Z. Physik B 30, 85–95 (1978)

Zeitschrift
für Physik B

On Derivations and Solutions of Master Equations and Asymptotic Representations*

P. Hänggi
Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

Received January 27, 1978

A linear semi-group for both: Markov processes and non-Markov processes as they occur in the description of macroscopic systems is introduced. The elegance of the semi-group approach is demonstrated by the derivation of the master equation for a Markov process which undergoes continuous and discontinuous jumps. By use of nonlinear transformations of stochastic processes a class of processes is found for which the whole stochastic kinetics reduces mainly to the kinetics of a general Gauss-Markov process. Further the convergence of sequences of Markov processes to a limiting Markov process is studied. In this context, a semi-group formulation for the validity of various expansion methods of master equations developed recently is given and the convergence of functionals of the original process to a limiting transformed process is investigated. Some results are illustrated for the behaviour of the stochastics in a bistable tunnel diode. A model for macroscopic irreversibility is introduced using a sequence of non-Markov processes which converges to a Fokker-Planck process. Finally a few accomplishments on some recent related works are given.

I. Introduction

In recent years the concept of master equations, stochastic differential equations and all that has become very fruitful for describing the behaviour of macroscopic systems as they occur in hydrodynamics [1–3], quantum optics [1, 4, 5], biology [6] and economic systems. In many of these cases the fluctuating macrovariables, $\mathbf{x}(t)=\{x_1(t),\ldots,x_n(t)\}$, can be represented fairly good by continuous functions. Hence, one suggests that a diffusion process, or equivalently a Fokker-Planck process offers a good approximation. The equation which governs the time-evolution of the probability $p(\mathbf{x}(t),t)$, the master equation, in most cases cannot be derived from the underlying microscopic dynamics. Normally one has to assume a kind of typical transition probabilities characterizing the macroscopic dynamics. The solutions of the resulting master equations usually cannot be given in an analytical closed form. This fact forces us to settle for tractable approximation procedures [7, 8]. On the other hand, it is worthwhile to work out classes of stochastic processes for which the whole dynamics is given in analytical form. Many master equations contain a parameter Ω in which the transition probabilities scale. Then the sequence of stochastic processes $\mathbf{x}_\Omega(t)$ may converge in the limit $\Omega\to\Omega_0$ to a typical process $\mathbf{x}(t)$ whose dynamics is of a transparent form. Expansions of master equations in terms of such a parameter Ω have found wide application [9–12]. Because macroscopic processes are only approximative Markov processes the concept of convergence of non-Markov processes to Markov processes plays an important role. The following paper treats some of these problems in more detail by use of the technique of semigroups.

In Section II we first introduce a linear semi-group for both Markov and non-Markov processes and study the derivation of the master equation for a Markov process which undergoes continuous and discontinuous jumps. Using nonlinear transfor-

* Work supported by U.S. Department of Health, Education and Welfare under Grant GM18051-06

0340-224X/78/0030/0085/$02.20

mations of stochastic processes we establish classes of Fokker-Planck processes whose dynamics reduces mainly to the dynamics of a general Gauss-Markov process. In Section III we investigate on a semi-group basis asymptotic representations of Markov processes. In particular we study the kind of convergence to a Markov process $\mathbf{x}(t)$ as well as the convergence of nonlinear transformations of the original process to a limiting Markov process $\mathbf{f}(\mathbf{x}(t))$. As an example we treat the fluctuations in a bistable Esaki diode. In Section IV we propose a model for macroscopic irreversibility and investigate the convergence of a sequence of non-Markov processes to a Fokker-Planck process. A discussion of the results obtained and a few accomplishements on some recent papers are given in Section V.

II. Semi-Group Approach and Solutions of Master Equations

The derivation of master equations for general stochastic processes can be performed using the concept of semi-groups [13] in connection with stochastic differential equations [14]. It seems that the semi-group technique is superior to any other method due to the fact that Markov chains, time-homogeneous and time-inhomogeneous processes defined over a discrete or continuous state space can be treated with one and the same method.

