Physica **96A** (1979) 435–453 © *North-Holland Publishing Co.*

COMPOSITE STOCHASTIC PROCESSES

N.G. van KAMPEN

Instituut voor Theoretische Fysica, Rijksuniversiteit Utrecht, Utrecht, The Netherlands

Received 27 December 1978

Certain problems in physics and chemistry lead to the definition of a class of stochastic processes. Although they are not Markovian they can be treated explicitly to some extent. In particular, the probability distribution for large times can be found. It is shown to obey a master equation. This demonstrates how a non-Markovian process can be approximated on a coarse time scale by a Markov description. The conditions for this approximation to be valid are discussed.

1. The problem

With the name "composite stochastic processes" we denote a certain class of random processes, which occurred in the following three physical models.

Giddings and Eyring[1]) described chromatography by assuming that molecules, independently of one another, can be either adsorbed on the wall or dissolved in the fluid. When dissolved they are carried along with the constant velocity of the fluid. The transitions from the adsorbed state into the dissolved state and vice versa occur at random time points, with a certain probability per unit time.

Friedman and Ben-Naim[2]) calculated the transport of "two-state Brownons", i.e., particles that can be in two states and in each of them are subject to diffusion, but with two different diffusion constants. A similar model had been used by Mysels[3]) for electrodiffusion.

Finally, Singwi and Sjölander[4]) described self-diffusion in water by assuming that each molecule can either be trapped in a cage of surrounding molecules or be free. When trapped it performs an oscillatory motion, when free it moves about by diffusion.

In these models a molecule can be in two states, in each of which it propagates in space, but the way in which it propagates is different in both states. This picture leads to the following general definition of composite stochastic processes.

Let $\mathcal{Y}$ be a probability space and $Y(t)$ a stochastic process with values $y \in \mathcal{Y}$. Let $\mathcal{X}$ be another probability space, and $X(t; y)$ a family of stochastic processes in it, whose members are distinguished by the parameter y. Then

the joint process $Z(t) \equiv \{X(t; Y(t)), Y(t)\}$ in the product space $\mathscr{Z} = \mathscr{X} \otimes \mathscr{Y}$ will be called a composite stochastic process. The essential point is that $Y(t)$ is an autonomous stochastic process, i.e., independent of X, while X does depend on Y.

In each of the above-mentioned models it was supposed that $Y(t)$ is a Markov process, and also that $X(t; y)$ for fixed y is Markovian (except for the oscillatory motion in the third model). As a consequence their probability densities obey master equations

$$\dot{p}(x, t) = \mathbb{A}(y)p(x, t), \tag{1}$$

$$\dot{q}(y, t) = Bq(y, t). \tag{2}$$

Here B is an operator acting on the y-dependence, and $\mathbb{A}(y)$ operates on the x-dependence while depending itself parametrically on y. It follows that $Z(t)$ is also Markovian and its probability density $P(x, y, t)$ obeys

$$\dot{P}(x, y, t) = \{\mathbb{A}(y) + B\}P(x, y, t). \tag{3}$$

Having defined this Markov process $Z(t)$ one may study its component $X(t; Y(t))$ separately as a stochastic process. It is *not* a Markov process, but has a memory with a relaxation time equal to the time in which $Y(t)$ approaches its equilibrium according to (2). The equation (1) for $p(x, t)$ is not a master equation owing to the occurrence of the stochastic parameter y. This way of constructing a non-Markovian process by imbedding it in a Markov process with more components has been used[5]) for describing line narrowing.

The authors mentioned above, however, did not use the master equation (3), but another approach. Weiss[6]) noticed that this alternative approach can also be used when $Y(t)$ is not Markovian. That might happen e.g. in an improved version of the chromatography model, in which one takes into account that a recently desorbed molecule has a larger probability per unit time for adsorption as it is more likely to be near to the wall. We want to investigate this case, i.e. *we take $Y(t)$ non-Markovian but still suppose that $X(t; y)$ for each fixed y is a Markov process governed by the master equation* (1). The joint process $Z(t)$, will no longer be Markovian, nor will its component $X(t; Y(t))$. Our aim is to show that nevertheless under a certain condition *it is possible to describe approximately $X(t; Y(t))$ as a Markov process* obeying a master equation. Roughly speaking, the condition is that the memory time τ_c of $Y(t)$ is short compared to the time on which p varies due to the propagators $\mathbb{A}$.

The reason for this investigation is that the Markov assumption is often used for stochastic processes in physics, although none of them are strictly Markovian. Even Brownian movement is not Markovian in time intervals

comparable with the duration of a single collision with a molecule of the surrounding fluid. Intuitively it is clear that the Markov approximation is justified when the memory is short compared to the time scale on which the over-all state of the system changes appreciably. Composite stochastic processes are a conveniently simple model on which to investigate how this approximate Markov character comes about. There is also a close connection with stochastic differential equations, see Appendix A.

2. Description of the processes to be considered

Throughout we shall take for $\mathcal{Y}$ a discrete set with elements labelled j to be called "levels" (although they need not have anything to do with quantum-mechanical energy levels). When Y has the value j it has a probability per unit time to jump to another level i equal to $\gamma_{ij}(\tau)$, where τ is the time it has sojourned in j. (For a Markov process γ_{ij} would be independent of τ.) The probability $u_j(\tau)$ that Y is still in the level j after a time τ since it arrived there is

$$u_j(\tau) = \exp\left[-\int_0^\tau \sum_i \gamma_{ij}(\tau')\,\mathrm{d}\tau'\right]. \tag{4a}$$

Normally one has $\gamma_{jj}(\tau) = 0$, but $\gamma_{jj}(\tau) > 0$ need not be excluded: it corresponds to a "transition" in which Y does not change its value, but is rejuvenated in the same level j. It is shown in Appendix B, however, that the latter case can be reduced to the former and we therefore take $\gamma_{jj}(\tau) = 0$.

