

and the DeWitt-Brehme method give the same results for slow motion, provided one does not neglect any term of the DeWitt-Brehme equation of motion when expanding it in a power series in $1/c$.

Still another point should be clarified: the meaning of neglecting R^k in the DeWitt-Brehme equation within the framework of accuracy of the seventh order in $1/c$.

We have already shown that the seventh-order term of R^k is given by Eq. (4.5); lower order terms do not exist. Equation (4.5) can also be written in the form

$${}_7R^k = \frac{2}{3}e^2(d/dt)[\ddot{\xi}^k - (M/r)_{,k}], \quad (4.12)$$

the right-hand side of which is (except for a coefficient) just the time-derivative of Newton's law of motion. Thus neglecting R^k in this approximation implies the use of Newton's law of motion, which is the lowest order of the undamped equation of motion. This alone enables us to write the result obtained, which is the traditional radiative damping force, in the form given

by Eq. (1.7), rather than conversely, as has been stated in reference 5.^{23,24}

ACKNOWLEDGMENTS

The author gratefully acknowledges the hospitality extended to him at Lehigh University by Professor P. Havas, and wishes to thank him and Dr. S. Bażański for helpful discussions.

Part of this work was done while the author was at the Technion, Haifa. Discussions with Professors N. Rosen and A. Peres are gratefully acknowledged.

I would like to thank also Professor B. S. DeWitt for his valuable criticism.

²³ The same procedure of using Newton's law of motion in order to write the traditional radiative damping force term in terms of the Riemann tensor was used by Peres (Ref. 15). It can easily be shown, however, that the form suggested by Peres [Eq. (12) in Ref. 15] for this term is *not* fully covariant.

²⁴ It will be noted that we conclude nothing about the phenomenon of preacceleration since we feel that the slow-motion approximation is not the proper tool for it.

Master Equations and Markov Processes

IRWIN OPPENHEIM*

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts

AND

KURT E. SHULER

National Bureau of Standards, Washington, D. C.

(Received 13 November 1964)

The processes described by generalized master equations (GME), derived from the Liouville equation on the basis of various physical and dynamic arguments, have been termed Markovian or non-Markovian depending upon whether the GME did not or did involve an explicit time integration. We show that these designations are not in accord with the (very specific) mathematical definitions of Markovian and non-Markovian processes. We demonstrate that the GME does not contain sufficient information to determine whether or not the stochastic process described by it is Markovian or non-Markovian.

I. INTRODUCTION

A NUMBER of generalized master equations have been derived recently.¹ Zwanzig² has shown that all of these equations are equivalent if not identical. These generalized master equations differ from the Pauli equation³ in that they involve an explicit time

integration. On the basis of this time integration, these equations have been termed "*non-Markovian*" by the various authors. We shall demonstrate that this nomenclature is quite misleading and may, in fact, lead to erroneous conclusions as to the nature of the stochastic processes since it does not agree with the well-defined *mathematical* usage of the terms *Markovian* and *non-Markovian*. We suggest, in view of the development to be presented below, that the physical generalized master equations (and the stochastic processes described by them) be termed non-Paulian to distinguish them from the Pauli equation (and processes) which do not involve explicit time integrations.

In Sec. II we present definitions of joint and conditional probabilities, we give the mathematical definition of Markov processes and we derive equations for the temporal development of the joint and con-

* This work was supported in part by the National Science Foundation.

¹ L. Van Hove, *Physica* **23**, 441 (1957); S. Nakajima *Progr. Theoret. Phys. (Kyoto)* **20**, 948 (1958); R. Zwanzig, *J. Chem. Phys.* **33**, 1338 (1960); *Lectures in Theoretical Physics (Boulder)* (Interscience Publishers, Inc., New York, 1960), Vol. III; I. Prigogine and P. Resibois, *Physica* **27**, 629 (1961); P. Resibois, *ibid.* **29**, 721 (1963); E. W. Montroll in *Fundamental Problems in Statistical Mechanics*, edited by E. G. D. Cohen (North-Holland Publishing Company, Amsterdam, 1962), pp. 230-249.

² R. Zwanzig, *Physica* **30**, 1109 (1964).

