable has the sharp value x'. We now derive a general expansion of the transition probability in three different ways.

First Way

If all the moments are given, we can construct the characteristic function (x' is to be considered as a parameter) (2.19, 21)

$$\begin{aligned} C(u, x', t, \tau) &= \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}u(x-x')} P(x, t+\tau|x', t)\, \mathrm{d}x \\ &= 1 + \sum_{n=1}^{\infty} (\mathrm{i}u)^n M_n(x', t, \tau)/n! \,. \end{aligned} \tag{4.3}$$

Because the characteristic function is the Fourier transform of the probability density and vice versa (2.22) we can express the transition probability by the moments M_n

$$\begin{aligned} P(x, t+\tau|x', t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}u(x-x')} C(u, x', t, \tau)\, \mathrm{d}u \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}u(x-x')} \left[1 + \sum_{n=1}^{\infty} (\mathrm{i}u)^n M_n(x', t, \tau)/n! \right] \mathrm{d}u \,. \end{aligned} \tag{4.4}$$

Because ($n \geqq 0$)

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} (\mathrm{i}u)^n \mathrm{e}^{-\mathrm{i}u(x-x')}\, \mathrm{d}u = \left(-\frac{\partial}{\partial x} \right)^n \delta(x-x') \tag{4.5}$$

and

$$\delta(x-x') f(x') = \delta(x-x') f(x) \,, \tag{4.6}$$

we have

$$P(x, t+\tau|x', t) = \left[1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial x} \right)^n M_n(x, t, \tau) \right] \delta(x-x') \,. \tag{4.7}$$

Second Way

Equation (4.7) may be derived without using the characteristic function in the following way [4.1]: starting from the identity

$$P(x, t+\tau|x', t) = \int \delta(y-x)\, P(y, t+\tau|x', t)\, \mathrm{d}y \tag{4.8}$$

and using the formal Taylor series expansion of the δ function in the form

$$\begin{aligned}\delta(y-x) &= \delta(x'-x+y-x') \\ &= \sum_{n=0}^{\infty} \frac{(y-x')^n}{n!} \left(\frac{\partial}{\partial x'}\right)^n \delta(x'-x) \\ &= \sum_{n=0}^{\infty} \frac{(y-x')^n}{n!} \left(-\frac{\partial}{\partial x}\right)^n \delta(x'-x)\,, \end{aligned} \tag{4.9}$$

we get

$$\begin{aligned} P(x,t+\tau|x',t) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial x}\right)^n \int (y-x')^n P(y,t+\tau|x',t)\,\mathrm{d}y\, \delta(x'-x) \\ &= \left[1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial x}\right)^n M_n(x',t,\tau)\right] \delta(x'-x) \\ &= \left[1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial x}\right)^n M_n(x,t,\tau)\right] \delta(x-x')\,. \end{aligned} \tag{4.10}$$

In deriving the second line of (4.10) we used (4.2) and for the last line $\delta(x-x') = \delta(x'-x)$ and (4.6).

Inserting (4.7) or (4.10) into (4.1) leads in both cases to

$$\begin{aligned} W(x,t+\tau) - W(x,t) &= \frac{\partial W(x,t)}{\partial t}\tau + O(\tau^2) \\ &= \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x}\right)^n \int \delta(x-x')\, M_n(x,t,\tau)\, W(x',t)\,\mathrm{d}x'/n! \\ &= \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x}\right)^n [M_n(x,t,\tau)/n!]\, W(x,t)\,. \end{aligned} \tag{4.11}$$

Third Way

The formal Taylor series expansion (4.9) is convenient for deriving (4.11). After multiplying (4.9) by a function of y and x' and then integrating the equation over y and x', we end with a Taylor series expansion of this function (only for this expansion can the Taylor series converge). Therefore (4.11) may be derived by avoiding any δ function and its derivatives and using only Taylor series expansion for the distribution function and the transition probability. This derivation of (4.11) runs as follows. Introducing $\Delta = x - x'$, the integrand in (4.1) may be expanded in a Taylor series according to

$$\begin{aligned} P(x,t+\tau|x',t)\, W(x',t) &= P(x-\Delta+\Delta, t+\tau|x-\Delta,t)\, W(x-\Delta,t) \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \Delta^n \left(\frac{\partial}{\partial x}\right)^n P(x+\Delta, t+\tau|x,t)\, W(x,t)\,. \end{aligned}$$

Inserting this expression in (4.1) and integrating over Δ we directly obtain (4.11). (The negative sign of the differential $\mathrm{d}\Delta = -\mathrm{d}x'$ may be absorbed into the integration boundaries.)

We now assume that the moments M_n can be expanded into a Taylor series with respect to τ ($n \geqq 1$)

$$M_n(x,t,\tau)/n! = D^{(n)}(x,t)\,\tau + O(\tau^2)\,. \tag{4.12}$$

The term with τ^0 must vanish, because for $\tau = 0$ the transition probability P has the initial value

$$P(x,t|x',t) = \delta(x-x')\,, \tag{4.13}$$

which leads to vanishing moments (4.2). By taking into account only the linear terms in τ we thus have

$$\frac{\partial W(x,t)}{\partial t} = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x}\right)^n D^{(n)}(x,t)\, W(x,t) = \boldsymbol{L}_{\mathrm{KM}} W\,, \tag{4.14}$$

where the differential symbol acts on $D^{(n)}(x,t)$ and $W(x,t)$. The Kramers-Moyal operator $\boldsymbol{L}_{\mathrm{KM}}$ is defined by

$$\boldsymbol{L}_{\mathrm{KM}}(x,t) = \sum_{n=1}^{\infty} (-\partial/\partial x)^n D^{(n)}(x,t)\,. \tag{4.15}$$

Equation (4.14) is the Kramers-Moyal expansion.

For non-Markovian processes, the conditional probability in (4.1) depends on the values of the stochastic variable $\xi(t')$ at all earlier times $t' < t$ (2.69). Hence also the moments (4.2) and their expansion coefficients $D^{(n)}$ which occur in (4.14) depend on these earlier times for non-Markovian processes. For Markov processes, $D^{(n)}$ do not depend on the values of $\xi(t')$ at these earlier times. With respect to time t, (4.14) is then a differential equation of first order and the distribution function $W(x,t)$ is uniquely determined by integration of (4.14) starting with the initial distribution $W(x,t_0)$ $(t > t_0)$ and for appropriate boundary conditions. Therefore we assume that the process described by the probability density $W(x,t)$ is a Markov process.

The transition probability $P(x,t\,|\,x',t')$ is the distribution $W(x,t)$ for the special initial condition $W(x,t') = \delta(x-x')$. Thus the transition probability must also obey (4.14), i.e.,

$$\partial P(x,t|x',t')/\partial t = \boldsymbol{L}_{\mathrm{KM}}(x,t)\, P(x,t|x',t')\,, \tag{4.16}$$

where the initial condition of P is given by (4.13) with t replaced by t'.

4.1.1 Formal Solution

A formal solution of (4.16) with the initial value (4.13) for time-independent $\boldsymbol{L}_{\mathrm{KM}}$ reads

$$P(x,t|x',t') = e^{L_{KM}(x)(t-t')}\delta(x-x') . \tag{4.17}$$

For time-dependent Kramers-Moyal operators we have to take into account that $\boldsymbol{L}_{KM}$ does not need to commute with itself for different times. The general solution of (4.16) with the initial value (4.13) may be found by iteration of (4.16) (Dyson series [4.2])

$$\begin{aligned} P(x,t|x',t') &= \delta(x-x') + \int_{t'}^{t} \boldsymbol{L}_{KM}(x,t_1)\,dt_1\,\delta(x-x') \\ &\quad + \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2 \boldsymbol{L}_{KM}(x,t_1)\boldsymbol{L}_{KM}(x,t_2)\,\delta(x-x') + \dots \\ &= \left[1 + \sum_{n=1}^{\infty} \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2 \dots \int_{t'}^{t_{n-1}} dt_n \boldsymbol{L}_{KM}(x,t_1)\dots\boldsymbol{L}_{KM}(x,t_n)\right] \\ &\quad \times \delta(x-x') . \end{aligned} \tag{4.18}$$

If we introduce the time-ordering operator $\overleftarrow{T}$ which interchanges the time-dependent operators in such a way that the operators with larger times stand to the left of operators with smaller times, (4.18) becomes [4.2]

$$\begin{aligned} P(x,t|x',t') &= \overleftarrow{T}\left[1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{t'}^{t} dt_1 \int_{t'}^{t} dt_2 \dots \int_{t'}^{t} dt_n \boldsymbol{L}_{KM}(x,t_1)\dots\boldsymbol{L}_{KM}(x,t_n)\right] \\ &\quad \times \delta(x-x') \\ &= \overleftarrow{T} \exp\left[\int_{t'}^{t} \boldsymbol{L}_{KM}(x,t'')\,dt''\right] \delta(x-x') . \end{aligned} \tag{4.19}$$

For small time differences $\tau = t - t'$ (4.18) reduces to

$$P(x,t+\tau|x',t) = [1 + \boldsymbol{L}_{KM}(x,t)\cdot\tau + O(\tau^2)]\,\delta(x-x') \tag{4.20}$$

in agreement with (4.7, 12, 15).

4.2 Kramers-Moyal Backward Expansion

In (4.15, 16) we derived an equation of motion for the transition probability $P(x,t|x',t')$. In (4.15, 16) differential operators with respect to x and t occur, i.e., with respect to the value of the stochastic variable $\xi(t)$ at the later time $t>t'$. Backward expansions are equations of motion for P where we differentiate with respect to x' and t', i.e., with respect to the value of the stochastic variable $\xi(t')$ at the earlier time $t'<t$. As shown at the end of this section, both equations lead to the same result for P and thus either one can be used for determining P.