The probability that Y ends its sojourn in j, after a time between τ and $\tau + \mathrm{d}\tau$ since it arrived in j, by means of a jump to a given level i, is $v_{ij}(\tau)\,\mathrm{d}\tau$, where

$$v_{ij}(\tau) = \gamma_{ij}(\tau) u_j(\tau). \tag{4b}$$

For future use we note that for each j one has the identities

$$\sum_i v_{ij}(\tau) = -\frac{\mathrm{d}u_j(\tau)}{\mathrm{d}\tau}, \tag{5}$$

$$\int_0^\infty \mathrm{d}\tau \sum_i v_{ij}(\tau) = 1. \tag{6}$$

The operators $\mathbb{A}(y)$ acting on functions in the space $\mathcal{X}$ may now be written $\mathbb{A}_j$ and will be called "propagators". As they describe the time variation of a probability density they have the property

$$\int \mathbb{A}_j \phi(x)\, dx = 0 \quad \text{for any } \phi \text{ and } j. \tag{7}$$

The joint probability density $P(x, y, t)$ may be written $P_j(x, t)$. While Y sojourns in level j the solution of (1) is

$$P_j(x, t) = e^{(t-t')\mathbb{A}_j} P_j(x, t') \quad (t \geqslant t'). \tag{8}$$

Finally we define two matrices

$$\mathsf{U}_{ij}(\tau) = \delta_{ij} u_j(\tau)\, e^{\tau \mathbb{A}_j}, \tag{9a}$$

$$\mathsf{V}_{ij}(\tau) = v_{ij}(\tau)\, e^{\tau \mathbb{A}_j} = \gamma_{ij}(\tau) u_j(\tau)\, e^{\tau \mathbb{A}_j}. \tag{9b}$$

The sans-serif indicates that they operate on the x-dependence.

Suppose at $t = 0$ the system is at x_0 in level j_0. *This does not fully describe its initial state*, because the process is not Markovian. The future also depends on how long it has already sojourned in j_0. For simplicity we here assume that it has just arrived in j_0; for the general case see Appendix C. Thus we want to compute the probability density $P_j(x, t)$ that a system is found at x in level j after a time t since it has arrived in j_0 at x_0.

It is possible for Y to get from j_0 to j by an arbitrary number $s = 1, 2, 3, \ldots$ of transitions, at arbitrary times

$$t > t_s > t_{s-1} > \cdots > t_2 > t_1 > 0, \tag{10a}$$

through an arbitrary sequence of intermediate levels

$$j_{s-1}, j_{s-2}, \ldots, j_2, j_1. \tag{10b}$$

The probability for this history to occur is

$$u_j(t - t_s) v_{jj_{s-1}}(t_s - t_{s-1}) v_{j_{s-1}j_{s-2}}(t_{s-1} - t_{s-2}) \ldots v_{j_2 j_1}(t_2 - t_1) v_{j_1 j_0}(t_1). \tag{11}$$

If $j = j_0$ one has the additional possibility that $s = 0$, with probability $u_{j_0}(t)$.

The probability density of X, j conditional on the particular history (10) is

$$\begin{aligned} P_j(x, t \mid \text{history}) = {} & \exp[(t - t_s)\mathbb{A}_j] \exp[(t_s - t_{s-1})\mathbb{A}_{j_{s-1}}] \ldots \\ & \times \exp[(t_2 - t_1)\mathbb{A}_{j_1}] \exp[t_1 \mathbb{A}_{j_0}] \delta(x - x_0). \end{aligned} \tag{12}$$

The total probability is obtained by multiplying this by the conditional probability (11) of the history and summing over all possible histories. With the aid of (9) it can be written

$$\begin{aligned} P_j(x, t) = \Big[\mathsf{U}_{jj_0}(t) + \sum_{s=1}^{\infty} \sum_{j_1, \ldots, j_s} \int dt_1 \ldots dt_s \mathsf{U}_{jj_s}(t - t_s) \mathsf{V}_{j_s j_{s-1}}(t_s - t_{s-1}) \ldots \\ \times \mathsf{V}_{j_2 j_1}(t_2 - t_1) \mathsf{V}_{j_1 j_0}(t_1) \Big] \delta(x - x_0), \end{aligned} \tag{13}$$

where the integration extends over the range (10a). The dummy summation index j_s serves to simplify the notation. The expression [ ] is an operator acting on the x-dependence of the delta function. It is not required that the $\mathbb{A}_j$ mutually commute, although they do in the applications mentioned.

The convolution integral in (13) can be reduced to a product by Laplace transformation. We suppose that the operators $\exp[\mathbb{A}_j t]$ are bounded in the sense specified in Appendix F, equation (93). We also suppose that there are positive constants C and τ_c such that

$$u_j(\tau) \leqslant C\, e^{-\tau/\tau_c}, \quad v_{ij}(\tau) \leqslant C\, e^{-\tau/\tau_c}. \tag{14}$$

Then for $\operatorname{Re} \lambda > -1/\tau_c$ the Laplace transforms of $\mathsf{U}(\tau)$ and $\mathsf{V}_{ij}(\tau)$ exist, e.g.,

$$\hat{\mathsf{V}}_{ij}(\lambda) = \int_0^\infty \mathsf{V}_{ij}(\tau)\, e^{-\lambda\tau}\, d\tau = \int_0^\infty v_{ij}(\tau)\, e^{\tau\mathbb{A}_j - \lambda\tau}\, d\tau. \tag{15}$$

The Laplace transform of (13) may now be written as a matrix product

$$\begin{aligned} \hat{P}_j(x, \lambda) &= \left[\hat{\mathsf{U}}(\lambda) + \sum_{s=1}^{\infty} \hat{\mathsf{U}}(\lambda)\{\hat{\mathsf{V}}(\lambda)\}^s\right]_{jj_0} \delta(x - x_0) \\ &= [\hat{\mathsf{U}}(\lambda)\{1 - \hat{\mathsf{V}}(\lambda)\}^{-1}]_{jj_0}\delta(x - x_0). \end{aligned} \tag{16}$$

In this result no approximation has yet been made. It is shown in Appendix D that for the Markov case it coincides with the solution of (3).

3. Example of a non-Markovian process

In the case of chromatography there are two levels, $j = 1$ (adsorbed) and $j = 2$ (dissolved). Furthermore $\mathbb{A}_1 = 0$ and $\mathbb{A}_2 = -w(\partial/\partial x)$, where w is the constant velocity of the solvent. Let the adsorption take place at random, $\gamma_{12}(\tau) = \text{constant} = \gamma$. On the other hand, suppose that each sojourn in the adsorbed level 1 lasts exactly a time τ_0, so that $\gamma_{21}(\tau) = \delta(\tau - \tau_0)$. This is not meant as a realistic description of chromatography, but as an example of a non-Markovian composite stochastic process.