³ W. Pauli, *Probleme der Modernen Physik* (S. Hirzel, Leipzig, 1928), p. 30.

ditional probabilities. In Sec. III we discuss the physical generalized master equation and describe the basis of its derivation. In Sec. IV we compare the equations obtained from the "mathematical" and "physical" derivations and critically discuss their relationships. On the basis of this comparison, we demonstrate that the physical generalized master equation does not contain sufficient information to determine whether or not the stochastic process described by it is *Markovian* or *non-Markovian*.

II. DERIVATION OF THE "MATHEMATICAL MASTER EQUATIONS"

We begin this section by defining joint and conditional probabilities. We consider a set of random variables $y(t)$ which can assume a continuous set of values as a function of the continuous time t . We denote the r th-order *joint probability* that the variable $y(t)$ is in the range x to $x+dx$ at time t , in the range y_1 to y_1+dy_1 at time s_1, \dots , in the range y_{r-1} to $y_{r-1}+dy_{r-1}$ at time s_{r-1} by $P_r(x,t; y_1, s_1; \dots; y_{r-1}, s_{r-1}) \times dx dy_1 \dots dy_{r-1}$. The function P_r has the properties

$$P_r \geq 0, \tag{2.1}$$

$$\int P_r dy_1 = P_{r-1}(x,t; y_2, s_2; \dots; y_{r-1}, s_{r-1}), \tag{2.2}$$

$$P_r(x,t; y_1, t; y_2, s_2; \dots; y_{r-1}, s_{r-1}) = P_{r-1}(x,t; y_2, s_2; \dots; y_{r-1}, s_{r-1}) \delta(x - y_1). \tag{2.3}$$

The r th-order *conditional probability* that the variable $y(t)$ will assume a value between x and $x+dx$ at time t if it had the value y_1 at time s_1 , and y_2 at time s_2, \dots , and y_{r-1} at time s_{r-1} , with $t > s_1 > s_2 > \dots > s_{r-1}$, is denoted by $W_r(x,t/y_1, s_1; y_2, s_2; \dots; y_{r-1}, s_{r-1}) dx$. The function W_r has the properties

$$W_r \geq 0, \tag{2.4}$$

$$\int W_r dx = 1, \tag{2.5}$$

$$W_r(x,t/y_1, t; y_2, s_2; \dots; y_{r-1}, s_{r-1}) = \delta(x - y_1). \tag{2.6}$$

The r th-order joint probability density P_r can be written as

$$P_r(x,t; y_1, s_1; \dots; y_{r-1}, s_{r-1}) = P_{r-1}(y_1, s_1; \dots; y_{r-1}, s_{r-1}) \times W_r(x,t/y_1, s_1; y_2, s_2; \dots; y_{r-1}, s_{r-1}). \tag{2.7}$$

From (2.2), it then follows that

$$P_{r-1}(x,t; y_2, s_2; \dots; y_{r-1}, s_{r-1}) = \int dy_1 P_{r-1}(y_1, s_1; \dots; y_{r-1}, s_{r-1}) \times W_r(x,t/y_1, s_1; \dots; y_{r-1}, s_{r-1}). \tag{2.8}$$

$r = 2, 3, \dots$

We shall now use Eq. (2.8) to derive an expression which describes the temporal behavior of the joint probability density P_{r-1} . It follows from Eq. (2.8) that⁴

$$\begin{aligned} \dot{P}_{r-1}(x,t; y_2, s_2; \dots; y_{r-1}, s_{r-1}) &= \int dy_1 A_r(x, y_1, t; y_2, s_2; \dots; y_{r-1}, s_{r-1}) \\ &\times P_{r-1}(y_1, t; y_2, s_2; \dots; y_{r-1}, s_{r-1}), \end{aligned} \tag{2.9}$$

where \dot{P}_{r-1} is the derivative of P_{r-1} with respect to t and where

$$A_r(x, y_1, t; y_2, s_2; \dots; y_{r-1}, s_{r-1}) = \lim_{\Delta \rightarrow 0} \left[\frac{W_r(x, t+\Delta/y_1, t; \dots; y_{r-1}, s_{r-1}) - \delta(x - y_1)}{\Delta} \right] \tag{2.10}$$