For the derivation we follow closely the procedure of the second way in Sect. 4.1

Starting from the Chapman-Kolmogorov equation (2.78) in the form ($t \geqq t' + \tau \geqq t'$)

$$P(x,t|x',t') = \int P(x,t|x'',t'+\tau)P(x'',t'+\tau|x',t')\,dx'' \tag{4.21}$$

we write as in (4.8)

$$P(x'',t'+\tau|x',t') = \int \delta(y-x'')\,P(y,t'+\tau|x',t')\,dy . \tag{4.22}$$

Furthermore, we make a Taylor series expansion of the δ function in the form

$$\begin{aligned} \delta(y-x'') &= \delta(x'-x''+y-x') \\ &= \sum_{n=0}^{\infty} \frac{(y-x')^n}{n!} \left(\frac{\partial}{\partial x'}\right)^n \delta(x'-x'') \end{aligned} \tag{4.23}$$

and obtain

$$\begin{aligned} P(x'',t'+\tau|x',t') &= \sum_{n=0}^{\infty} \frac{1}{n!} \int (y-x')^n P(y,t'+\tau|x',t')\,dy \left(\frac{\partial}{\partial x'}\right)^n \delta(x'-x'') \\ &= \left[1 + \sum_{n=1}^{\infty} \frac{1}{n!} M_n(x',t',\tau)\left(\frac{\partial}{\partial x'}\right)^n\right] \delta(x'-x'') . \end{aligned} \tag{4.24}$$

Inserting (4.24) in (4.21) yields

$$\begin{aligned} P(x,t|x',t') - P(x,t|x',t'+\tau) &= -\frac{\partial P(x,t|x',t')}{\partial t'}\tau + O(\tau^2) \\ &= \sum_{n=1}^{\infty} \frac{1}{n!} M_n(x',t',\tau)\left(\frac{\partial}{\partial x'}\right)^n P(x,t|x',t'+\tau) \\ &= \tau \sum_{n=1}^{\infty} D^{(n)}(x',t')\left(\frac{\partial}{\partial x'}\right)^n P(x,t|x',t') + O(\tau^2) . \end{aligned} \tag{4.25}$$

In deriving the last line we used (4.12). By taking into account only the linear terms in τ we get

$$\frac{\partial P(x,t|x',t')}{\partial t'} = -\boldsymbol{L}^{+}_{KM}(x',t')\,P(x,t|x',t') \tag{4.26}$$

with

$$\boldsymbol{L}^{+}_{KM}(x',t') = \sum_{n=1}^{\infty} D^{(n)}(x',t')(\partial/\partial x')^n . \tag{4.27}$$

As may be easily checked, (4.27) is the adjoint operator of (4.15). Equations (4.26, 27) form the desired Kramers-Moyal backward expansion.

4.2.1 Formal Solution

A formal solution of (4.26) with the initial value (4.13) reads for time-independent L_{KM}^{+}

$$P(x,t|x',t') = e^{L_{KM}^{+}(x')(t-t')}\,\delta(x-x')\,. \tag{4.28}$$

For a time-dependent operator we have the Dyson series

$$\begin{aligned} P(x,t|x',t') &= \left[1+\sum_{n=1}^{\infty}\int_{t'}^{t}dt_1\int_{t_1}^{t}dt_2\ldots\int_{t_{n-1}}^{t}dt_n L_{KM}^{+}(x',t_1)\ldots L_{KM}^{+}(x',t_n)\right]\\ &\quad\times\delta(x-x')\\ &= \vec{T}\left[1+\sum_{n=1}^{\infty}\frac{1}{n!}\int_{t'}^{t}dt_1\int_{t'}^{t}dt_2\ldots\int_{t'}^{t}dt_n L_{KM}^{+}(x',t_1)\ldots L_{KM}^{+}(x',t_n)\right]\\ &\quad\times\delta(x-x')\\ &= \vec{T}\exp\left[\int_{t'}^{t}L_{KM}^{+}(x',t'')\,dt''\right]\delta(x-x')\,. \end{aligned} \tag{4.29}$$

In (4.29) the time-ordering operator $\vec{T}$ arranges the operators $L_{KM}^{+}(x',t)$ so that the time in the products of L_{KM}^{+} increases from left to right. For small time differences $\tau = t-t'$ (4.29) reduces to

$$P(x,t+\tau|x',t) = [1+L_{KM}^{+}(x',t)\,\tau+O(\tau^2)]\,\delta(x-x')\,. \tag{4.30}$$

4.2.2 Equivalence of the Solutions of the Forward and Backward Equations

To show the equivalence of (4.28, 29 and 30) with (4.17, 19 and 20), respectively, we first derive the relation

$$A(x)\,\delta(x-x') = A^{+}(x')\,\delta(x-x')\,. \tag{4.31}$$

Here $A(x)$ is a general real operator containing only differential operators with respect to x and functions depending only on x. For a derivation of (4.31) we first observe that $A(x)\,\varphi(x)$ can be written in two different ways:

$$\begin{aligned} A(x)\,\varphi(x) &= A(x)\int\delta(x-x')\,\varphi(x')\,dx'\\ &= \int A(x)\,\delta(x-x')\,\varphi(x')\,dx'\\ &= \int\varphi(x')\,A(x)\,\delta(x-x')\,dx'\,, \end{aligned} \tag{4.32}$$

$$\begin{aligned} A(x)\,\varphi(x) &= \int\delta(x-x')\,A(x')\,\varphi(x')\,dx'\\ &= \int\varphi(x')\,A^{+}(x')\,\delta(x-x')\,dx'\,. \end{aligned} \tag{4.33}$$

By subtracting both expressions we get

$$0 = \int\varphi(x')\,[A(x)\,\delta(x-x')-A^{+}(x')\,\delta(x-x')]\,dx'\,. \tag{4.34}$$

Because $\varphi(x)$ is an arbitrary function the bracket in (4.34) must be zero.

The equivalence of (4.20, 30) follows immediately from (4.31) for $A(x) = L_{KM}(x)$. Furthermore, one easily shows the equivalence of (4.28, 29) with (4.17, 19) by using (4.31) for

$$A(x) = e^{L_{KM}(x)(t-t')}\,;\qquad A^{+}(x) = e^{L_{KM}^{+}(x)(t-t')} \tag{4.35}$$

and for

$$\begin{aligned} A(x) &= \vec{T}\exp\left[\int_{t'}^{t}L_{KM}(x,t'')\,dt''\right]\\ A^{+}(x) &= \vec{T}\exp\left[\int_{t'}^{t}L_{KM}^{+}(x,t'')\,dt''\right]. \end{aligned} \tag{4.36}$$

The last relation follows from the fact that the adjoint of a product reverses its order

$$(ABC\ldots)^{+} = \ldots C^{+}B^{+}A^{+}\,. \tag{4.37}$$

4.3 Pawula Theorem

For the solution of (4.14) it is important to know how many terms of expansion (4.15) must be taken into account. We first derive the theorem of *Pawula* [4.3], which states that for a positive transition probability P, the expansion (4.15) may stop either after the first term or after the second term, if it does not stop after the second term it must contain an infinite number of terms. If expansion (4.15) stops after the second term, (4.15, 16) are then called the Fokker-Planck or forward Kolmogorov equation, and (4.26, 27) is then called the backward Kolmogorov equation.

To derive the Pawula theorem we need the generalized Schwartz inequality

$$\left[\int f(x)\,g(x)\,P(x)\,dx\right]^2 \leqq \int f^2(x)\,P(x)\,dx\int g^2(x)\,P(x)\,dx\,. \tag{4.38}$$

In (4.38) $P(x)$ is a nonnegative function and $f(x)$ and $g(x)$ are arbitrary functions. The inequality may be derived from

$$\iint [f(x)\,g(y) - f(y)\,g(x)]^2 P(x)\,P(y)\,\mathrm{d}x\,\mathrm{d}y \geqq 0\,,$$

which obviously holds for nonnegative P. We now apply (4.38) with $(n, m \geqq 0)$

$$f(x) = (x - x')^n\,; \qquad g(x) = (x - x')^{n+m}\,;$$
$$P(x) = P(x, t + \tau | x', t')$$

and thus obtain for the moments (4.2) the inequality

$$M_{2n+m}^2 \leqq M_{2n} \cdot M_{2n+2m}\,. \tag{4.39}$$

For $n = 0$ we have $M_m^2 \leqq M_{2m}$. This relation is obviously fulfilled for $m = 0$ ($M_0 = 1$). For $m \geqq 1$ no restriction follows from this relation for the short time expansion coefficients $D^{(n)}$ of M_n (4.12). For $m = 0$, $M_{2n}^2 \leqq M_{2n}^2$, which is obviously fulfilled for every n. Thus we need to consider (4.39) only for $n \geqq 1$ and $m \geqq 1$. By inserting (4.12) into (4.39), dividing the resulting inequality by τ^2 and taking the limit $\tau \to 0$ we then obtain the following inequality for the expansion coefficients $D^{(n)}$ ($n \geqq 1$, $m \geqq 1$):

$$[(2n+m)!\,D^{(2n+m)}]^2 \leqq (2n)!\,(2n+2m)!\,D^{(2n)} D^{(2n+2m)}\,. \tag{4.40}$$

If $D^{(2n)}$ is zero, $D^{(2n+m)}$ must be zero, too, i.e.,

$$D^{(2n)} = 0 \Rightarrow D^{(2n+1)} = D^{(2n+2)} = \ldots = 0 \qquad (n \geqq 1)\,. \tag{4.41}$$

Furthermore if $D^{(2n+2m)}$ is zero, $D^{(2n+m)}$ must be zero, too, i.e.,

$$D^{(2r)} = 0 \Rightarrow D^{(r+n)} = 0 \qquad (n = 1, \ldots, r-1)\,, \qquad \text{i.e.,}$$
$$D^{(2r-1)} = \ldots = D^{(r+1)} = 0 \qquad (r \geqq 2)\,. \tag{4.42}$$

From (4.41) and the repeated use of (4.42), one concludes that if any $D^{(2r)} = 0$ for $r \geqq 1$ all coefficients $D^{(n)}$ with $n \geqq 3$ must vanish, i.e.,

$$D^{(2r)} = 0 \Rightarrow D^{(3)} = D^{(4)} = \ldots = 0 \qquad (r \geqq 1)\,. \tag{4.43}$$

The Pawula theorem immediately follows from the last statement. (In contrast to (4.43) for even coefficients a vanishing odd coefficient does not lead to restrictions.)

The Pawula theorem, however, does not say that expansions truncated at $n \geqq 3$ are of no use. As we shall discuss in Sect. 4.6 for a simple example, one may very well use Kramers-Moyal expansions truncated at $n \geqq 3$ for calculating distribution functions. Though the transition probability must then have negative values at least for sufficiently small times, these negative values may be very small. For the example discussed in Sect. 4.6, the distribution function obtained by the Kramers-Moyal expansion truncated at a proper $n \geqq 3$ is in better agreement with the exact distribution than the distribution function following from the Kramers-Moyal expansion truncated at $n = 2$.

4.4 Fokker-Planck Equation for One Variable

If the Kramers-Moyal expansion (4.14) stops after the second term we get the Fokker-Planck equation ($\partial/\partial t$ is denoted by a dot)

$$\dot{W}(x,t) = L_{\mathrm{FP}} W(x,t)\,, \tag{4.44}$$

$$L_{\mathrm{FP}} = -\frac{\partial}{\partial x} D^{(1)}(x,t) + \frac{\partial^2}{\partial x^2} D^{(2)}(x,t)\,. \tag{4.45}$$

For the nonlinear Langevin equation (3.67) with (3.68) the drift coefficient $D^{(1)}$ and the diffusion coefficient $D^{(2)}$ are given by (3.93, 94) in terms of the function occurring in (3.67). All higher Kramers-Moyal coefficients $D^{(n)}$ with $n \geqq 3$ are zero [see the last equation in (3.95)] and therefore (4.44) with L_{FP} given by (4.45) is the exact equation for the probability density $W(x,t)$. For another derivation, see App. A5.