Owing to the translational invariance it is possible to apply a Fourier transformation in space

$$P_j(x, t) = \int_{-\infty}^{\infty} P_j(q, t)\, e^{iqx}\, dq \tag{17}$$

and to consider each value of q separately. Then the operator $\mathbb{A}_2$ reduces to the numerical factor $-iwq$ and

$$\hat{U}_{11}(\lambda) = \frac{1 - e^{-\lambda\tau}}{\lambda}, \quad \hat{U}_{22}(\lambda) = \frac{1}{\lambda + \gamma + iwq}, \tag{18a}$$

$$\hat{V}_{21}(\lambda) = e^{-\lambda\tau_0}, \qquad \hat{V}_{12}(\lambda) = \frac{\gamma}{\lambda + \gamma + iwq}. \tag{18b}$$

Suppose at $t = 0$ the particle is dissolved at position $x = 0$, and one asks for the probability at $t > 0$ to find it adsorbed at x. Its Fourier–Laplace transform is obtained by inserting (18) into the general formula (16), so that

$$P_1(q, t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{\lambda t}\, d\lambda\, \frac{1 - e^{-\lambda\tau_0}}{\lambda} \frac{\gamma}{\lambda + \gamma + iwq - \gamma e^{-\lambda\tau_0}}$$

$$= \sum_n e^{\lambda_n t} \frac{1 - e^{-\lambda_n\tau_0}}{\lambda_n(1 - \gamma\tau_0 e^{-\lambda_n\tau_0})}. \tag{19}$$

Here the λ_n are the poles to be found from

$$\lambda + \gamma + iwq = \gamma e^{-\lambda\tau_0}. \tag{20}$$

Now the lifetimes τ_0 and γ^{-1} of both levels are supposed to be short of order τ_c, but $\gamma\tau_0$ may have any value. It is shown in Appendix E that then the dominant pole is the one near zero, whose value is to second order in τ_c given by

$$\lambda = \frac{-iwq}{1 + \gamma\tau_0} - \frac{\gamma\tau_0^2}{2(1 + \gamma\tau_0)^3}(wq)^2 + \mathcal{O}(\tau_c^3). \tag{21}$$

All other poles have a real part less than $-\gamma$. Hence for $t \gg \gamma^{-1}$ the dominant contribution is

$$P_1(x, t) = \text{const.} \int \exp\left[iq\left(x - \frac{w}{1 + \gamma\tau_0} t\right)\right] \exp\left[-\frac{\gamma\tau_0^2}{2(1 + \gamma\tau_0)^3}(wq)^2 t\right] dq. \tag{22}$$

One recognizes propagation with an average or renormalized velocity

$$\bar{w} = \frac{\gamma^{-1}}{\gamma^{-1} + \tau_0} w, \tag{23}$$

which may be understood as the velocity of the solvent reduced by the fraction of the time spent in that solvent. In addition a Gaussian broadening of the probability density appears in the nature of a diffusion, due to the random character of the process. In fact, (22) satisfies the diffusion equation

$$\frac{\partial P_1}{\partial t} = -iw \frac{\partial P_1}{\partial x} + \frac{\gamma\tau_0^2 w^2}{2(1 + \gamma\tau_0)^3} \frac{\partial^2 P_1}{\partial x^2}. \tag{24}$$

Two approximations have been used. First the contributions of other poles have been omitted, which amounts to omitting terms that differ from the main one by a factor of type e^{-t/τ_c}. This approximation is responsible for the fact that P_1 obeys a first order differential equation in time of the form of a master equation. Secondly, the expansion of λ_0 in powers of τ_c has been cut off. This had the effect that (24) is a diffusion equation; higher orders would bring in higher derivatives on the right, but it would still have the form of a master equation.

It appears that the result is an expansion in powers of $\tau_0 wq$. (Of course this is the only dimensionless parameter apart from $\gamma\tau_0$.) Thus the expansion does not require y to be Markovian, but it requires that its sojourn times are short on the time scale determined by the propagators $\mathbb{A}$. Under that condition $X(t, Y(t))$ approximately obeys a master equation with a renormalized propagator. This renormalized propagator can be determined in successive orders; for the present example the first two are displayed in (24).

4. The long-time behavior of the general case

Returning to the general formula (16) we see that the value of $P_j(x, t)$ for large t is determined by that singularity of the analytic function [ ] of λ whose real part is largest. Both $\hat{U}(\lambda)$ and $\hat{V}(\lambda)$ are analytic down to $\mathrm{Re}\,\lambda = -1/\tau_c$, but the reciprocal will give rise to new singularities. The reciprocal is singular when the equation

$$\{1 - \hat{V}(\lambda)\}f = 0 \tag{25}$$

has a non-zero solution $f_j(x)$. More explicitly this equation reads

$$\sum_j \hat{V}_{ij}(\lambda) f_j(x) = f_i(x). \tag{26}$$

Any isolated value λ for which this equation has a solution is a pole of $\{1 - \hat{V}(\lambda)\}^{-1}$ and the residue is proportional to that solution. If there is a continuous spectrum of such values it is still true that $\{1 - \hat{V}(\lambda)\}^{-1}$ tends to infinity when λ approaches a point in the spectrum and that the coefficient is proportional to the corresponding solution $f_j(x)$.