is the transition probability per unit time for a transition from y_1 to x at time t if the variable $y(t)$ had the value y_2 at time s_2, \dots , and y_{r-1} at time s_{r-1} . The transition probability A_r has the properties

$$A_r(x, y_1, t; y_2, s_2; \dots; y_{r-1}, s_{r-1}) \geq 0 \text{ for } y_1 \neq x \tag{2.11}$$

and

$$\int dx A_r(x, y_1, t; y_2, s_2; \dots; y_{r-1}, s_{r-1}) = 0. \tag{2.12}$$

Equation (2.9) is the general form of the kinetic equation for the joint probabilities P_r . It is the starting point for the special cases discussed below.

When $r = 2$, Eq. (2.9) reduces to

$$\dot{P}(x,t) = \int dy_1 A(x, y_1, t) P(y_1, t), \tag{2.13}$$

where we write $P_1 \equiv P$ and $A_2 \equiv A$. We shall refer to Eq. (2.13) as the " P -mathematical master equation" (PMME). Note that the development leading to Eq. (2.13) requires a knowledge of the conditional probability $W_2(x, t+\Delta/y_1, t)$ [see e.g., Eq. (2.10)] or, alternatively, of the joint probability $P_2(x, t+\Delta; y_1, t)$ which is related to W_2 by Eq. (2.7). In the derivation of Eq. (2.13) no assumptions have been made about the nature of the probabilistic process. In fact, Eq. (2.13) applies for all probabilistic processes for which A [Eq. (2.10)] exists.

For $r = 3$, Eq. (2.9) reduces to

$$\begin{aligned} \dot{P}_2(x,t; y_2, s_2) &= \int dy_1 A_3(x, y_1, t; y_2, s_2) \\ &\times P_2(y_1, t; y_2, s_2). \end{aligned} \tag{2.14}$$

Using Eq. (2.7), the above can be rewritten in terms

⁴ A. Kolmogorov, Math. Ann. 104, 415 (1931).

of the conditional probabilities as

$$\dot{W}(x,t/y_2,s_2) = \int dy_1 A_3(x,y_1,t; y_2,s_2) W(y_1,t/y_2,s_2) \quad (2.15)$$

with $W_2 \equiv W$.

A Markov process is defined by⁵

$$W_r(x,t/y_1,s_1; y_2,s_2; \dots y_{r-1},s_{r-1}) = W(x,t/y_1,s_1) \quad (2.16)$$

for all r . Equation (2.16) is the precise mathematical statement for a process in which the conditional probability W_r that the stochastic variable $y(t)$ will assume a value between x and $x+dx$ at time t depends only on the fact that the stochastic variable y has the value y_1 at time s_1 . The conditional probability W_r does not depend on the values of the stochastic variable y at times previous to the time s_1 . The usual identification of "processes without memory" with Markov processes, as found so frequently in the physical literature, is based on a loose interpretation and extension of the condition (2.16). In the subsequent sections, we shall show that this loose identification often leads to erroneous conclusions as to the nature of the stochastic processes involved.

Equations (2.16) and (2.10) imply that for a Markov process

$$A_r(x,y,t; y_2,s_2; \dots; y_{r-1},s_{r-1}) = A(x,y,t). \quad (2.17)$$

If Eq. (2.17) is substituted into Eq. (2.15) we obtain

$$\dot{W}(x,t/y_2,s_2) = \int dy_1 A(x,y_1,t) W(y_1,t/y_2,s_2). \quad (2.18)$$

By analogy with (2.13), Eq. (2.18) is the " W -mathematical master equation" (WMME). Equation (2.18) is frequently used as a criterion for a Markov process.⁶ The WMME (2.18) is a closed equation for the conditional probabilities W since A in turn is defined in terms of W [Eq. (2.10)]. Thus W is completely determined by Eq. (2.18) and the condition (2.6).