Equations (4.44, 45) may be written in the form

$$\frac{\partial W}{\partial t} + \frac{\partial S}{\partial x} = 0\,, \tag{4.46}$$

$$S(x,t) = \left[D^{(1)}(x,t) - \frac{\partial}{\partial x} D^{(2)}(x,t)\right] W(x,t)\,. \tag{4.47}$$

Because (4.46) is a continuity equation for a probability distribution, S has to be interpreted as a probability current. If this probability current vanishes at the boundaries $x = x_{\min}$ and $x = x_{\max}$, (4.46) then guarantees that the normalization is preserved

$$\int_{x_{\min}}^{x_{\max}} W(x,t)\,\mathrm{d}x = \text{const}\,. \tag{4.48}$$

For natural boundary conditions ($x_{\min} = -\infty, x_{\max} = \infty$), $W(x,t)$ and the probability current (4.47) also vanish at $x = \pm\infty$.

For a stationary process the probability current must be constant. With natural boundary conditions, the probability current must be zero. To demonstrate the usefulness of the Fokker-Planck equation we calculate the stationary distribution function for the Brownian motion process described by the Langevin equation (3.1) with (3.2). Here we have

$$D^{(1)} = -\gamma v\,, \qquad D^{(2)} = q/2 = \gamma kT/m \tag{4.49}$$

and we immediately get from

$$S = \left(-\gamma v - \frac{\gamma kT}{m}\frac{\partial}{\partial v}\right) W = 0 \tag{4.50}$$

and from the normalization condition the Maxwell distribution (3.30), i.e.,

$$W(v) = \sqrt{\frac{m}{2\pi kT}} \exp\left(-\frac{mv^2}{2kT}\right). \tag{4.51}$$

4.4.1 Transition Probability Density for Small Times

We now derive an expression for the transition probability density for small τ in another form than (4.20) specialized for the Fokker-Planck operator, i.e.,

$$P(x, t+\tau | x', t) = [1 + L_{\mathrm{FP}}(x,t)\tau + O(\tau^2)]\,\delta(x-x') \tag{4.52}$$

with

$$L_{\mathrm{FP}}(x,t) = -\frac{\partial}{\partial x} D^{(1)}(x,t) + \frac{\partial^2}{\partial x^2} D^{(2)}(x,t). \tag{4.53}$$

Inserting (4.53) into (4.52) we get up to corrections of the order τ^2

$$\begin{aligned} P(x,t+\tau|x',t) &= \left[1 - \frac{\partial}{\partial x} D^{(1)}(x',t)\tau + \frac{\partial^2}{\partial x^2} D^{(2)}(x',t)\tau\right]\delta(x-x') \\ &= \exp\left[-\frac{\partial}{\partial x} D^{(1)}(x',t)\tau + \frac{\partial^2}{\partial x^2} D^{(2)}(x',t)\tau\right]\delta(x-x'). \end{aligned} \tag{4.54}$$

In deriving (4.54) we replaced x by x' (4.6) in the drift and diffusion coefficients. If we now introduce the representation of the δ function in terms of a Fourier integral, we obtain for small τ

$$\begin{aligned} P(x,t+\tau|x',t) &= \exp\left[-\frac{\partial}{\partial x} D^{(1)}(x',t)\tau + \frac{\partial^2}{\partial x^2} D^{(2)}(x',t)\tau\right]\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}u(x-x')}\mathrm{d}u \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp[-\mathrm{i}uD^{(1)}(x',t)\tau - u^2 D^{(2)}(x',t)\tau + \mathrm{i}u(x-x')]\,\mathrm{d}u \\ &= \frac{1}{2\sqrt{\pi D^{(2)}(x',t)\tau}} \exp\left(-\frac{[x-x'-D^{(1)}(x',t)\tau]^2}{4D^{(2)}(x',t)\tau}\right). \end{aligned} \tag{4.55}$$

For drift and diffusion coefficients independent of x and t, (4.55) is not only valid for small τ, but for arbitrary $\tau > 0$. [The last line in (4.54) is then the formal solution (4.17).] We now want to check that (4.55) leads to the correct moments

$$M_n(x',t,\tau) = \int (x-x')^n P(x,t+\tau|x',t)\,\mathrm{d}x.$$

Using [4.4]

$$\int_{-\infty}^{\infty} x^n \exp[-(x-\beta)^2]\,\mathrm{d}x = (2\mathrm{i})^{-n}\sqrt{\pi}\,\mathrm{H}_n(\mathrm{i}\beta), \tag{4.56}$$

where $\mathrm{H}_n(x)$ are the Hermite polynomials ($\mathrm{H}_0 = 1$, $\mathrm{H}_1 = 2x$, $\mathrm{H}_2 = 4x^2 - 2, \ldots$) we obtain from (4.55)

$$\begin{aligned} M_n(x',t,\tau) = {} & [-\mathrm{i}\sqrt{D^{(2)}(x',t)\tau}]^n \\ & \times \mathrm{H}_n\{\tfrac{1}{2}\mathrm{i}D^{(1)}(x',t)\sqrt{\tau/D^{(2)}(x',t)}\}. \end{aligned} \tag{4.57}$$

For the expansion coefficients of M_n linear in τ we therefore have ($M_0 = 1$)

$$\lim_{\tau\to 0}\frac{1}{\tau} M_n(x',t,\tau)/n! = \begin{cases} D^{(1)}(x',t) & \quad n=1 \\ D^{(2)}(x',t) & \text{for} \quad n=2 \\ 0 & \quad n \geqq 3. \end{cases}$$

Thus (4.55) [as well as (4.52, 53)] leads to the correct drift and diffusion coefficients, i.e., it leads to expectation values which are correct up to terms linear in τ.

The form (4.55) is not unique. A class of equivalent forms has been derived [4.5, 6]. One of these forms may be obtained as follows: by performing the differentiation for the drift and diffusion coefficient in (4.53) we get

$$\begin{aligned} L_{\mathrm{FP}}(x,t) = {} & -\frac{\partial D^{(1)}(x,t)}{\partial x} + \frac{\partial^2 D^{(2)}(x,t)}{\partial x^2} \\ & - \left[D^{(1)}(x,t) - 2\frac{\partial D^{(2)}(x,t)}{\partial x}\right]\frac{\partial}{\partial x} + D^{(2)}(x,t)\frac{\partial^2}{\partial x^2}. \end{aligned} \tag{4.53a}$$

If we insert this expression into (4.52) and replace $\partial/\partial x$ by $-\partial/\partial x'$, we can perform the same steps as before, leading for small time τ to

$$\begin{aligned} P(x,t+\tau|x',t) = {} & \frac{1}{2\sqrt{\pi D^{(2)}(x,t)\tau}} \exp\left(-\frac{\partial D^{(1)}(x,t)}{\partial x}\tau + \frac{\partial^2 D^{(2)}(x,t)}{\partial x^2}\tau\right. \\ & \left. - \frac{\{x-x'-[D^{(1)}(x,t) - 2\partial D^{(2)}(x,t)/\partial x]\tau\}^2}{4D^{(2)}(x,t)\tau}\right). \end{aligned} \tag{4.55a}$$

Notice that here x instead of x' appears in the drift and diffusion coefficients.

4.4.2 Path Integral Solutions

The transition probabilities are needed for the path integral solutions [1.14, 4.5–12]. They are derived as follows: by repeatedly applying the Chapman-Kolmogorov equation (2.78) we can express the evolution of $W(x,t)$ from the initial distribution $W(x_0, t_0)$ in terms of the transition probability. Dividing the time difference $t-t_0$ in N small time intervals of length $\tau=(t-t_0)/N$, we have $(t_n = t_0 + n\tau)$

$$\begin{aligned}W(x,t) = &\int dx_{N-1}\int dx_{N-2}\dots\int dx_0\\ &P(x,t|x_{N-1},t_{N-1})\,P(x_{N-1},t_{N-1}|x_{N-2},t_{N-2})\dots\\ &P(x_1,t_1|x_0,t_0)\,W(x_0,t_0)\,. \qquad (4.58)\end{aligned}$$

For $N\to\infty$ we may use for the transition probability the expression (4.55) for small τ, which then gives correct expectation values of $W(x,t)$ in the limit $N\to\infty$. (Every integral is correct up to the order $1/N^2$ and the product of the $N+1$ integrals is then correct up to the order $1/N$ [4.6].) Inserting (4.55) into (4.58) and taking the limit $N\to\infty$ we obtain with $x_N = x$, $[\tau=(t-t_0)/N]$

$$\begin{aligned}W(x,t) = &\lim_{N\to\infty}\underbrace{\int\dots\int}_{N\text{ times}}\prod_{i=0}^{N-1}\{[4\pi D^{(2)}(x_i,t_i)\tau]^{-1/2}dx_i\}\\ &\times\exp\left(-\sum_{i=0}^{N-1}\frac{[x_{i+1}-x_i-D^{(1)}(x_i,t_i)\tau]^2}{4D^{(2)}(x_i,t_i)\tau}\right)W(x_0,t_0)\,. \qquad (4.59)\end{aligned}$$

If we use (4.55a) instead of (4.55) in (4.58), we obtain a slightly different expression.

Positivity of the Distribution Function

Because in (4.59) all the factors in front of $W(x_0,t_0)$ are positive, the distribution function must remain positive if we start with a positive distribution $W(x_0,t_0)$.

Generalized Onsager-Machlup Function

By writing

$$x_{i+1}-x_i = \dot{x}(t_i)\,\tau$$

we may put the negative term in the exponent in (4.59) for the limit $N\to\infty$ in the form

$$\sum_{i=0}^{N-1}\frac{[\dot{x}(t_i)-D^{(1)}(x_i,t_i)]^2}{4D^{(2)}(x_i,t_i)}\tau = \int_{t_0}^{t}\frac{[\dot{x}(t')-D^{(1)}(x(t'),t')]^2}{4D^{(2)}(x(t'),t')}dt'\,. \qquad (4.60)$$

The function under the integral is called a generalized Onsager-Machlup function. (*Onsager* and *Machlup* [4.7] investigated such forms for a linear drift coefficient and a constant diffusion coefficient.) Expression (4.59), where the sum in the exponent is replaced by (4.60), seems at first glance to be quite evident. For small diffusion $D^{(2)}$, for instance, only the pathes near the deterministic solution of

$$\dot{x} = D^{(1)}(x,t)\,,$$

contribute to W. It was pointed out in [1.14, 4.6], however, that this and similar other continuous forms are meaningless if the discretization process is not specified. Hence, only discrete forms such as (4.59) should be used.

4.5 Generation and Recombination Processes

To exemplify a process containing an infinite number of Kramers-Moyal coefficients $D^{(n)}$ we consider a process in which the stochastic variable $\xi(t)$ can take on only the discrete values $x_m = lm$ $(m=1,\dots,M)$ and in which only transitions to nearest-neighbor states occur. If the transition rate from state x_m to state x_{m+1} (generation rate) is denoted by $G(x_m,t)$ and if the transition rate from state x_m to state x_{m-1} (recombination rate) is denoted by $R(x_m,t)$, the equation of motion for the probability $W(x_m,t)$ of state x_m is given by the following master equation [special case of (1.34) for nearest-neighbor transitions]

$$\begin{aligned}\dot{W}(x_m,t) = &\,G(x_{m-1},t)\,W(x_{m-1},t) - G(x_m,t)\,W(x_m,t)\\ &+R(x_{m+1},t)\,W(x_{m+1},t) - R(x_m,t)\,W(x_m,t)\,. \qquad (4.61)\end{aligned}$$

This equation may be easily read off Fig. 4.1. For $x_m = m$, $G(m) = \mu m$, $R(m) = \nu m$, (4.61) describes a birth and death process, whereas for $x_m = m$, $G(m) = \mu$, $R(m) = 0$, (4.61) describes a Poisson process. Exact solutions of (4.61) for various processes are given in Table 2.1 of [1.12]; for multidimensional generation and recombination processes, see [1.11c].