It is shown in Appendix F that (25) has no solution when $\mathrm{Re}\,\lambda > 0$. We shall here find a solution with λ close to zero. For this purpose expand (15):

$$\begin{aligned}\hat{V}_{ij}(\lambda) &= \int_0^\infty v_{ij}(\tau)\{1 + \tau(\mathbb{A}_j - \lambda) + \tfrac{1}{2}\tau^2(\mathbb{A}_j - \lambda)^2 + \cdots\}\,d\tau \\ &= V_{ij}^{(0)} + V_{ij}^{(1)}(\mathbb{A}_j - \lambda) + V_{ij}^{(2)}(\mathbb{A}_j - \lambda)^2 + \cdots.\end{aligned} \tag{27}$$

The justification is that $v_{ij}(t)$ decreases exponentially on account of (14) and it is easily seen that the successive terms in (27) are of successive orders in τ_c. All quantities $V_{ij}^{(n)}$ are non-negative and according to (6)

$$\sum_i V_{ij}^{(0)} = 1. \tag{28}$$

Hence $V_{ij}^{(0)}$ has the left eigenvector $(1, 1, \ldots)$ with unit eigenvalue, and consequently a right eigenvector ξ_j. The theorem of Perron and Frobenius[7]) asserts that this is the only eigenvector and that its components are non-negative, so that we may normalize by requiring

$$\sum_j \xi_j = 1. \tag{29}$$

To zeroth order in τ_c equation (26) reduces to

$$\sum_j V_{ij}^{(0)} f_j(x) = f_i(x). \tag{30}$$

This equation has the solutions

$$f_j(x) = \xi_j \phi(x) \tag{31}$$

with arbitrary $\phi(x)$. To include the next order we set

$$f_j(x) = \xi_j \phi(x) + \psi_j(x) + \mathcal{O}(\tau_c^2), \tag{32}$$

and find

$$\sum_j \{V_{ij}^{(0)} - \delta_{ij}\}\psi_j(x) = -\sum_j V_{ij}^{(1)}(\mathbb{A}_j - \lambda)\xi_j \phi(x). \tag{33}$$

This equation only admits of a solution ψ_j provided that the right-hand side is orthogonal on the left eigenvector of $V_{ij}^{(0)} - \delta_{ij}$. This requirement yields an additional condition on ϕ,

$$\sum_{ij} V_{ij}^{(1)}(\mathbb{A}_j - \lambda)\xi_j \phi(x) = 0. \tag{34}$$

Denoting the average life time of each level j by T_j:

$$\sum_i V_{ij}^{(1)} = \int_0^\infty \tau \sum_i v_{ij}(\tau)\, \mathrm{d}\tau = T_j, \tag{35}$$

one has

$$\lambda \phi(x) = \frac{\sum_j T_j \xi_j \mathbb{A}_j}{\sum_j T_j \xi_j} \phi(x) \equiv \bar{\mathbb{A}} \phi(x). \tag{36}$$

The operator $\bar{\mathbb{A}}$ is an average of the several $\mathbb{A}_j$; the weights of this average are the life times multiplied with the probabilities of the several levels.

The conclusion is that (26) has – to lowest order in τ_c – solutions $f_j(x) = \xi_j \phi(x)$ when λ is an eigenvalue of $\bar{\mathbb{A}}$ and $\phi(x)$ the corresponding eigenfunction. If there are such eigenvalues with $\mathrm{Re}\,\lambda > -1/\tau_c$ they are points where $\{1 - \hat{\mathbb{V}}(\lambda)\}^{-1}$ is singular and determine the long-time behavior of

$$P_j(x, t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{t\lambda} P_j(x, \lambda)\, d\lambda. \tag{37}$$

On contracting the integration path around these singularities one is left with a sum or integral of contributions each having the form

$$e^{t\lambda} \xi_j \phi(x). \tag{38}$$

As λ and $\phi(x)$ are solutions of (36) this sum has the form

$$P_j(x, t) = \xi_j P(x, t) \tag{39}$$

with

$$\frac{\partial P(x, t)}{\partial t} = \bar{\mathbb{A}} P(x, t). \tag{40}$$

Thus *the distribution over the levels has become stationary and the common spatial distribution is governed by an averaged transport equation having the form of a master equation.*

5. Higher approximations

As the approximation scheme is of somewhat unfamiliar nature we formulate it first independently of the present context. Let $M(\lambda)$ be a matrix (or operator) involving a parameter λ and acting on vectors (or functions) f. The aim is to find an f and a λ such that $M(\lambda)f = 0$. Suppose that $M(\lambda)$ can be expanded formally in a small quantity ϵ; accordingly we expand f and λ, so that the equation to be solved is, omitting terms beyond ϵ^2,

$$[M_0(\lambda_0 + \epsilon\lambda_1 + \epsilon^2\lambda_2 + \cdots) + \epsilon M_1(\lambda_0 + \epsilon\lambda_1 + \cdots) + \epsilon^2 M_2(\lambda_0 + \cdots) + \cdots]$$
$$(f_0 + \epsilon f_1 + \epsilon^2 f_2 + \cdots) = 0. \tag{41}$$

Assume that the zeroth-order equation

$$M_0(\lambda_0) f_0 = 0 \tag{42}$$

has been solved. Then the terms of order ϵ yield

$$M_0(\lambda_0)f_1 + \{M_1(\lambda_0) + \lambda_1 M_0'(\lambda_0)\}f_0 = 0. \tag{43}$$

The prime denotes differentiation with respect to λ and henceforth we omit the argument λ_0. As M_0 has a right eigenvector f_0 according to (42), it also has a left eigenvector $\tilde{f}_0$. Multiplying by this one obtains the condition

$$(\tilde{f}_0, M_1 f_0) + \lambda_1(\tilde{f}_0, M_0' f_0) = 0. \tag{44}$$

This can be satisfied by choosing λ_1 appropriately, provided that

$$(\tilde{f}_0, M_0' f_0) \neq 0. \tag{45}$$

Subsequently one obtains f_1 from (44); the result is unique up to an additional term proportional to f_0, which remains arbitrary. The scheme is easily extended to higher orders, subject to the condition (45).

Our case, however, does not enter into this general scheme because $M_0(\lambda)$ does not actually contain λ and therefore (45) is not satisfied. The dependence of λ only comes in through the next order term $\epsilon M_1(\lambda)$, so that one has to do with a singular perturbation problem. The scheme therefore has to be modified as follows.