⁵ See e.g., A. T. Barucha-Reid, *Elements of the Theory of Markov Processes and Their Applications* (McGraw-Hill Book Company, Inc., New York, 1960), p. 11; J. L. Doob, *Stochastic Processes* (John Wiley & Sons, Inc., New York, 1953), p. 80; W. Feller, *Probability Theory and its Application* (John Wiley & Sons, Inc., New York, 1956), p. 338; M. Kac, *Probability and Related Topics in Physical Sciences* (Interscience Publishers, Inc., New York, 1959), p. 145; M. Rosenblatt, *Random Processes* (Oxford University Press, New York, 1962), p. 122; R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach Publishers, New York, 1963), Vol. 1, p. 55; M. C. Wang and G. E. Uhlenbeck, *Rev. Mod. Phys.* 17, 323, 325 (1945). In none of these references, and a number of others consulted but not listed here, is there to be found any reference which specifically associates A. A. Markov with the condition (2.16).

⁶ We do not consider here so called "pseudo-Markovian" processes whose conditional probabilities obey the WMME (2.18) but do not obey condition (2.16) for all r . For the purpose of this paper, Eq. (2.18) can be taken as the definition of a Markov process. For a discussion of pseudo-Markovian processes see, e.g., P. Levy, *Compt. Rend.* 228, 2004 (1949); W. Feller, *Ann. Math. Stat.* 30, 1252 (1959).

We can now make the following statements which form part of the central thesis of this paper:

(1) *Given $A(x,y,t)$ and no other information*, it is not possible to determine whether the stochastic process is Markovian or non-Markovian. It is clear [compare Eqs. (2.15) and (2.18)] that it is not possible to find $W(x,t/y,s)$. On the other hand, it follows from Eq. (2.13), that it is possible to determine $P(x,t)$ once $P(x,0)$ is given.

(2) *Given $A(x,y,t)$ and the information that the process is non-Markovian* [Eq. (2.17) is not obeyed] it is not possible to determine $W(x,t/y,s)$ since its time dependence is described by Eq. (2.15) which involves $A_3(x,y_1,t; y_2,s_2)$. Again, $P(x,t)$ can be determined from Eq. (2.13) for a given $P(x,0)$.

(3) *Given $A(x,y,t)$ and the information that the process is Markovian* [as defined in Eq. (2.16)] it is possible to determine $W(x,t/y,s)$ by making use of Eq. (2.18) and condition (2.6). Also, $P(x,t)$ can be determined from Eq. (2.13) and $P(x,0)$.

The above statements imply that it is not possible to ascertain whether a stochastic process is Markovian or non-Markovian if the only information available is that the probability density $P(x,t)$ obeys the PMME (2.13). On the other hand, if it is known that the conditional probability $W(x,t/y,s)$ obeys the WMME (2.18) then the stochastic process is Markovian. If $W(x,t/y,s)$ does not obey the WMME (2.18) but instead some equation such as (2.15), the underlying process is clearly non-Markovian.

We now consider briefly the form of the time dependence of the conditional probability W for a non-Markovian process for which the probability density P at some *special* time s is a delta function, i.e.,

$$P(y,s) = \delta(y-z), \quad (2.19)$$

where z is some fixed value of the stochastic variable. When Eq. (2.19) is substituted into Eq. (2.8) for $r=2$ one finds

$$P(x,t) = W(x,t/z,s). \quad (2.20)$$

Substitution of Eq. (2.20) into (2.13) yields

$$\dot{W}(x,t/z,s) = \int A(x,y,t) W(y,t/z,s) dy. \quad (2.21)$$

While Eq. (2.21) is of the same form as the WMME (2.18), it should be noted that it is valid only for the *special* time s at which the initial condition (2.19) holds. The WMME (2.18), which has been derived from the Markovian condition (2.16), however, holds for *all* times s_1 .

III. THE "PHYSICAL" GENERALIZED MASTER EQUATION

The "mathematical" master equations presented in the previous section have been derived by purely

probabilistic arguments involving the existence of various types of probability densities and of a transition probability density per unit time which can be obtained from the conditional probability densities by a limiting process. The “physical” master equations, to be discussed in this section, have been derived from an entirely different viewpoint and by entirely different methods.