First we introduce a linear semi-group for Markov processes as well as for non-Markov processes, $\mathbf{x}(t)$, $t\in[0,\infty)$ defined on a probability space with an increasing family of sub σ-algebras, $\{\mathscr{F}_t\}$, characterizing the whole prehistory of the process for times $s\leqq t$. If we denote the linear space of real valued processes $f(t)$, progressively measurable with respect to $\{\mathscr{F}_t\}$, the $\mathscr{L}$ we can define an operator $\mathscr{T}(s)$ acting on the elements of $\mathscr{L}$:

$$\mathscr{T}(s)f(t)=\langle f(t+s)|\mathscr{F}_t\rangle \qquad s>0. \tag{2.1}$$

Here $\langle\,|\,\rangle$ denotes the conditional expectation which fulfills for the two events X, Y and condition Σ the properties [14]:

$$\langle\langle X|\Sigma\rangle\rangle=\langle X\rangle \tag{2.2}$$

$$\langle aX+bY|\Sigma\rangle=a\langle X|\Sigma\rangle+b\langle Y|\Sigma\rangle \tag{2.3}$$

$$\langle f(X)|X\rangle=f(X) \tag{2.4}$$

and for $\Sigma_1 c\Sigma_2$:

$$\langle\langle X|\Sigma_1\rangle|\Sigma_2\rangle=\langle\langle X|\Sigma_2\rangle|\Sigma_1\rangle=\langle X|\Sigma_1\rangle. \tag{2.5}$$

Hence we obtain from

$$\begin{aligned}\mathscr{T}(r)\mathscr{T}(s)f(t)&=\langle\langle f(t+s+r)|\mathscr{F}_t\rangle|\mathscr{F}_{t+s}\rangle; \qquad s,r>0\\ &=\langle f(t+s+r)|\mathscr{F}_t\rangle=\mathscr{T}(r+s)f(t),\end{aligned} \tag{2.6}$$

a linear semi-group of operators $\mathscr{T}(s)$ on $\mathscr{L}$, the linear semi-group of conditional shifts. Note that this semi-group is substantially different from the propagators acting on the space of probabilities introduced in recent works [15, 16]. What is now specific for a Markov process $\mathbf{x}(t)$ in terms of the semi-group $\mathscr{T}(s)$? To answer this questions let M_1 denote the linear space of processes $f(\mathbf{x}(t),t)$ and similarly M_2 the linear space of processes $f(\mathbf{x}(t))$. Using $f_0\in M_1$ and the concept of conditional transition probabilities R we have in general

$$\begin{aligned}\mathscr{T}(s)f_0(t)&=\langle f_0(t+s)|\mathscr{F}_t\rangle, \qquad s>0\\ &=\int f_0(\mathbf{y},t+s)\,R(\mathbf{y}(t+s),t+s|\text{prehistory})\,d\mathbf{y}\end{aligned} \tag{2.7}$$

whereas in the Markov case

$$\begin{aligned}\mathscr{T}(s)\,f_0(t)&=\bar{f}_0(\mathbf{x}_0,t+s)\\ &=\int f_0(\mathbf{y},t+s)\,R(\mathbf{y}(t+s),t+s|\mathbf{x}_0(t),t)\,d\mathbf{y}.\end{aligned} \tag{2.8}$$

In terms of stochastic processes, $\bar{f}_0(\mathbf{x}_0;t+s)$ is the conditional expectation of $f_0(\mathbf{x}(t+s),t+s)$ on hypothesis that $\mathbf{x}(t)=\mathbf{x}_0$. We observe that $\mathbf{x}(t)$ is Markov if $\mathscr{T}(s)$ leaves M_1 invariant and is Markov and temporally homogeneous if $\mathscr{T}(s)$ leaves M_2 invariant.

Dealing with Markov processes one can use the whole powerful mathematical apparaturs for semi-groups developed in this field [13]. It is well known [13, 14] that the probabilities are the solutions of the Kolmogrov-forward equation and the conditional expectations $\langle f_0(\mathbf{x}(t),t)|\mathbf{x}(s)\rangle$ the solutions of the Kolmogrov-backward equation with respect to differention of the former time s. With differentiation with respect to the later time t, the conditional expectations fulfill the equation [17]: (we use usual operator notation)

$$\frac{\partial}{\partial t}\langle f(t)|s\rangle=\langle\Gamma^+(t)f(t)|s\rangle+\left\langle\frac{\partial f(t)}{\partial t}\middle|s\right\rangle \qquad s<t. \tag{2.9}$$