The zeroth order equation is, rather than (42),

$$M_0 f_0 = 0. \tag{46}$$

No parameter λ enters and this equation either has a solution f_0 or not. If not, one cannot expect to find solutions by a perturbation method. If, however, (46) has a solution one can again expand f and obtains to first order

$$M_0 f_1 + M_1(\lambda) f_0 = 0. \tag{47}$$

Multiplication with the left eigenvector of M_0 yields

$$(\tilde{f}_0, M_1(\lambda) f_0) = 0. \tag{48}$$

This determines the zeroth order eigenvalue λ_0, and subsequently

$$f_1 = -M_0^{-1} M_1(\lambda_0) f_0. \tag{49}$$

To obtain higher orders we expand λ around λ_0 and again omit the argument λ_0. To second order

$$M_0 f_2 + M_1 f_1 + \lambda_1 M_1' f_0 + M_2 f_0 = 0. \tag{50}$$

Multiplication with $\tilde{f}_0$ yields

$$-(\tilde{f}_0, M_1 M_0^{-1} M_1 f_0) + \lambda_1(\tilde{f}_0, M_1' f_0) + (\tilde{f}_0, M_2 f_0) = 0. \tag{51}$$

This determines λ_1 provided that

$$(\tilde{f}_0, M_1' f_0) \neq 0. \tag{52}$$

We shall now apply this scheme to our composite stochastic processes. All operators are acting on functions over $\mathscr{X}$, but to simplify the problem we suppose that there is translational invariance with respect to x. Then the dependence on x can be extracted by setting $f_j(x) = e^{iqx} f_j$, where f_j is a vector in $\mathscr{Y}$. The operators $\mathbb{A}_j$ reduce to numbers A_j depending on q, and all remaining operators are matrices M_{ij}. Thus we have

$$M(\lambda) = \hat{V}(\lambda) - 1, \quad \epsilon = \tau_c, \quad M_0 = V^{(0)} - 1,$$
$$M_1(\lambda) = V_{ij}^{(1)}(A_j - \lambda), \quad M_2(\lambda) = V_{ij}^{(2)}(A_j - \lambda)^2. \tag{53}$$

The zeroth order equation (46) is

$$\sum_j V_{ij}^{(0)} f_j = f_i, \tag{54}$$

with solution $f_j = \xi_j$; $\tilde{f}_j = \tilde{\xi}_j = (1, 1, \ldots)$. The eigenvalue equation (48) is

$$\sum_j V_{ij}^{(1)} \xi_j (A_j - \lambda_0) = 0, \tag{55}$$

which is the same as (34) but written for a single Fourier component.

To obtain the next order λ_1 we need M_0^{-1}. Let the other eigenvalues of $V^{(0)} - 1$ be $-\mu_a < 0$ with left eigenvectors $\tilde{\xi}^a$ and right eigenvectors ξ^a, normalized so as to have $(\tilde{\xi}^a, \xi^b) = \delta_{ab}$. Then (49) can be made more explicit:

$$f_{1i} = \sum_a \mu_a^{-1} \xi_i^a \sum_j \tilde{\xi}_j^a V_{jk}^{(1)} (A_k - \bar{A}) \xi_k. \tag{56}$$

With its aid the equation (51) is seen to be

$$\lambda_1 \left(\sum_j T_j \xi_j \right) = \sum_{ij} V_{ij}^{(2)} \xi_j (A_j - \bar{A})^2 + \sum_a \mu_a^{-1} \sum_{hijk} V_{hi}^{(1)} \xi_i^a \tilde{\xi}_j^a V_{jk}^{(1)} \times \xi_k (A_i - \bar{A})(A_k - \bar{A}). \tag{57}$$

This result will now be applied to the example of section 3. Expansion of (18) yields

$$V^{(0)} - 1 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \quad V^{(1)} = \begin{pmatrix} 0 & \gamma^{-1} \\ \tau_0 & 0 \end{pmatrix}, \quad V^{(2)} = \begin{pmatrix} 0 & \gamma^{-2} \\ \frac{1}{2}\tau_0^2 & 0 \end{pmatrix}. \tag{58}$$

Hence $\tilde{\xi} = (1, 1)$, $\xi = (\frac{1}{2}, \frac{1}{2})$, $\mu_a = -2$, $\tilde{\xi}^a = (1, -1)$, $\xi^a = (\frac{1}{2}, -\frac{1}{2})$. According to (35) one has $T_1 = \tau_0$, $T_2 = \gamma^{-1}$ and therefore

$$\lambda_0 = -iwq/(1 + \gamma\tau_0), \tag{59}$$

which is the first term in (21). Subsequently inserting the above data into (57) one obtains for λ_1 the second term of (21).

The conclusion formulated in section 3 now turns out to be general. One

sees from (58) that the expansion in ϵ is an expansion in $\tau_c\mathbb{A}$. Hence the condition under which the random parameter y in (1) can be eliminated is *not* that Y be Markovian, but that Y varies fast on the time scale on which X varies through the propagators $\mathbb{A}$.

6. The distribution of residence times

As a preliminary we verify that (16) conserves total probability, for which we have to show that

$$\int \mathrm{d}x \sum_j \hat{P}_j(x, \lambda) = \int_0^\infty \mathrm{e}^{-\lambda t}\,\mathrm{d}t \int \mathrm{d}x \sum_j P_j(x, t) = \frac{1}{\lambda}. \tag{60}$$

Owing to (7) the integration over x reduces all factors $\mathrm{e}^{\tau\mathbb{A}j}$ to unity, and one is left with an equation of matrices acting in $\mathcal{Y}$ alone

$$\int \hat{P}_j(x, \lambda)\,\mathrm{d}x = \hat{u}_j(\lambda)\{1 - \hat{v}(\lambda)\}^{-1}_{jj_0}, \tag{61}$$

where $\hat{u}$ and $\hat{v}$ are the Laplace transforms of (4). According to (5)

$$\sum_i \hat{v}_{ij}(\lambda) = 1 - \lambda\hat{u}_j(\lambda). \tag{62}$$

Substituting this value of $\hat{u}_j(\lambda)$ in (61) and summing over j one proves (60).