Because

$$f(x\pm l) = \exp(\pm l\,\partial/\partial x)f(x)$$

we may immediately write the master equation (4.61) in form of the Kramers-Moyal expansion (4.14). Denoting the variable x_m by x we have

$$\begin{aligned}\dot{W}(x,t) &= [\exp(-l\,\partial/\partial x)-1]\,G(x,t)\,W(x,t) + [\exp(l\,\partial/\partial x)-1]\,R(x,t)\,W(x,t)\\ &= \sum_{n=1}^{\infty}(-\partial/\partial x)^n D^{(n)}(x,t)\,W(x,t)\,, \qquad (4.62)\end{aligned}$$

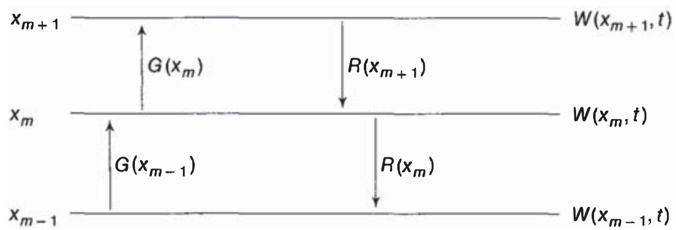

Fig. 4.1. Transition rates leading to the master equation (4.61)

where the Kramers-Moyal coefficients are given by

$$D^{(n)}(x,t) = (l^n/n!)[G(x,t) + (-1)^n R(x,t)]\,. \tag{4.63}$$

In particular, the drift and diffusion coefficients $D^{(1)}$ and $D^{(2)}$ read

$$\begin{aligned} D^{(1)} &= l(G-R) = l(\text{rate in} - \text{rate out}) \\ D^{(2)} &= (l^2/2)(G+R) = (l^2/2)(\text{rate in} + \text{rate out})\,. \end{aligned} \tag{4.63a}$$

If the difference l between the discrete steps becomes smaller, higher Kramers-Moyal coefficients also become smaller and we may truncate expansion (4.62) at some finite value n. For an actual system we cannot change l. If, for instance, x describes electric charges, l will be the elementary charge e, which cannot be changed. We may of course increase the size of system. If we increase the size of the system by a factor L, i.e. $m = 1, \dots, ML$, extensive quantities will also increase by this factor, i.e., $x = ml = Lx_{\text{nor}}$. If the rates G and R and the probability depend only on the intensive quantities $x_{\text{nor}} = x/L = (m/L)l$, then we get

$$\dot{W}(x_{\text{nor}}, t) = \sum_{n=1}^{\infty} (-\partial/\partial x_{\text{nor}})^n D^{(n)}(x_{\text{nor}}, t)\, W(x_{\text{nor}}, t)$$

$$D^{(n)}(x_{\text{nor}}, t) = (\alpha^n/n!)[G(x_{\text{nor}}, t) + (-1)^n R(x_{\text{nor}}, t)]$$

with

$$\alpha^n = (l/L)^n\,. \tag{4.64}$$

Thus by increasing the size of the system the Kramers-Moyal coefficients also decrease more rapidly in n ($1/\Omega$ expansion by *van Kampen* [1.24]). Thus, if we truncate expansion (4.62) after the second term we obtain the Fokker-Planck equation (4.44, 45) with drift and diffusion coefficients given by (4.64). Other possibilities to truncate (4.62) are discussed in the following section for the Poisson process.

4.6 Application of Truncated Kramers-Moyal Expansions

A continuous stochastic variable obeying the Langevin equation (3.67) with δ-correlated Gaussian Langevin forces (3.68) leads to a Fokker-Planck equation, i.e., to the Kramers-Moyal expansion (4.14), which stops after the second term. We have seen in the last section that for a generation and recombination process, where the stochastic variable takes on only discrete values, the Kramers-Moyal expansion has an infinite number of terms. An equation with an infinite number of terms cannot be treated numerically and the question arises whether one can approximate the infinite Kramers-Moyal expansion by a Kramers-Moyal expansion truncated at a finite order. One may conclude from the Pawula theorem

(Sect. 4.3) that the Kramers-Moyal expansion can be truncated only after the first or second terms because the transition probability calculated from the Kramers-Moyal expansion truncated at some finite term of the order $N \geqq 3$ must have negative values at least for small enough times. However, an approximate distribution function does not need to be positive everywhere. As long as the negative values and the region where they occur are small this approximate distribution function may be very useful.

We now want to investigate the different approximations of expansion (4.14) for the simple example [4.13] of the Poisson process, for which the master equation (4.61) reduces to ($l = 1$, $x_m = m \geqq 0$, $G(m) = \mu$, $R(m) = 0$)

$$\dot{W}(m,t) = \mu W(m-1,t) - \mu W(m,t)\,. \tag{4.65}$$

The solution of (4.65) with the initial value

$$W(m,0) = \delta_{m,0} \tag{4.66}$$

is the Poisson distribution

$$W(m,t) = \tau^m e^{-\tau}/m! \quad \text{with} \quad \tau = \mu t\,. \tag{4.67}$$

The cumulants K_n (2.21, 25) of this Poisson distribution are all equal ($K_n = \tau$ for $n \geqq 1$). If m is substituted by the continuous variable $x(-\infty < x < \infty)$ and $W(x-1,t)$ is expanded into a Taylor series we get the infinite Kramers-Moyal expansion

$$\dot{W}(x,t) = \sum_{n=1}^{\infty} \mu(-\partial/\partial x)^n W(x,t)/n!\,. \tag{4.68}$$

If we truncate the expansion (4.68) after the Nth term we have

$$\dot{W}_N(x,t) = \sum_{n=1}^{N} \mu(-\partial/\partial x)^n W_N(x,t)/n!\,. \tag{4.69}$$

In the continuous case we should use as initial condition

$$W(x,0) = \delta(x)\,. \tag{4.70}$$

In order to see how (4.69) approximates (4.67), we have to solve (4.69). By making a Fourier transform with respect to x it is easily seen that the solution of (4.69) with the initial condition (4.70) is given by

$$W_N(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left[ikx + \sum_{n=1}^{N} (-ik)^n \mu t/n!\right] dk\,. \tag{4.71}$$

By performing the integration we easily get for $N = 1$ and $N = 2$

$$W_1(x,t) = \delta(x-\tau)\,, \tag{4.72}$$

$$W_2(x,t) = (2\pi\tau)^{-1/2}\mathrm{e}^{-(x-\tau)^2(2\tau)^{-1}}\,. \tag{4.73}$$

For higher N the integration cannot be done analytically. For a numerical integration we write (4.71) in the real form

$$W_N(x,t) = \frac{1}{\pi}\int_0^\infty \exp\left[\sum_{m=1}^{[N/2]}(-k^2)^n\tau/(2n)!\right] \times \cos\left[kx - k\tau\sum_{n=0}^{[(N-1)/2]}(-k^2)^n/(2n+1)!\right]\mathrm{d}k\,. \tag{4.74}$$

Here $[a]$ is the integer part of the number a and the sum has to be omitted if the lower index is larger than the upper one. Due to the exponential function in (4.71), however, only the approximations for $N = 1, 2, 3, 6, 7, 10, 11, \ldots$ exist.

To compare (4.74) with the exact result (4.67) it is convenient to treat n as a continuous variable in (4.67). We therefore use [$\Gamma(x)$ is the gamma function]

$$W(x,t) = \tau^x\mathrm{e}^{-\tau}/\Gamma(x+1)\,, \tag{4.75}$$

which agrees with (4.67) for integer $x \geqq 0$. From the argument of positivity of the distribution function we conclude that (4.65) can be approximated only by truncation at $N = 2$, i.e., by a Fokker-Planck equation or the exact solution of (4.65). Figure 4.2 shows the exact solution (4.75) together with (4.73) and higher-order

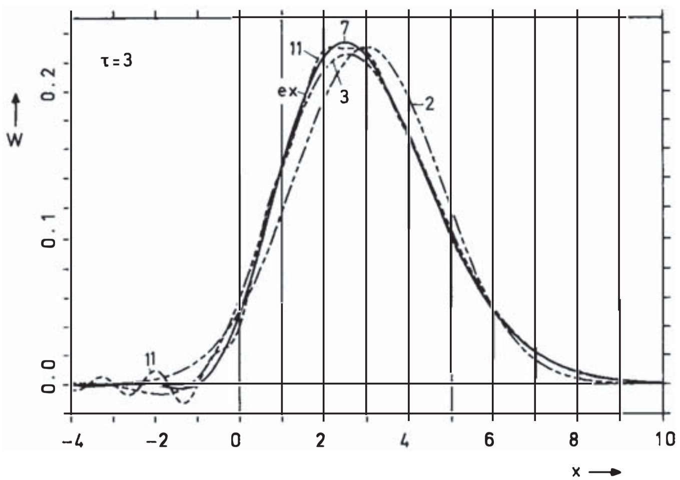

Fig. 4.2. Plot of the exact distribution (ex) and of the finite-order approximate distributions (4.71) for $N = 2, 3, 7, 11$ and $\tau = 3$. The approximation for $N = 7$ agrees with the exact distribution within the linewidth. For the Poisson process only the positive integer values of x have to be considered

Table 4.1. The exact normalization $M(0)$ and the exact first five moments $M(p)$, $p = 1, \ldots, 5$ and their successive approximations (4.76) for $N = 2, 3, 6, 7, 10, 11$ and $\tau = 3$

Approx.	$M(0)$	$M(1)$	$M(2)$	$M(3)$	$M(4)$	$M(5)$
Exact	1.000	3.000	12.000	57.000	309.000	1866.000
2	0.980	3.025	11.963	54.063	269.835	1457.907
3	1.002	2.992	12.017	56.968	306.013	1817.669
6	1.000	3.000	12.000	57.004	309.088	1867.829
7	1.000	3.000	12.000	57.000	308.993	1865.893
10	1.001	2.976	11.522	47.337	115.036	−1994.323
11	0.994	3.006	12.004	56.976	308.914	1866.233

approximations W_N calculated numerically [4.13]. It may be seen that the main virtue of W_2 is to be positive everywhere. Some higher approximations are closer to the exact solution in the sense of least-squared deviation, as seen especially for $N = 7$ where no difference is perceptible. Like the exact solution (4.75), the distribution W_7 is negative for some negative x values. For large x there are also very small negative values of W_7. As suggested from the numerical results even the approximation $N = 3$, that is significantly better than that for $N = 2$, seems to stay positive for all $x \geqq 0$ and therefore has properties similar to (4.75). As is seen, furthermore, terms of order higher than $N = 7$ tend to have larger mean-squared deviations; so the approximation (4.71) seems to be a semiconvergent series, converging only for $\tau \to \infty$ in the strict sense (for smaller τ lower approximations seem to be better, i.e., $N = 2$ for $\tau \approx 0.1$). Table 4.1 shows the moments if they are calculated either analytically (exact) or numerically by summing up the approximations (4.71) at the integer values $x = 0, 1, 2, \ldots$

$$M_N(p) = \sum_{m=0}^{\infty} m^p W_N(m,t)\,. \tag{4.76}$$

It is seen that the first higher-order approximations lead to more accurate moments. This shows that (4.71) also converges to the exact distribution (4.75); also this convergence seems to be asymptotic. It was found in [4.13] that the even-numbered approximations to the distribution oscillate more than the odd-numbered ones. The negative value of the fifth moment for $N = 10$ is a result of negative values for large x. If the moments are calculated by integration

$$\bar{M}_N(p) = \int_{-\infty}^{\infty} x^p W_N(x,t)\,\mathrm{d}x$$

it may be seen that the cumulants up to the order $p = N$ are identical to the exact ones and that higher cumulants vanish. Therefore the moments $\bar{M}_N(p)$ agree with the exact ones up to the order $p = N$.