In (2.9) $\Gamma^+(t)$ denotes the transpose of the forward generator $\Gamma(t)$. To elucidate more the technique of semi-groups and to present the advantages of an equation like (2.9) we study the Markov process composed of continuous and discontinuous jumps with the (Ito)-stochastic differential equation

$$d\mathbf{x}(t)=\mathbf{a}(\mathbf{x}(t),t)\,dt+\mathbf{B}(\mathbf{x}(t),t)\,d\mathbf{w}(t)+d\boldsymbol{\xi}_p(t). \tag{2.10}$$

Here, $\mathbf{w}(t)$ denotes the standard $(n\times 1)$ normalized vector Wiener process and $\boldsymbol{\xi}_p$ is a $(n\times 1)$-generalized Poisson process with rate vector $\boldsymbol{\lambda}(t)$ and jump probabilities $\mathbf{p}=(p_{b_1},\ldots,p_{b_k},\ldots,p_{b_n})$:

$$\xi_p^{(i)}(t)=\sum_{k=1}^{m_i(t)}[b_i^{(k)}=y]\,\theta(t-\tau_k), \qquad i=1,\ldots,n. \tag{2.11}$$

$m_i(t)$ is the simple Poisson counting process where with $\bar{\lambda}_i(t)=\int_0^t \lambda_i(\tau)\,d\tau$, $t>r$

$$p((m_i(t)-m_i(r))=k) = \frac{(\bar{\lambda}_i(t)-\bar{\lambda}_i(r))^k}{k!}\exp(\bar{\lambda}_i(r)-\bar{\lambda}_i(t)). \tag{2.12}$$

The independent random variables $\{b_i, i=1\ldots n\}$ are independent of $\{m_i(t)\}$ and distributed with probability $p_{b_i}(y)$. $\theta(t-\tau_i)$ is the unit step function and $\{\tau_i\}$ are the arrival times of the Poisson counting process $m(t)$. In the case $\bar{\lambda}_i(t)=\bar{\lambda}_i t$, $\forall\, i$, $\xi_p(t)$ is time-homogeneous. If we choose $f_0(\mathbf{x})=\exp i\omega\mathbf{x}$, we obtain

$$\frac{\partial}{\partial t}\langle e^{i\omega\mathbf{x}(t)}|x(s)\rangle = \lim_{\Delta t\to 0^+}\frac{1}{\Delta t}[\langle e^{i\omega\mathbf{x}(t+\Delta t)}|\mathbf{x}(s)\rangle - \langle e^{i\omega\mathbf{x}(t)}|\mathbf{x}(s)\rangle]. \tag{2.13}$$

With

$$\langle\exp i\omega\mathbf{x}(t+\Delta t)|\mathbf{x}(s)\rangle = \iint R(\mathbf{y}, t+\Delta t|\mathbf{v}, t)\,R(\mathbf{v}, t|\mathbf{x}, s)\cdot\exp(i\omega[\mathbf{y}-\mathbf{v}+\mathbf{v}])\,d\mathbf{v}\,d\mathbf{y} = \langle\exp i\omega\mathbf{x}(t)\langle\exp i\omega(\mathbf{x}(t+\Delta t)-\mathbf{x}(t))|\mathbf{x}(t)\rangle|\mathbf{x}(s)\rangle \tag{2.14}$$

we obtain by use of (2.4)

$$\frac{\partial}{\partial t}\langle\exp i\omega\mathbf{x}(t)|\mathbf{x}(s)\rangle = \langle\exp i\omega\mathbf{x}(t)\,\gamma(\omega, \mathbf{x}(t), t)|\mathbf{x}(s)\rangle, \tag{2.15}$$

where

$$\gamma(\omega, \mathbf{x}(t), t) = \lim_{\Delta t\to 0^+}\langle(\exp i\omega\, d\mathbf{x}(t))-1|\mathbf{x}(t)\rangle/\Delta t. \tag{2.16}$$

Using again (2.4) and the fact that the two noise processes are independent we have from (2.10)

$$\langle(\exp i\omega\, d\mathbf{x}(t))|\mathbf{x}(t)\rangle = \exp i\omega\mathbf{a}(\mathbf{x}(t)\,dt\,\langle\exp i\omega\,\mathbf{B}(\mathbf{x}(t), t)\,d\mathbf{w}(t)\rangle\cdot\langle\exp i\omega d\xi_p(t)\rangle. \tag{2.17}$$