After this preliminary we show that the same formulation with a slight modification serves to answer the following question: Suppose one has a non-Markovian process Y, *what is the probability distribution of the total times θ_k it sojourns in any level k during the time interval* $(0, t)$? The question is further specified by selecting only those processes that start from j_0 and at t are in j. The probability for a process starting in j_0 to end up in j having sojourned in each level k during a total time θ_k will be denoted by

$$\Pi(j, t|\theta_1, \theta_2, \ldots|j_0, 0). \tag{63}$$

It has a multivariate Fourier transform

$$\Gamma(j, t|\omega_1, \omega_2, \ldots|j_0, 0) = \int \Pi(j, t|\theta_1, \theta_2, \ldots|j_0, 0) \times \mathrm{e}^{\mathrm{i}(\omega_1\theta_1 + \omega_2\theta_2 + \cdots)}\,\mathrm{d}\theta_1\,\mathrm{d}\theta_2, \ldots. \tag{64}$$

To find this function consider again a history (10). The contribution of this history to (64) is

$$\exp[\mathrm{i}\omega_j(t - t_s)]\exp[\mathrm{i}\omega_{j_{s-1}}(t_s - t_{s-1})]\ldots\exp[\mathrm{i}\omega_{j_1}(t_2 - t_1)]\exp[\mathrm{i}\omega_{j_0}t_1]. \tag{65}$$

This is the same formula as occurs in (12) if one replaces $\mathscr{X}$ by a single point and the propagators $\mathbb{A}_j$ by the variables $i\omega_j$. As a result the expression (9) for $\hat{\mathbb{V}}(\lambda)$ reduces to

$$\int_0^\infty v_{ij}(\tau)\, e^{(i\omega_j-\lambda)\tau}\, d\tau = \hat{v}_{ij}(\lambda - i\omega_j), \tag{66}$$

and similarly $\hat{\mathbb{U}}(\lambda)$ reduces to $\delta_{ij}\hat{u}_j(\lambda - i\omega_j)$. Hence (16) states

$$\hat{\Gamma}(j, \lambda|\omega_1, \omega_2, \ldots|j_0, 0) = \hat{u}_j(\lambda - i\omega_j)\{1 - \hat{v}\}^{-1}_{jj_0}, \tag{67}$$

where $\hat{v}$ is an abbreviation for the matrix (66). Summation over j yields the characteristic function of the distribution of the θ's regardless of the final level

$$\hat{\Gamma}(\lambda|\omega_1, \omega_2, \ldots|j_0, 0) = \sum_j \hat{\Gamma}(j, \lambda|\omega_1, \omega_2, \ldots|j_0, 0). \tag{68}$$

When $t \gg \tau_c$ the value of (64) is again determined by the pole with largest real part. To lowest order in τ_c it is given by (55)

$$\lambda_0 = i \sum_j T_j\xi_j\omega_j \Big/ \sum_j T_j\xi_j. \tag{69}$$

The corresponding eigenvector is ξ_j, and therefore

$$\Gamma(j, t|\omega_1, \omega_2, \ldots|j_0, 0) = \xi_j \exp\Big[it \sum_k T_k\xi_k\omega_k \Big/ \sum_k T_k\xi_k\Big]. \tag{70}$$

The exponential factor in this expression is just the characteristic function regardless of final level, comp. (68). The factor ξ_j represents the relative probabilities of the levels. The average time spent in level k is

$$\langle\theta_k\rangle = t\,\frac{T_k\xi_k}{\sum_j T_j\xi_j}. \tag{71}$$

From this result one can immediately understand why the renormalized operator $\bar{\mathbb{A}}$ was given by the average (38).

The variances and covariances of the θ's only enter through the next order of τ_c. We compute them for the two-level process $Y(t)$ defined in section 3. With the aid of (57) one obtains for large t

$$\Gamma(t|\omega_1, \omega_2|j_0, 0) = \exp\Big[it\,\frac{\tau_0\omega_1 + \gamma^{-1}\omega_2}{\tau_0 + \gamma^{-1}} - \frac{t}{2}\,\frac{\tau_0^2\gamma^{-2}}{(\tau_0 + \gamma^{-1})^3}(\omega_1 - \omega_2)^2\Big]. \tag{72}$$

For $\omega_1 = 0$, $\omega_2 = -iwq$ one recognizes (22). In this order the distribution of

θ_1, θ_2 is Gaussian and their variances and covariance, indicated by double brackets, are

$$\langle\langle\theta_1^2\rangle\rangle = \langle\langle\theta_2^2\rangle\rangle = \langle\langle\theta_1\theta_2\rangle\rangle = t\,\frac{\tau_0^2\gamma^{-2}}{(\tau_0+\gamma^{-1})^3}. \tag{73}$$

Appendix A

The connection with stochastic differential equations

A *stochastic differential equation* is a differential equation whose coefficients are *given* stochastic variables or functions, that is, their stochastic properties do not depend on the solution of the equation[8]). Clearly (1) is such an equation, because the stochastic parameter y in it is determined by (2) independently of x. Thus $X(t; Y(t))$ is a stochastic process whose probability density $p(x, t)$ obeys the stochastic differential equation (1). Although this equation looks like a master equation, the stochastic element in it has the effect that $X(t; Y(t))$ is not a Markov process, unless $Y(t)$ has zero memory time, i.e., is delta correlated.

Vice versâ, it can be seen that any stochastic differential equation may be treated as a special type of composite stochastic process with Markovian $X(t; y)$. For, let $\{x_\nu\}$ be a set of variables obeying the differential equations

$$\dot{x}_\nu = f_\nu(\{x\}; y), \tag{74}$$

where y is a stochastic function of t, or a collection of such functions. Then for each fixed value of y the probability density of $\{x\}$ obeys the Liouville equation

$$\dot{p}(\{x\}, t; y) = -\sum_\nu \frac{\partial}{\partial x_\nu} f_\nu(\{x\}; y)p(\{x\}, t; y). \tag{75}$$

This is an equation of type (1), albeit rather special inasmuch as for fixed y it describes a deterministic process.

Appendix B

Elimination of the rejuvenation

Let $Y(t)$ be a non-Markovian process with transition probabilities $\gamma_{ij}(\tau)$ in which $\gamma_{ij}(\tau) > 0$. The probability for finding it still in level j after a time τ since it arrived, without intervening transition or rejuvenation, is $u_j(\tau)$ given

by (4a). The probability per unit time for making a first transition or rejuvenation after a time τ is $v_{ij}(\tau)$ given by (4b). The probability for finding it still in j without intervening transition, but having undergone rejuvenations at times $t_1, t_2, \ldots$ is obtained in analogy with (11)

$$u_j^*(\tau) = \sum_s \int \mathrm{d}t_1 \ldots \mathrm{d}t_s u_j(\tau - t_s) v_{jj}(t_s - t_{s-1}) v_{jj}(t_{s-1} - t_{1-2}) \ldots v_{jj}(t_1). \tag{76}$$

The probability for making its first transition to a different level i after a time τ is

$$v_{ij}^*(\tau) = \sum_s \int \mathrm{d}t_1 \ldots \mathrm{d}t_s v_{ij}(\tau - t_s) v_{jj}(t_s - t_{s-1}) \ldots v_{jj}(t_1). \tag{77}$$

The quantities u_j^* and v_{ij}^* define a new process $Y^*(t)$ having no rejuvenations but with the same transition probabilities as $Y(t)$.