The usual starting point of these derivations is the Liouville equation

$$i(\partial\rho(t)/\partial t)=L\rho(t). \quad (3.1)$$

For classical systems, $\rho=\rho(p,q,t)$ is the phase-space distribution function; for quantum systems, ρ is the density matrix with elements ρ_{mn} . In either case, ρ is an N -particle distribution function. Equation (3.1) is a *deterministic* equation of motion which describes the time evolution of the quantity ρ . The Liouville operator L is a purely dynamical quantity whose explicit form depends entirely on the Hamiltonian of the system under consideration.

A number of derivations of the generalized master equation (GME) from the Liouville Eq. (3.1) can be found in the literature.^{1,7} These derivations all contain, as an essential element, a decomposition of the Liouville operator into an unperturbed part and a perturbation and thus they rest on a physical, mechanistic basis. We shall not repeat these derivations here but instead present the final result and then discuss its form and implications. We will follow here primarily the exposition of Zwanzig.¹

The GME derived by Zwanzig is

$$\frac{d\rho_{mm}(t)}{dt}=-\int_0^t dt_1 \sum_n K_{mmnn}(t_1)\rho_{nn}(t-t_1). \quad (3.2)$$

Here, ρ_{mm} is a diagonal element of the density matrix ρ and $K_{mmnn}(t)$ is the “memory kernel” which can be written in the form

$$K_{mmnn}(t)=[L_1 e^{-it(L-D)L}(1-D)L_1]_{mmnn}, \quad (3.3)$$

where D is a projection operator [such that $(D\rho)_{mn}=\rho_{mm}\delta_{mn}$] and L_1 is the perturbed part of the Liouville operator, $L=L_0+L_1$ with L_0 the unperturbed part. The kernel K is a purely deterministic dynamical quantity since it is obtained directly from the application of a projection operator to the Liouville operator.

Equation (3.2) has been derived for a specific initial condition, namely, that at the *special* time $t=0$ the density matrix is diagonal. This initial condition, which is commonly referred to as the assumption of initial random phases, implies that there are no phase correlations at time $t=0$. When more general initial conditions are assumed, the kinetic equation for the

evolution of the elements of the density matrix (or the phase-space density, or its Fourier-expansion coefficients) can no longer be written in the form of (3.2). *Probabilistic concepts enter into Eq. (3.2) only through this initial condition of random phases at the special time $t=0$.*

It will be noted that the kernel K_{mmnn} which governs the rate of transitions from state n to m is an explicit function of the time-integration variable t_1 as is the element ρ_{nn} of the density matrix. The evaluation of ρ_{mm} at time t requires a time integration from the initial time $t_1=0$ to $t_1=t$. The value and the evolution in time of ρ_{mm} at time t thus is not determined by the knowledge of the values of K_{mmnn} and ρ_{nn} at time t alone; instead the time history of the kernel and of the distribution function must be explicitly taken into account. The future ($t'>t$) therefore does not depend solely on the present ($t'=t$) but also involves the past ($0\leq t'<t$). It is essentially on the basis of these considerations that GME of the form (3.2) have been termed *non-Markovian*.^{1,2}

The passage from a so-called “non-Markovian” form to a so-called “Markovian” form of the physical GME is very instructive in illuminating the usage and meaning of these terms as employed in the above context. Such a passage has been discussed in detail by Prigogine and his co-workers^{7,8} and we shall outline their approach here. We consider that the “memory” contained in the kernel K and distribution function ρ is of the order of the duration t_{int} of a mechanical interaction, such as a collision, for instance. If we are now interested only in the long-time behavior of the matrix element ρ_{mm} for times $t\gg t_{\text{int}}$, we can rewrite Eq. (3.2) as

$$\frac{d\rho_{mm}(t)}{dt}\simeq-\int_0^\infty \sum_n K_{mmnn}(t_1)dt_1\rho_{nn}(t). \quad (3.4)$$

Under conditions spelled out clearly by Prigogine (7.8) and Zwanzig² (weak coupling, $t_{\text{int}}\ll$ relaxation times, etc.) the integration in (3.4) can be performed explicitly for a given physical system or model to yield

$$\frac{d\rho_{mm}(t)}{dt}=-\sum_n K_{mmnn}\rho_{nn}(t). \quad (3.5)$$

In the literature cited above, Eq. (3.5) is termed a “Markovian” master equation (ME) on the basis that it no longer involves the time history of the kernel or of the distribution function. The evolution of ρ_{nn} for all times $t'>t$ is now uniquely determined by the values of K and ρ at time t . It should again be noted that K_{mmnn} is a purely physical “dynamic” quantity and has no probabilistic antecedents, and that the above

⁷ For a very comprehensive account of the work of the Brussels school see I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Interscience Publishers, Inc., New York, 1962).