Thus for certain parameters in the Poisson process the absolute amount of negative values of the distribution function calculated by (4.69) for appropriate $N \geqq 3$ gets extremely small in the relevant region of variables, and the solution of the Fokker-Planck equation [i.e., (4.69) for $N=2$] deviates from the exact solution much more than the solution of (4.69) deviates for some suitable $N \geqq 3$ values. From this example we conclude that for approximate calculations of distribution functions, Kramers-Moyal expansion truncated at some suitable $N \geqq 3$ term may sometimes be used. Because the convergence seems to be asymptotic, its N value should not be chosen too large. (To estimate the appropriate N value without knowing the exact result will, however, be a difficult task.)

4.7 Fokker-Planck Equation for N Variables

For N stochastic variables

$$\{\xi\} = \xi_1, \xi_2, \dots, \xi_N \tag{4.77}$$

we proceed similarly to the one-variable case. We start with the extension of (4.1) for N variables, i.e., with

$$W(\{x\}, t+\tau) = \int P(\{x\}, t+\tau|\{x'\}, t)\, W(\{x'\}, t)\, d^N x' \,. \tag{4.78}$$

In (4.78) the volume element is denoted by

$$d^N x' = dx_1' \, dx_2' \dots dx_N' \tag{4.79}$$

and N integrations have to be performed over the N variables (only one integration sign is written down). Denoting the δ function for the N variables by

$$\delta(\{x\}) = \delta(x_1)\, \delta(x_2) \dots \delta(x_N)\,, \tag{4.80}$$

we may write

$$P(\{x\}, t+\tau|\{x'\}, t) = \int \delta(\{y\} - \{x\})\, P(\{y\}, t+\tau|\{x'\}, t)\, d^N y \,. \tag{4.81}$$

It is now convenient to use the summation convention, i.e., we perform the summation over latin indices appearing twice in the expressions without writing down the summation signs. A Taylor series expansion at $\{y\} = \{x'\}$ of the δ function appearing in (4.81) then has the form

$$\begin{aligned} \delta(\{y\} - \{x\}) &= \delta(\{x'\} - \{x\} + \{y\} - \{x'\}) \\ &= \sum_{\nu=0}^{\infty} \frac{1}{\nu!} (y_{j_1} - x'_{j_1})(y_{j_2} - x'_{j_2}) \dots (y_{j_\nu} - x'_{j_\nu}) \frac{\partial^\nu}{\partial x'_{j_1} \partial x'_{j_2} \dots \partial x'_{j_\nu}} \delta(\{x'\} - \{x\}) \\ &= \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \frac{(-\partial)^\nu}{\partial x_{j_1} \partial x_{j_2} \dots \partial x_{j_\nu}} (y_{j_1} - x'_{j_1})(y_{j_2} - x'_{j_2}) \dots (y_{j_\nu} - x'_{j_\nu})\, \delta(\{x'\} - \{x\}) \,. \end{aligned} \tag{4.82}$$

In deriving the last line we used

$$(\partial/\partial x_i')\, \delta(\{x'\} - \{x\}) = (-\partial/\partial x_i)\, \delta(\{x'\} - \{x\}) \,.$$

The summation convention implies that we have to sum over the indices $j_1, j_2, \dots j_\nu$. Inserting (4.82) into (4.81) yields

$$\begin{aligned} &P(\{x\}, t+\tau|\{x'\}, t) \\ &= \left[1 + \sum_{\nu=1}^{\infty} \frac{1}{\nu!} \frac{(-\partial)^\nu}{\partial x_{j_1} \partial x_{j_2} \dots \partial x_{j_\nu}} M^{(\nu)}_{j_1, j_2, \dots, j_\nu}(\{x\}, t, \tau) \right] \delta(\{x\} - \{x'\}) \,, \end{aligned} \tag{4.83}$$

where the νth moment is defined by

$$\begin{aligned} M^{(\nu)}_{j_1, j_2, \dots j_\nu}(\{x'\}, t, \tau) = \int &(y_{j_1} - x'_{j_1})(y_{j_2} - x'_{j_2}) \dots (y_{j_\nu} - x'_{j_\nu}) \\ &\times P(\{y\}, t+\tau|\{x'\}, t)\, d^N y \,. \end{aligned} \tag{4.84}$$

In deriving (4.83) we used in accordance with the one-dimensional case (4.10) the extension of (4.6) to the N-variable δ function and $\delta(\{x\} - \{x'\}) = \delta(\{x'\} - \{x\})$. Expanding the moments for small τ (4.12)

$$M^{(\nu)}_{j_1, j_2, \dots, j_\nu}(\{x\}, t, \tau)/\nu! = D^{(\nu)}_{j_1, j_2, \dots, j_\nu}(\{x\}, t)\, \tau + O(\tau^2)\,, \tag{4.85}$$

we obtain the forward Kramers-Moyal expansion for N variables by inserting (4.83) into (4.78), dividing the resulting equation by τ and taking the limit $\tau \to 0$:

$$\partial W(\{x\}, t)/\partial t = \sum_{\nu=1}^{\infty} \frac{(-\partial)^\nu}{\partial x_{j_1} \dots \partial x_{j_\nu}} D^{(\nu)}_{j_1, \dots, j_\nu}(\{x\}, t)\, W(\{x\}, t) \,. \tag{4.86}$$

The solution of (4.86) with the initial condition

$$W(\{x'\}, t') = P(\{x\}, t'|\{x'\}, t') = \delta(\{x\} - \{x'\}) \tag{4.87}$$

is the transition probability P. Thus the forward equation for this probability density reads

$$\partial P(\{x\}, t|\{x'\}, t')/\partial t = L_{\rm KM}(\{x\}, t)\, P(\{x\}, t|\{x'\}, t') \tag{4.88}$$

with

$$L_{\rm KM}(\{x\}, t) = \sum_{\nu=1}^{\infty} \frac{(-\partial)^\nu}{\partial x_{j_1} \dots \partial x_{j_\nu}} D^{(\nu)}_{j_1, \dots, j_\nu}(\{x\}, t) \,. \tag{4.89}$$

The corresponding backward equation takes the form

$$\partial P(\{x\}, t|\{x'\}, t')/\partial t' = -L^{+}_{\rm KM}(\{x'\}, t')\, P(\{x\}, t|\{x'\}, t') \,, \tag{4.90}$$

$$L^{+}_{\mathrm{KM}}(\{x\},t)=\sum_{\nu=1}^{\infty} D^{(\nu)}_{j_1,\ldots,j_\nu}(\{x\},t)\frac{\partial^\nu}{\partial x_{j_1}\ldots\partial x_{j_\nu}}, \tag{4.91}$$

where the initial condition reads

$$P(\{x\},t\,|\{x'\},t)=\delta(\{x\}-\{x'\}). \tag{4.92}$$

The backward equation may easily be derived by extending the derivation in Sect. 4.2 to the N-variable case. Formal solutions of (4.88, 90) with initial conditions (4.87, 92) are given by (4.17 – 19, 28, 29), where one has to replace x and x' by $\{x\}$ and $\{x'\}$. The equivalence of the formal solutions of the forward and backward equations may be shown by using the N-variable version of (4.31), i.e.,

$$A(\{x\})\,\delta(\{x\}-\{x'\})=A^{+}(\{x'\})\,\delta(\{x\}-\{x'\}), \tag{4.93}$$

as was done for the one-variable case. In (4.93) $A(\{x\})$ is an operator containing functions and derivatives of the variables $x_1,\ldots,x_N$.

For a process which is described by the Langevin equation (3.110) with δ-correlated Gaussian Langevin forces (3.111) all coefficients $D^{(\nu)}$ with $\nu\geqq 3$ vanish (3.120). The transition probability then satisfies the equations (summation convention, $t\geqq t'$).

Fokker-Planck or Forward Kolmogorov Equation

$$\partial P(\{x\},t\,|\{x'\},t')/\partial t=L_{\mathrm{FP}}(\{x\},t)\,P(\{x\},t\,|\{x'\},t'), \tag{4.94}$$

$$L_{\mathrm{FP}}(\{x\},t)=-\frac{\partial}{\partial x_i}D_i(\{x\},t)+\frac{\partial^2}{\partial x_i\partial x_j}D_{ij}(\{x\},t), \tag{4.95}$$

Backward Kolmogorov Equation

$$\partial P(\{x\},t\,|\{x'\},t')/\partial t'=-L^{+}_{\mathrm{FP}}(\{x'\},t')\,P(\{x\},t\,|\{x'\},t'), \tag{4.96}$$

$$L^{+}_{\mathrm{FP}}(\{x'\},t')=D_i(\{x'\},t')\frac{\partial}{\partial x'_i}+D_{ij}(\{x'\},t')\frac{\partial^2}{\partial x'_i\,\partial x'_j}. \tag{4.97}$$

The initial condition in both cases is

$$P(\{x\},t\,|\{x'\}t)=P(\{x\},t'\,|\{x'\},t')=\delta(\{x\}-\{x'\}). \tag{4.98}$$

If we multiply (4.94) by $W(\{x'\},t')$ and integrate over x' we obtain the Fokker-Planck equation for the probability density $W(\{x\},t)$, i.e.,

$$\partial W(\{x\},t)/\partial t=L_{\mathrm{FP}}(\{x\},t)\,W(\{x\},t). \tag{4.94a}$$

In (4.95, 97) we omitted the upper index 1 in the drift coefficient and the upper index 2 in the diffusion coefficient, because both coefficients are distinguished by the number of lower indices. The drift coefficient or drift vector, the diffusion coefficient or diffusion matrix are defined by [cf. (4.84, 85)]:

drift vector

$$D_i(\{x\},t)=\lim_{\tau\to 0}\frac{1}{\tau}\langle\xi_i(t+\tau)-\xi_i(t)\rangle\Big|_{\xi_k(t)=x_k}, \tag{4.99}$$

diffusion matrix

$$\begin{aligned}D_{ij}(\{x\},t)&=D_{ji}(\{x\},t)\\&=\frac{1}{2}\lim_{\tau\to 0}\frac{1}{\tau}\langle[\xi_i(t+\tau)-\xi_i(t)]\,[\xi_j(t+\tau)-\xi_j(t)]\rangle\Big|_{\xi_k(t)=x_k},\end{aligned} \tag{4.100}$$

where $|_{\xi_k(t)=x_k}$ means that the stochastic variable ξ_k at time t has the sharp value x_k $(k=1,2,\ldots,N)$.