To estimate the long-time behavior of u_j^* and v_{ij}^* take the Laplace transform of (76)

$$\hat{u}_j^*(\lambda) = \hat{u}_j(\lambda)\{1 - \hat{v}_{jj}(\lambda)\}^{-1}. \tag{78}$$

If u_j and v_{ij} obey (14) the dominant pole is given by

$$1 = \hat{v}_{jj}(\lambda) = \int_0^\infty v_{jj}(\tau)\, \mathrm{e}^{-\lambda\tau}\, \mathrm{d}\tau \leqslant \int_0^\infty v_{jj}(\tau)\, \mathrm{d}\tau \quad (\mathrm{Re}\, \lambda \geqslant 0). \tag{79}$$

The quantity on the right is less than unity on account of (6) (unless no transitions to other levels are possible). Hence the dominant pole lies between $-1/\tau_c$ and 0, so that $u_j^*(\tau)$ decreases exponentially. The same is true for $v_{ij}^*(\tau)$, so that there is still an inequality of the type (14), although with a new $\tau_c^* > \tau_c$.

Appendix C

The initial conditions for a non-Markov process

The initial state is fully determined by the initial x_0, the initial level j_0, and the time θ since the system arrived in j_0. The probability that the system sojourns in j_0 for a period τ after $t = 0$ is the conditional probability

$$u_{j_0}(\tau \mid \theta) = \frac{u_{j_0}(\tau + \theta)}{u_{j_0}(\theta)}. \tag{80}$$

The probability per unit time for a transition to another level j_1 is

$$v_{ij_0}(\tau \mid \theta) = \gamma_{ij_0}(\tau + \theta)u_{j_0}(\tau \mid \theta). \tag{81}$$

This quantity has to be substituted for the last factor in (11). For $s = 0$ one has to substitute $u_{j_0}(t \mid \theta)$ for $u_{j_0}(t)$.

To amend the result (16) accordingly we define

$$\mathbb{U}_{ij_0}(t \mid \theta) = \delta_{ij_0}u_{j_0}(t \mid \theta), \tag{82a}$$

$$\mathbb{V}_{i_1j_0}(t \mid \theta) = v_{i_1j_0}(t \mid \theta)\, e^{t\mathbb{A}_{j_0}}. \tag{82b}$$

Then (16) becomes

$$\begin{aligned}\hat{P}_j(x, \lambda) &= \left[\hat{U}(\lambda \mid \theta) + \sum_{s=1}^{\infty} \hat{U}(\lambda)\{\hat{\mathbb{V}}(\lambda)\}^{s-1}\hat{\mathbb{V}}(\lambda \mid \theta)\right]_{ij_0}\delta(x - x_0) \\ &= [\hat{U}(\lambda \mid \theta) + \hat{U}(\lambda)\hat{\mathbb{V}}(\lambda)\{1 - \hat{\mathbb{V}}(\lambda)\}^{-1}\hat{\mathbb{V}}(\lambda \mid \theta)]_{ij_0}\delta(x - x_0).\end{aligned} \tag{83}$$

This is still exact. The long-time behavior, however, is again determined by the singularities of $\{1 - \hat{\mathbb{V}}(\lambda)\}^{-1}$ and is therefore the same as found in the text. It is not surprising that for long times the precise initial situation is irrelevant.

Appendix D

The Markov case

In this case γ_{ij} is independent of τ, so that

$$\hat{U}_{ij}(\lambda) = \frac{\delta_{ij}}{\lambda + \gamma_j - \mathbb{A}_j}, \quad \hat{\mathbb{V}}_{ij}(\lambda) = \frac{\gamma_{ij}}{\lambda + \gamma_j - \mathbb{A}_j}, \tag{84}$$

where we have put

$$\sum_i \gamma_{ij} = \gamma_j. \tag{85}$$

On the other hand, equation (3) holds and may be written

$$\dot{P}_i(x, t) = \mathbb{A}_i P_i(x, t) + \sum_j \gamma_{ij}P_j(x, t) - \gamma_i P_i(x, t). \tag{86}$$

Laplace transformation of this equation:

$$(\lambda + \gamma_i - \mathbb{A}_i)\hat{P}_i(x, \lambda) = \sum_j \gamma_{ij}\hat{P}_j(x, \lambda) + P_i(x, 0). \tag{87}$$

Comparison with (84) shows that (87) is solved by setting

$$\hat{P}_i(x, \lambda) = \sum_j \hat{U}_{ij}(\lambda)f_j(x, \lambda) \tag{88}$$

when $f_j(x, \lambda)$ obeys

$$f_j(x, \lambda) = \sum_{j_0} (1 - \hat{\mathbb{V}})^{-1}_{jj_0} P_{j_0}(x, 0). \tag{89}$$

This is precisely the expression (16).

Incidentally, if one considers the Markov process Y by itself, without connection with another process X, all operators $\mathbb{A}_j$ reduce to unity. The equation (13) then expresses the solution $P_j(t)$ of the master equation for Y as a sum of histories. This sum is the analog for master equations of the path integral for diffusion equations. It has been used as a computational technique by Gillespie[9]).

Appendix E

The solutions of equation (20)

Set $\lambda = \lambda' + i\lambda''$ (with λ', λ'' real) so that (20) is equivalent with

$$\lambda' + \gamma = \gamma\, e^{-\lambda'\tau_0} \cos \lambda''\tau_0, \tag{90a}$$

$$\lambda'' + wq = \gamma\, e^{-\lambda'\tau_0} \sin \lambda''\tau_0. \tag{90b}$$

From (90a) follows $\lambda' \leqslant 0$. Hence the factor $\gamma\, e^{-\lambda'\tau_0}$ on the right of (90b) is large of order τ_c^{-1}. Supposing wq of order unity one sees that either λ'' is large of sin $\lambda''\tau_0$ must be small. The latter possibility can be realized by taking $\lambda''\tau_0$ of order γ^{-1}. Or by taking $\lambda''\tau_0$ near a multiple of π, but then again λ'' is large of order τ_c^{-1}.