⁸ I. Prigogine, P. Resibois, and G. Severne, in *Proceedings of the International Seminar on the Transport Properties of Gases* (Brown University, Providence, Rhode Island, 1964).

development depends crucially on the special diagonal properties of ρ at time $t=0$.

IV. COMPARISON OF "MATHEMATICAL" AND "PHYSICAL" MASTER EQUATIONS

We are now in a position to study the connection between the "physical" and "mathematical" master equations. The first thing to note is that the phase distribution function $\rho(t)$ is the physical analog of the probability density $P(x,t)$ defined in Sec. II. The equation describing the time dependence of $P(x,t)$, i.e., the PMME (2.13), contains the transition probability density $A(x,y,t)$, which is related to the conditional probability density W by Eq. (2.10) for $r=2$, and the probability density $P(y,t)$. The equation describing the time dependence of $\rho(t)$, i.e., the GME (3.2), contains the convolution of the dynamical kernel $K(t')$ and the phase-space distribution function $\rho(t-t')$. To show the relationship between the GME and the PMME, it is necessary to establish the connection between A and K . Formally, the equivalence between the PMME (2.13) and the GME (3.2) would imply the necessary and sufficient condition

$$\int A(x,y,t)P(y,t)dy = \int \int_0^t dt_1 K(x,y,t_1)\rho(y,t-t_1)dy, \quad (4.1)$$

where we replaced the summation over discrete states in Eq. (3.2) by an integration over the continuous variable y . The transition probability $A(x,y,t)$ may well involve a time integration of the form

$$A(x,y,t) = \frac{1}{P(y,t)} \int_0^t dt_1 K(x,y,t_1)\rho(y,t-t_1), \quad (4.2)$$

i.e., it may well involve information about the time history of the dynamics and of the probability density. As a specific example of the connection between a

transition probability $A(x,y,t)$ and the dynamics of the process we may refer to the work of Montroll and Shuler⁹ and of Toda¹⁰ and George¹¹ on the derivation of the PMME and the GME for the time evolution of a system of harmonic oscillator.

The GME (3.2) or the ME (3.5) describes the time dependence of $\rho(x,t)$ and provides information, via Eq. (4.1), concerning the form of $A(x,y,t)$. *We maintain the the GME contains no other information about the nature of the stochastic process* and therefore statement (1) of Sec. II applies. To reiterate, given $A(x,y,t)$ and no other information, it is not possible to determine whether the process is Markovian or non-Markovian. In order to determine whether the physical process described by the Liouville equation is Markovian or non-Markovian it would be necessary to derive a kinetic equation for the conditional probability W or the joint probability density P_2 rather than a GME (3.2) or a ME (3.5) for the probability density $P(x,t)$. *In summary, neither the GME (3.2) which involves an integration over time nor the ME (3.5) which contains no explicit time integration contain sufficient information to determine whether the processes described by them are Markovian or non-Markovian.*

In light of the above discussion we suggest the following nomenclature for "physical" master equations to replace the terms Markovian and non-Markovian. We propose that processes described by master equations of the form of Eq. (3.5) be termed *Pauli processes*. This is in accord with the fact that the first detailed physical derivation of the ME (3.5) was presented by Pauli.³ Processes described by a GME of the form of Eq. (3.2) involving explicit time integration could then be termed *non-Paulian* processes.

ACKNOWLEDGMENTS

We wish to thank G. H. Weiss and R. Zwanzig for a number of helpful discussions.

⁹ E. W. Montroll and K. E. Shuler, J. Chem. Phys. **26**, 544 (1957).

¹⁰ M. Toda, J. Phys. Soc. (Japan) **13**, 1266 (1958).

¹¹ C. George, Physica **26**, 453 (1960).