As seen from the definition, the diffusion matrix is a symmetric matrix. Furthermore, it is semidefinite, which follows from (a_i is an arbitrary vector, $\tau>0$)

$$\begin{aligned}2D_{ij}a_ia_j&=\lim_{\tau\to 0}\frac{1}{\tau}\langle\{[\xi_i(t+\tau)-\xi_i(t)]\,a_i[\xi_j(t+\tau)-\xi_j(t)]\,a_j\}\rangle\Big|_{\xi_k(t)=x_k}\\&=\lim_{\tau\to 0}\frac{1}{\tau}\langle\{[\xi_i(t+\tau)-\xi_i(t)]\,a_i\}^2\rangle\Big|_{\xi_k(t)=x_k}\geqq 0.\end{aligned} \tag{4.101}$$

Sometimes we assume that D_{ij} is positive definite, i.e.,

$$D_{ij}a_ia_j>0 \quad \text{for} \quad a_ia_i>0. \tag{4.102}$$

Then the inverse of the diffusion matrix will exist.

4.7.1 Probability Current

The Fokker-Planck equation (4.94a) with (4.95) may be written in the form of the continuity equation

$$\frac{\partial W}{\partial t}+\frac{\partial S_i}{\partial x_i}=0, \tag{4.103}$$

where the probability current S_i is defined by

$$S_i=D_iW-(\partial/\partial x_j)D_{ij}W. \tag{4.104}$$

If the probability current vanishes at an $N-1$ dimensional surface F of the N-dimensional space, the continuity equation (4.103) ensures that the total probability remains constant inside this surface F. If it is normalized to 1 at time $t=t'$, the normalization will always be 1, i.e.,

$$\int_{V(F)} W(\{x\},t)\,\mathrm{d}^N x = 1\,. \tag{4.105}$$

In (4.105) $V(F)$ is the volume inside the surface F. For natural boundary conditions the probability W and the current S_i vanish at infinity and therefore the normalization condition reads

$$\int W(\{x\},t)\,\mathrm{d}^N x = 1\,. \tag{4.105a}$$

(If we do not indicate any integration boundaries, an integration from $-\infty$ to $+\infty$ is understood.)

4.7.2 Joint Probability Distribution

As discussed in Sect. 2.4.1, the complete information of a Markov process is contained in the joint probability distribution $W_2(\{x\},t;\{x'\},t')$ which can be expressed by the transition probability density (4.98) and the distribution at time t',

$$W_2(\{x\},t;\{x'\},t') = P(\{x\},t\,|\,\{x'\},t')\,W(\{x'\},t')\,. \tag{4.106}$$

If the drift and diffusion coefficients do not depend on time, a stationary solution may exist. In this case, P can depend only on the time difference $t-t'$, and we may write for the joint probability distribution in the stationary state for

$t \geqq t'$

$$W_2(\{x\},t;\{x'\},t') = P(\{x\},t-t'\,|\,\{x'\},0)\,W_{\mathrm{st}}(\{x'\})\,, \tag{4.106a}$$

$t \leqq t'$

$$W_2(\{x\},t;\{x'\},t') \equiv W_2(\{x'\},t';\{x\},t) = P(\{x'\},t'-t\,|\,\{x\},0)\,W_{\mathrm{st}}(\{x\})\,. \tag{4.106b}$$

4.7.3 Transition Probability Density for Small Times

The extension of (4.52) for the N-variable case reads

$$P(\{x\},t+\tau\,|\,\{x'\},t) = [1+L_{\mathrm{FP}}(\{x\},t)\cdot\tau+O(\tau^2)]\,\delta(\{x\}-\{x'\})\,. \tag{4.107}$$

If we insert the operator (4.95) here we may write up to terms of the order τ^2

$$\begin{aligned}P(\{x\},t+\tau\,|\,\{x'\},t) &\approx \left[1-\frac{\partial}{\partial x_i}D_i(\{x'\},t)\,\tau+\frac{\partial^2}{\partial x_i\partial x_j}D_{ij}(\{x'\},t)\,\tau\right]\delta(\{x\}-\{x'\})\\ &\approx \exp\left[-\frac{\partial}{\partial x_i}D_i(\{x'\},t)\,\tau+\frac{\partial^2}{\partial x_i\partial x_j}D_{ij}(\{x'\},t)\,\tau\right]\delta(\{x\}-\{x'\})\,.\end{aligned}$$

Here we replaced $\{x\}$ by $\{x'\}$ in the drift and diffusion coefficient. By inserting the δ function expression

$$\delta(\{x\}-\{x'\}) = (2\pi)^{-N}\int \exp[\mathrm{i}\,u_j(x_j-x_j')]\,\mathrm{d}^N u \tag{4.108}$$

we obtain the extension of (4.55) to N variables

$$\begin{aligned}P(\{x\},t+\tau\,|\,\{x'\},t) &= (2\sqrt{\pi\tau})^{-N}[\mathrm{Det}\{D_{sl}(\{x'\},t)\}]^{-1/2}\\ &\times \exp\left\{-\frac{1}{4\tau}[D^{-1}(\{x'\},t)]_{jk}[x_j-x_j'-D_j(\{x'\},t)\,\tau]\,[x_k-x_k'-D_k(\{x'\},t)\,\tau]\right\}.\end{aligned} \tag{4.109}$$

In (4.109) we assumed that the diffusion matrix is positive definite so that the inverse of the diffusion matrix exists and $\mathrm{Det}\{D_{sl}\} \neq 0$. It may be shown in a way similar to the one-variable case that the drift vector (4.99) and the diffusion matrix (4.100) are recovered from (4.109), whereas all higher Kramers-Moyal expansion coefficients vanish. Path integral solutions may be derived from the transition probability density for small τ, i.e., from (4.109), in the same way as in the one-variable case, Sect. 4.4.2. With this path integral solution it can again be shown that the solution of the Fokker-Planck equation stays positive, if it was initially positive.

4.8 Examples for Fokker-Planck Equations with Several Variables

We now list a few examples of Fokker-Planck equations with more than one variable.

4.8.1 Three-Dimensional Brownian Motion without Position Variable

The equation of motion for the velocity of a particle without any external force is the Langevin equation (3.21) with a Gaussian δ-correlated Langevin force, whose strength is given by (3.22, 13). Therefore we now have 3 variables and the drift and diffusion coefficients read

$$D_i = -\gamma v_i\,, \qquad D_{ij} = \tfrac{1}{2}q\,\delta_{ij} = (\gamma kT/m)\,\delta_{ij}\,. \tag{4.110}$$

As is seen, the diffusion matrix is positive definite. The Fokker-Planck equation takes the form $[W = W(v_1, v_2, v_3, t)]$

$$\frac{\partial W}{\partial t} = \gamma\left(\frac{\partial}{\partial v_i} v_i + \frac{kT}{m}\frac{\partial^2}{\partial v_i \partial v_i}\right) W$$
$$= \gamma\left(\nabla_v v + \frac{kT}{m}\Delta_v\right) W . \qquad (4.111)$$

In the last line of (4.111) we have introduced vector notation, the ∇ operator and the Laplace operator act with respect to the velocity coordinate. Equation (4.111) describes a special Ornstein-Uhlenbeck process. The general solution of this process will be given in Sect. 6.5.

4.8.2 One-Dimensional Brownian Motion in a Potential

The equations of motion for the velocity and position coordinate for the Brownian motion of a particle in the potential $mf(x)$ are given by (3.130) and the corresponding drift and diffusion coefficients by (3.131). In this case the diffusion matrix is singular. The corresponding Fokker-Planck equation

$$\frac{\partial W(x,v,t)}{\partial t} = \left\{-\frac{\partial}{\partial x} v + \frac{\partial}{\partial v}[\gamma v + f'(x)] + \frac{\gamma kT}{m}\frac{\partial^2}{\partial v^2}\right\} W(x,v,t) \qquad (4.112)$$

is often called Kramers equation. In (4.112) $mf'(x) = -F(x)$ is the negative force. This equation is investigated further in Chap. 10.

4.8.3 Three-Dimensional Brownian Motion in an External Force

For three-dimensional Brownian motion in an external field of force $\boldsymbol{F}(\boldsymbol{x})$ there are 6 coordinates. The Fokker-Planck equation then reads $[W = W(x_1, x_2, x_3, v_1, v_2, v_3, t)]$

$$\frac{\partial W}{\partial t} = \left[-\frac{\partial}{\partial x_i} v_i + \frac{\partial}{\partial v_i}\left(\gamma v_i - \frac{F_i}{m}\right) + \frac{\gamma kT}{m}\frac{\partial^2}{\partial v_i \partial v_i}\right] W$$
$$= \left[-\nabla_x \boldsymbol{v} + \nabla_v\left(\gamma \boldsymbol{v} - \frac{\boldsymbol{F}}{m}\right) + \frac{\gamma kT}{m}\Delta_v\right] W . \qquad (4.113)$$

In the last expression we have used vector notation.

4.8.4 Brownian Motion of Two Interacting Particles in an External Potential

For Brownian motion of two particles with mass m_1, m_2 in one dimension, the Langevin equations are

$$\dot{x}_1 = v_1 ; \quad \dot{v}_1 = -\gamma_1 v_1 - f_a'(x_1) - \frac{m_1+m_2}{m_1}\frac{\partial}{\partial x_1} f_w(x_1 - x_2) + \Gamma_1 ,$$
$$\dot{x}_2 = v_2 ; \quad \dot{v}_2 = -\gamma_2 v_2 - f_a'(x_2) - \frac{m_1+m_2}{m_2}\frac{\partial}{\partial x_2} f_w(x_1 - x_2) + \Gamma_2 . \qquad (4.114)$$

In (4.114) $x_1(x_2)$ and $v_1(v_2)$ are the position and velocity of the first (second) particle; $mf_a(x_1)$ is the external force for the two particles and $(m_1+m_2)f_w(x_1-x_2)$ 2 is the interaction potential of the two particles. If we assume that the Langevin forces Γ_1 and Γ_2 acting on particles 1 and 2 are not correlated, then

$$\langle \Gamma_\nu(t)\Gamma_\mu(t')\rangle = 2\frac{\gamma_\nu kT}{m_\nu}\delta_{\nu\mu}\delta(t-t') \qquad (4.115)$$

(no summation convention) and the Fokker-Planck equation for the distribution function $W = W(x_1, v_1; x_2, v_2; t)$ takes the form

$$\partial W/\partial t = \left\{-\frac{\partial}{\partial x_1} v_1 + \frac{\partial}{\partial v_1}\left[f_a'(x_1) + \frac{m_1+m_2}{m_1}\frac{\partial f_w(x_1-x_2)}{\partial x_1} + \gamma_1 v_1\right]\right.$$
$$+ \gamma_1 \frac{kT}{m_1}\frac{\partial^2}{\partial v_1^2} - \frac{\partial}{\partial x_2} v_2 + \frac{\partial}{\partial v_2}$$
$$\left.\times\left[f_a'(x_2) + \frac{m_1+m_2}{m_2}\frac{\partial f_w(x_1-x_2)}{\partial x_2} + \gamma_2 v_2\right] + \gamma_2 \frac{kT}{m_2}\frac{\partial^2}{\partial v_2^2}\right\} W . \qquad (4.116)$$

A numerical solution of this equation for an external cos-potential and some models for the interaction potential are given in [4.14].