On the other hand, from (90) one finds that all solutions of (20) must lie on the curve defined by

$$(\lambda' + \gamma)^2 + (\lambda'' + wq)^2 = \gamma^2\, e^{-2\lambda'\tau_0}. \tag{91}$$

A sketch shows immediately that large λ'' implies small λ'. Hence the pole with the largest real part is the one near zero, which can be found by expanding (90). As $\lambda''\tau_0$ has to be of order τ_c the cosine in (90a) differs from unity by an amount of order τ_c^2, and therefore λ' differs from zero only in order τ_c^2. A consistent expansion of (90) yields therefore

$$\lambda' + \gamma = \gamma\{1 - \lambda'\tau_0 - \tfrac{1}{2}(\lambda''\tau_0)^2\} + \mathcal{O}(\tau_c^4), \tag{92a}$$

$$\lambda'' + wq = \gamma\tau_0\lambda'' + \mathcal{O}(\tau_c^3). \tag{92b}$$

These equations immediately lead to (21).

Appendix F

Proof that (26) *has no solution for* $\mathrm{Re}\,\lambda > 0$

For complex functions $f(x)$ over $\mathscr{X}$ we introduce L_1 norm and assume that in this norm the $\mathbb{A}_j$ have the property

$$\|e^{\tau\mathbb{A}_j} f\|_1 \leqslant \|f\|_1 \quad (0 \leqslant \tau < \infty). \tag{93}$$

This is true in our applications and whenever $\mathbb{A}_j$ is the operator of a master equation in $\mathscr{X}$. In the space of complex functions $f_j(x)$ over $\mathscr{Z}$ we introduce the norm

$$\|f_j(x)\|_{\mathscr{Z}} = \sum_j \|f_j(x)\|_1 = \sum_j \int |f_j(x)|\,dx. \tag{94}$$

The norm of the left-hand side of (26) is

$$\sum_i \Big\| \sum_j \int_0^\infty v_{ij}(\tau)\, e^{-\lambda\tau + \tau\mathbb{A}_j} f_j(x)\, d\tau \Big\|_1. \tag{95}$$

Using the triangle inequality and the fact that $v_{ij}(\tau)\, e^{-\lambda'\tau}$ is positive (where $\lambda' = \mathrm{Re}\,\lambda$) one finds for this

$$\leqslant \sum_{ij} \int_0^\infty v_{ij}(\tau)\, e^{-\lambda'\tau}\, d\tau \|e^{-i\lambda''\tau + \tau\mathbb{A}_j} f_j(x)\|_1 \tag{96a}$$

$$\leqslant \sum_{ij} \int_0^\infty v_{ij}(\tau)\, e^{-\lambda'\tau}\, d\tau \|f_j(x)\|_1. \tag{96b}$$

Owing to (6) one has for $\lambda' > 0$

$$< \sum_{ij} \int_0^\infty v_{ij}(\tau)\, d\tau \|f_j(x)\|_1 = \sum_j \|f_j(x)\|_1 = \|f_j(x)\|_{\mathscr{Z}}. \tag{96c}$$

Hence the left-hand side of (26) has a smaller $\mathscr{Z}$-norm than the right-hand side and cannot be equal to it.

For $\lambda' = 0$ the inequality in (96c) becomes an equality. A solution may then be possible if the other estimates also reduce to equalities. For the inequality in (96b) that requires

$$\|e^{\tau\mathbb{A}_j} f_j(x)\|_1 = \|f_j(x)\|_1 \quad \text{for all } j > \tau > 0. \tag{97}$$

Owing to the condition (93) the only possibility is that $f_j(x)$ is an eigenfunction of $\mathbb{A}_j$ belonging to a purely imaginary eigenvalue ia_j. To reduce the triangle

inequality in (96a) to an equality all these eigenfunctions must be the same function $\phi(x)$ apart from a constant factor. Moreover, on substituting $f_j(x) = c_j\phi(x)$ in (95) en (96a) one obtains the additional requirement for each i

$$\left|\sum_j \int_0^\infty v_{ij}(\tau)\, e^{-i\lambda''\tau + ia_j\tau}\, d\tau \cdot c_j\right| = \sum_j \int_0^\infty v_{ij}(\tau)\, d\tau \cdot |c_j|. \tag{98}$$

But each integral on the left is less than the one on the right unless $\lambda'' = a_j$. Hence the only possibility that (26) has a solution with purely imaginary $\lambda = i\lambda''$ is that all $\mathbb{A}_j$ have a common imaginary eigenvalue with a common eigenfunction.

Conversely, if that condition is satisfied one does obtain a solution by taking for λ that common eigenvalue and taking $f_j(x) = c_j\phi(x)$ where c_j is the positive solution for

$$\sum_j \int_0^\infty v_{ij}(\tau)\, d\tau \cdot c_j = c_i, \tag{99}$$

which exists on account of (6). In our applications, this case did not occur, although it did happen in section 3 that λ' is of higher order than λ''.

References

1) J.C. Giddings and H. Eyring, J. Phys. Chem. **59** (1955) 416; J.C. Giddings, J. Chem. Phys. **26** (1957) 169 and **31** (1959) 1462; J. Chromatography **3** (1960) 443.
2) H.L. Friedman and A. Ben-Naim, J. Chem. Phys. **48** (1968) 120.
3) K. Mysels, J. Chem. Phys. **24** (1956) 371.
4) K.S. Singwi and A. Sjölander, Phys. Rev. **119** (1960) 863.
5) R. Kubo, in: *Stochastic Processes in Chemical Physics*, K.E. Shuler, ed. (Interscience, New York, 1969). For further literature see ref. 8, p. 221.
6) G.H. Weiss, J. Statist. Phys. **8** (1973) 221.
7) F.R. Gantmacher, *Applications of the Theory of Matrices* (Interscience, New York, 1959); D.R. Cox and H.D. Miller, *The Theory of Stochastic Processes* (Methuen, London, 1965).
8) N.G. van Kampen, Physics Reports **24C** (1976) 172.
9) D.T. Gillespie, J. Phys. Chem. **81** (1977) 2340.