4.9 Transformation of Variables

If instead of the N variables $\{x\} = x_1, \dots, x_N$ we use other new N variables $\{x'\} = x_1', \dots, x_N'$ which are given by the old variables in the form

$$x_i' = x_i'(\{x\}, t) = x_i'(x_1, \dots, x_N, t) ; \quad i = 1, \dots, N , \qquad (4.117)$$

the Fokker-Planck equation (4.94a, 95) may be expressed in terms of the new variables. It is the purpose of this chapter to find the transformation of the old drift and diffusion coefficients to the new ones. Though these new drift and diffusion coefficients were already obtained in Sect. 3.4.2 by transforming the Langevin equation (3.110), we now want to derive this transformation by using only the Fokker-Planck equation (4.94a).

By going over from one variable system to another the probability in the volume element does not change, i.e.,

$$W\mathrm{d}^N x = W'\mathrm{d}^N x' \; . \tag{4.118}$$

Because the volume elements are transformed according to the Jacobian J

$$\mathrm{d}^N x/\mathrm{d}^N x' = J = |\mathrm{Det}\{\partial x_i/\partial x_j'\}| = 1/J' = 1/|\mathrm{Det}\{\partial x_i'/\partial x_j\}| \; , \tag{4.119}$$

the probability densities W and W' are connected by

$$W' = JW = W/J' \; . \tag{4.120}$$

To find the transformation of the Fokker-Planck equation we must first know the derivative of the Jacobian. Because

$$\frac{\partial x_i'}{\partial x_j}\frac{\partial x_j}{\partial x_k'} = \delta_{ik} \; , \tag{4.121}$$

the cofactor or minor A^{jk} of the element $a_{ji} = \partial x_i'/\partial x_j$ is given by

$$A^{jk} = J'\frac{\partial x_j}{\partial x_k'} \; . \tag{4.122}$$

Therefore we may express the derivative of J' with respect to the element $a_{jk} = \partial x_k'/\partial x_j$ by

$$\frac{\partial J'}{\partial a_{jk}} = A^{jk} = J'\frac{\partial x_j}{\partial x_k'} \; . \tag{4.123}$$

Using the chain rule we thus obtain

$$\begin{aligned} -\frac{1}{J}\frac{\partial J}{\partial x_i} &= -\frac{\partial \ln J}{\partial x_i} = \frac{\partial \ln J'}{\partial x_i} = \frac{1}{J'}\frac{\partial J'}{\partial x_i} = \frac{1}{J'}\frac{\partial J'}{\partial a_{jk}}\frac{\partial a_{jk}}{\partial x_i} \\ &= \frac{\partial x_j}{\partial x_k'}\frac{\partial}{\partial x_i}\frac{\partial x_k'}{\partial x_j} = \frac{\partial x_j}{\partial x_k'}\frac{\partial}{\partial x_j}\frac{\partial x_k'}{\partial x_i} \\ &= \frac{\partial}{\partial x_k'}\frac{\partial x_k'}{\partial x_i} \; . \end{aligned} \tag{4.124}$$

Similarly, we obtain for the time derivative of J

$$\begin{aligned} -\frac{1}{J}\left(\frac{\partial J}{\partial t}\right)_x &= \frac{1}{J'}\left(\frac{\partial J'}{\partial t}\right)_x = \frac{1}{J'}\frac{\partial J'}{\partial a_{jk}}\left(\frac{\partial a_{jk}}{\partial t}\right)_x \\ &= \frac{\partial x_j}{\partial x_k'}\left(\frac{\partial}{\partial t}\right)_x\frac{\partial x_k'}{\partial x_j} = \frac{\partial x_j}{\partial x_k'}\frac{\partial}{\partial x_j}\left(\frac{\partial x_k'}{\partial t}\right)_x \\ &= \frac{\partial}{\partial x_k'}\left(\frac{\partial x_k'}{\partial t}\right)_x \; . \end{aligned} \tag{4.125}$$

The index x indicates that the old variables are kept constant. This index is necessary if the transformation (4.117) depends on t. We obviously have

$$\left(\frac{\partial}{\partial t}\right)_x = \left(\frac{\partial}{\partial t}\right)_{x'} + \left(\frac{\partial x_k'}{\partial t}\right)_x \frac{\partial}{\partial x_k'} \; . \tag{4.126}$$

To express the derivative $\partial/\partial x_i$ in terms of the derivatives of the new variables, we, again, use the chain rule to get

$$\frac{\partial}{\partial x_i} = \frac{\partial x_k'}{\partial x_i}\frac{\partial}{\partial x_k'} = \frac{\partial}{\partial x_k'}\frac{\partial x_k'}{\partial x_i} - \left[\frac{\partial}{\partial x_k'}\frac{\partial x_k'}{\partial x_i}\right] = \frac{\partial}{\partial x_k'}\frac{\partial x_k'}{\partial x_i} + \frac{1}{J}\frac{\partial J}{\partial x_i} \; ,$$

where the bracket indicates that the operator does not act outside this bracket. Because

$$\frac{\partial}{\partial x_i} = \frac{1}{J}\frac{\partial}{\partial x_i}J - \frac{1}{J}\frac{\partial J}{\partial x_i} \; ,$$

we get the useful result

$$\frac{\partial}{\partial x_i} = \frac{1}{J}\frac{\partial}{\partial x_k'}\frac{\partial x_k'}{\partial x_i}J \; . \tag{4.127}$$

Applying (4.127) twice we obtain

$$\begin{aligned} \frac{\partial^2}{\partial x_i \partial x_j} &= \frac{1}{J}\frac{\partial}{\partial x_k'}\frac{\partial x_k'}{\partial x_i}J\frac{1}{J}\frac{\partial}{\partial x_r'}\frac{\partial x_r'}{\partial x_j}J \\ &= \frac{1}{J}\frac{\partial^2}{\partial x_k'\partial x_r'}\frac{\partial x_k'}{\partial x_i}\frac{\partial x_r'}{\partial x_j}J - \frac{1}{J}\frac{\partial}{\partial x_k'}\left[\frac{\partial}{\partial x_r'}\frac{\partial x_k'}{\partial x_i}\right]\frac{\partial x_r'}{\partial x_j}J \\ &= \frac{1}{J}\frac{\partial^2}{\partial x_k'\partial x_r'}\frac{\partial x_k'}{\partial x_i}\frac{\partial x_r'}{\partial x_j}J - \frac{1}{J}\frac{\partial}{\partial x_k'}\frac{\partial^2 x_k'}{\partial x_i \partial x_j}J \; . \end{aligned} \tag{4.128}$$

For the derivative with respect to t we have similarly

$$\left(\frac{\partial}{\partial t}\right)_x = \frac{1}{J}\left(\frac{\partial}{\partial t}\right)_x J - \frac{1}{J}\left(\frac{\partial J}{\partial t}\right)_x$$

$$= \frac{1}{J}\left(\frac{\partial}{\partial t}\right)_{x'} J + \frac{1}{J}\left(\frac{\partial x'_k}{\partial t}\right)_x \frac{\partial}{\partial x'_k} J - \frac{1}{J}\left(\frac{\partial J}{\partial t}\right)_x ,$$

$$\left(\frac{\partial x'_k}{\partial t}\right)_x \frac{\partial}{\partial x'_k} = \frac{\partial}{\partial x'_k}\left(\frac{\partial x'_k}{\partial t}\right)_x - \left[\frac{\partial}{\partial x'_k}\left(\frac{\partial x'_k}{\partial t}\right)_x\right]$$

$$= \frac{\partial}{\partial x'_k}\left(\frac{\partial x'_k}{\partial t}\right)_x + \frac{1}{J}\left(\frac{\partial J}{\partial t}\right)_x ,$$

i.e., we get

$$\left(\frac{\partial}{\partial t}\right)_x = \frac{1}{J}\left(\frac{\partial}{\partial t}\right)_{x'} J + \frac{1}{J}\frac{\partial}{\partial x'_k}\left(\frac{\partial x'_k}{\partial t}\right)_x J . \tag{4.129}$$

By inserting (4.127 – 129) into the Fokker-Planck equation (4.94a, 95) we easily obtain the Fokker-Planck equation for the new variables [$W' = JW$, (4.120)]

$$\left(\frac{\partial W'}{\partial t}\right)_{x'} = \left(-\frac{\partial}{\partial x'_k} D'_k + \frac{\partial^2}{\partial x'_k \partial x'_r} D'_{kr}\right) W' , \tag{4.130}$$

$$D'_k = \left(\frac{\partial x'_k}{\partial t}\right)_x + \frac{\partial x'_k}{\partial x_i} D_i + \frac{\partial^2 x'_k}{\partial x_i \partial x_j} D_{ij} , \tag{4.131}$$

$$D'_{kr} = \frac{\partial x'_k}{\partial x_i}\frac{\partial x'_r}{\partial x_j} D_{ij} . \tag{4.132}$$

These transformed drift and diffusion coefficients agree with (3.128, 129).

4.10 Covariant Form of the Fokker-Planck Equation

The transformation to new variables may be seen best by writing the Fokker-Planck equation in covariant form, i.e., in a form where only scalars, contravariant or covariant vectors and tensors and covariant derivatives occur. In this chapter we restrict ourselves to coordinate transformations, which are independent of time. If we go over to new coordinates x'^i which are functions of the old coordinates x^i and vice versa, i.e.,

$$x'^i = x'^i(x^1, \dots, x^N) , \quad i = 1, \dots, N$$
$$x^i = x^i(x'^1, \dots, x'^N) , \tag{4.133}$$

contravariant vectors A^i and covariant vectors A_i are transformed according to

$$A'^i = \frac{\partial x'^i}{\partial x^j} A^j , \quad A'_i = \frac{\partial x^j}{\partial x'^i} A_j . \tag{4.134}$$

The coordinate differential $\mathrm{d}x^i$ is a contravariant vector, i.e.,

$$\mathrm{d}x'^i = \frac{\partial x'^i}{\partial x^j} \mathrm{d}x^j \tag{4.135}$$

(therefore one usually puts the index of the coordinate in the upper place), whereas the gradient of a scalar is a covariant vector, i.e.,

$$\frac{\partial \varphi}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i}\frac{\partial \varphi}{\partial x^j} . \tag{4.136}$$

A scalar is not changed by a coordinate transformation, i.e.,

$$\varphi' = \varphi . \tag{4.137}$$

For the transformation of a tensor of rank n with p contravariant and $q = n - p$ covariant indices we have to use $\partial x'/\partial x$ p times and $\partial x/\partial x'$ q times [4.15 – 17]. As seen from (4.132), the diffusion tensor $\bar{D}^{ij} \equiv D_{ij}$ $(n = p = 2)$ is a purely contravariant tensor

$$\bar{D}'^{kr} = \frac{\partial x'^k}{\partial x_i}\frac{\partial x'^r}{\partial x_j} \bar{D}^{ij} \tag{4.138}$$

and the indices should therefore be put in the upper place. Obviously, the probability density is not a scalar because it transforms according to (4.120) and furthermore the drift vector is not a contravariant vector because of the last term in (4.131) ($\partial x'_k/\partial t$ is zero because in this section we assume that transformation (4.133) does not depend on time). Thus we first have to find a scalar $\bar{W}$ which may be used instead of the probability density W and a contravariant vector $\bar{D}^i$ which may be used instead of the drift vector. Following *Graham* [4.18] (see also [4.19]), we introduce a scalar $\bar{W}$ defined by

$$\bar{W} = \sqrt{\mathrm{Det}}\, W \quad \text{with} \quad \mathrm{Det} = \mathrm{Det}\{\bar{D}^{ij}\} . \tag{4.139}$$

This transformation can be done only if Det > 0, i.e., if the diffusion matrix is positive definite. Because of the transformation (4.138) it is easily seen that Det transforms according to

$$\begin{aligned}\mathrm{Det}' &= \mathrm{Det}\{\bar{D}'^{ij}\} = (\mathrm{Det}\,\partial x'^k/\partial x_i)^2\,\mathrm{Det}\{\bar{D}^{sj}\}\\ &= J^{-2}\,\mathrm{Det}\{\bar{D}^{sj}\} = J^{-2}\,\mathrm{Det}\,,\end{aligned} \tag{4.140}$$

where J is the Jacobian (4.119). We therefore have ($W' = JW$)

$$\bar{W}' = \sqrt{\mathrm{Det}'}\,W' = J^{-1}\sqrt{\mathrm{Det}}\,J W = \bar{W}\,, \tag{4.141}$$

which shows that $\bar{W}$ is indeed a scalar. Next we introduce the contravariant drift vector [4.18]

$$\bar{D}^i = D_i - \sqrt{\mathrm{Det}}\,\frac{\partial}{\partial x^j}\,\frac{\bar{D}^{ij}}{\sqrt{\mathrm{Det}}}\,, \tag{4.142}$$

which transforms in accordance with (4.134), as shown by the following equation:

$$\begin{aligned}\frac{\partial x'^k}{\partial x^i}\bar{D}^i &= \frac{\partial x'^k}{\partial x^i}D_i - \frac{\partial x'^k}{\partial x^i}\sqrt{\mathrm{Det}}\,\frac{\partial}{\partial x^j}\,\frac{\bar{D}^{ij}}{\sqrt{\mathrm{Det}}}\\ &= D'_k - \frac{\partial^2 x'^k}{\partial x^i \partial x^j}\bar{D}^{ij} - \frac{\partial x'^k}{\partial x^i}\,\frac{\sqrt{\mathrm{Det}}}{J}\,\frac{\partial}{\partial x'^r}\,\frac{\partial x'^r}{\partial x^j}J\,\frac{\bar{D}^{ij}}{\sqrt{\mathrm{Det}}}\\ &= D'_k - \frac{\partial^2 x'^k}{\partial x^i \partial x^j}\bar{D}^{ij} - \sqrt{\mathrm{Det}'}\,\frac{\partial}{\partial x'^r}\,\frac{\partial x'^k}{\partial x^i}\,\frac{\partial x'^r}{\partial x^j}\,\frac{\bar{D}^{ij}}{\sqrt{\mathrm{Det}'}}\\ &\quad + \sqrt{\mathrm{Det}'}\left(\frac{\partial}{\partial x'^r}\,\frac{\partial x'^k}{\partial x^i}\right)\frac{\partial x'^r}{\partial x^j}\,\frac{\bar{D}^{ij}}{\sqrt{\mathrm{Det}'}}\\ &= D'_k - \sqrt{\mathrm{Det}'}\,\frac{\partial}{\partial x'^r}\,\frac{\bar{D}'^{kr}}{\sqrt{\mathrm{Det}'}} = \bar{D}'^k\,.\end{aligned} \tag{4.143}$$

In deriving (4.143), in the second line we used (4.127, 131) and in the third line, (4.140), the chain rule, $(\partial x'^r/\partial x^j)\,\partial/\partial x'^r = \partial/\partial x^j$ and (4.138).

Instead of (4.104) we now use the contravariant probability current

$$\bar{S}^i = \bar{D}^i\bar{W} - \bar{D}^{ij}\frac{\partial \bar{W}}{\partial x^j}\,. \tag{4.144}$$

The contraction of a contravariant tensor of rank 2 with a covariant vector $A_i = \partial\bar{W}/\partial x^j$ [which appears in (4.144)] is a contravariant vector as may be seen by the transformation law

$$\begin{aligned}B'^i = \bar{D}'^{ij}A'_j &= \frac{\partial x'^i}{\partial x^k}\,\frac{\partial x'^j}{\partial x^r}\,\frac{\partial x^s}{\partial x'^j}\,\bar{D}^{kr}A_s = \frac{\partial x'^i}{\partial x^k}\,\delta_{rs}\bar{D}^{kr}A_s\\ &= \frac{\partial x'^i}{\partial x^k}B^k\,.\end{aligned}$$

Next we must find an expression for the divergence of a vector. We have already seen that the derivative of a scalar, i.e.,

$$\bar{W}_{;i} \equiv \frac{\partial\bar{W}}{\partial x^i}\,, \tag{4.145}$$

is a covariant vector. It is easy to show that

$$\bar{S}^i_{;i} \equiv \sqrt{\mathrm{Det}}\,\frac{\partial}{\partial x^i}\,\frac{\bar{S}^i}{\sqrt{\mathrm{Det}}} \tag{4.146}$$

is a scalar, which is the desired expression for the divergence of a vector. From (4.127, 140) we have

$$\begin{aligned}\bar{S}^i_{;i} &= \frac{\sqrt{\mathrm{Det}}}{J}\,\frac{\partial}{\partial x'^k}\,\frac{\partial x'^k}{\partial x^i}\,\frac{J}{\sqrt{\mathrm{Det}}}\bar{S}^i\\ &= \sqrt{\mathrm{Det}'}\,\frac{\partial}{\partial x'^k}\,\frac{\bar{S}'^k}{\sqrt{\mathrm{Det}'}} = \bar{S}'^k_{;k}\,.\end{aligned} \tag{4.147}$$

Thus the equation

$$\begin{aligned}\partial\bar{W}/\partial t &= -\bar{S}^i_{;i} = [-\bar{D}^i\bar{W} + \bar{D}^{ij}\bar{W}_{;j}]_{;i}\\ &= \sqrt{\mathrm{Det}}\,\frac{\partial}{\partial x^i}\,\frac{1}{\sqrt{\mathrm{Det}}}\left(-\bar{D}^i\bar{W} + \bar{D}^{ij}\frac{\partial\bar{W}}{\partial x^j}\right)\end{aligned} \tag{4.148}$$

has the correct covariant form, i.e., it has same the form for every coordinate transformation. Using (4.139, 142), it is easy to show that (4.148) is identical to the Fokker-Planck equation (4.94a), i.e.,

$$\begin{aligned}\sqrt{\mathrm{Det}}\,\frac{\partial W}{\partial t} &= \sqrt{\mathrm{Det}}\,\frac{\partial}{\partial x^i}\,\frac{1}{\sqrt{\mathrm{Det}}}\left\{-D_i\sqrt{\mathrm{Det}}\,W + \sqrt{\mathrm{Det}}\left[\frac{\partial}{\partial x^j}\,\frac{\bar{D}^{ij}}{\sqrt{\mathrm{Det}}}\right]\right.\\ &\quad\left.\times\sqrt{\mathrm{Det}}\,W + D^{ij}\frac{\partial\sqrt{\mathrm{Det}}\,W}{\partial x^j}\right\},\\ \frac{\partial W}{\partial t} &= \frac{\partial}{\partial x^i}\left\{-D_iW + \left[\frac{\partial}{\partial x^j}\,\frac{\bar{D}^{ij}}{\sqrt{\mathrm{Det}}}\right]\sqrt{\mathrm{Det}}\,W + \frac{1}{\sqrt{\mathrm{Det}}}\bar{D}^{ij}\frac{\partial\sqrt{\mathrm{Det}}\,W}{\partial x^j}\right\}\\ &= \frac{\partial}{\partial x^i}\left(-D_iW + \frac{\partial}{\partial x^j}D_{ij}W\right).\end{aligned}$$

The comparison with tensor analysis [4.15 – 17] shows that the diffusion matrix (4.100) may serve as the contravariant metric tensor [4.18], i.e.,

$$g^{ij} = \bar{D}^{ij} = D_{ij} \,. \tag{4.149}$$

The covariant metric tensor $g_{ij} = \bar{D}_{ij}$ is then the inverse of the diffusion matrix (4.100), i.e., $g_{ij} = (D^{-1})_{ij}$. The Christoffel symbols of first and second kinds and the Riemann curvature tensor are expressed by the metric tensor in the following way [4.15 – 17]:

$$[ij,k] = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}\right), \tag{4.150}$$

$$\left\{ {i \atop jk} \right\} = g^{il}[jk,l] \,, \tag{4.151}$$

$$R^l_{ijk} = \frac{\partial}{\partial x^j}\left\{ {l \atop ik} \right\} - \frac{\partial}{\partial x^k}\left\{ {l \atop ij} \right\} + \left\{ {l \atop mj} \right\}\left\{ {m \atop ik} \right\} - \left\{ {l \atop mk} \right\}\left\{ {m \atop ij} \right\}. \tag{4.152}$$

If the Riemann curvature tensor vanishes the space is Euclidean. By using a proper coordinate transformation, the metric tensor and therefore also the diffusion tensor can then be reduced to the metric tensor of Euclidean cartesian coordinates [4.17], i.e., to

$$g^{ij} = g_{ij} = D^{ij} = \delta_{ij} \,. \tag{4.153}$$

If the Riemann curvature tensor does not vanish, it is impossible to find a transformation where (4.153) is valid globally (it may then be valid only locally, i.e., near some fixed point $\{x_0\}$). If we have only one variable, the Riemann curvature tensor always vanishes. Then we can find a transformation so that the diffusion coefficient $D^{(2)} > 0$ is normalized to unity, see also Sect. 5.1. For two variables with $D_{12} = D_{21} = 0$ for instance the Riemann curvature tensor vanishes only if

$$\frac{\partial}{\partial x_1}\sqrt{D_{11}D_{22}}\frac{\partial}{\partial x_1}\frac{1}{D_{22}} + \frac{\partial}{\partial x_2}\sqrt{D_{11}D_{22}}\frac{\partial}{\partial x_2}\frac{1}{D_{11}} = 0 \tag{4.154}$$

is fulfilled.