Because $d\mathbf{w}(t)$ is Gaussian, we obtain

$$\langle\exp i\omega\,\mathbf{B}(\mathbf{x}(t), t)\,d\mathbf{w}(t)\rangle = \exp -\tfrac{1}{2}\omega\,\mathbf{D}(\mathbf{x}(t), t)\,\omega\,dt, \tag{2.18}$$

with

$$\mathbf{D}(\mathbf{x}(t), t) = \mathbf{B}(\mathbf{x}(t), t)\,\mathbf{B}^+(\mathbf{x}(t), t). \tag{2.19}$$

Further the probability, P, of two and more Poisson jumps is of the order $o(dt)$ such that

$$\begin{aligned}\langle\exp i\omega d\xi_p(t)\rangle &= 1\,P[\text{no jumps}] + \sum_{j=1}^n\langle\exp i\omega_j b_j\rangle\,P[\text{only one jump in } d\xi_p^{(j)}]\\ &= \prod_{j=1}^n(1-\lambda_j(t)\,dt) + \sum_{j=1}^n\langle\exp i\omega_j b_j\rangle\,\lambda_j(t)\,dt\prod_{k\neq j}^n(1-\lambda_k(t)\,dt)\\ &= 1-\sum_{j=1}^n\lambda_j(t)\,dt[1-\langle\exp i\omega_j b_j\rangle]+o(dt).\end{aligned} \tag{2.20}$$

The final result for (2.13) reads therefore

$$\begin{aligned}\frac{\partial}{\partial t}&\langle\exp i\omega\mathbf{x}(t)|\mathbf{x}(s)\rangle\\ &=\langle\exp i\omega\mathbf{x}(t)\,[i\omega\mathbf{a}(\mathbf{x}(t), t)-\tfrac{1}{2}\omega\,\mathbf{D}(\mathbf{x}(t), t)\,\omega\\ &\quad-\sum_{j=1}^n\lambda_j(t)\,[1-\langle\exp i\omega_j b_j\rangle]|\mathbf{x}(s)\rangle.\end{aligned} \tag{2.21}$$

From (2.21) one immediately can read off with (2.9) the form of the generator $\Gamma(t)$ in the master equation (or by use of an inverse Fourier transform of the characteristic function for the condition probability) yielding

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = -\nabla\cdot\mathbf{a}(\mathbf{x}, t)\,p(\mathbf{x}, t)+\tfrac{1}{2}\nabla\nabla:\ \mathbf{D}(\mathbf{x}, t)\,p\mathbf{x}, t) + \sum_{i=1}^n\lambda_i(t)\,[\int p_{b_i}(x_i-y_i)p(x_1, \ldots, y_i, \ldots, x_n, t)\,dy_i - p(\mathbf{x}\,t)]. \tag{2.22}$$

Setting $p_{b_i}(y)=\delta(y)$ for all i (zero jump amplitudes) we obtain the general Fokker-Planck equation for a time-inhomogeneous diffusion process whereas with $n=1$, $a_1=0$ and $p_{b_1}(y)=\delta(y-1)$ one obtains the usual Poisson death process.

All the detailed dynamics (finite dimensional probabilities) of a Markov process is determined if the Green's function of (2.22), the conditional probability, and the initial single-event probability $p(\mathbf{x}, 0)$ are known. Unfortunately an analytical solution for the conditional probability is in most cases not available. For (multivariable) Markov systems only a few classes of processes (time-homogeneous Gaus-Markov processes [1], linear birth and death processes [6]) are known, where one has succeeded in the solution for the whole dynamics. Hence, approximative perturbation schemes play an important role [7–11]. However, in a lot of cases one is not necessarily interested in the detailed form of the probabilities either because some appropriate mean values or some low order correlation functions describe the system behavior adequately well. In such cases the above difficulty can be overcome in constructing a

cumulant expansion [18] or a continued fraction expansion [17, 19] which yields normally a satisfactory analytical continuation of a badly converging series (e.g. a Talor series of a correlation function for large times). Nevertheless, for some processes the complete solution of the master equation may be obtained by use of a nonlinear transformation $\mathbf{y}(t) = \mathbf{f}(\mathbf{x}(t), t)$ of the original process $\mathbf{x}(t)$. If the transformed process $\mathbf{y}(t)$ fulfills a simpler master equation with conditional probability $\hat{R}(t|s)$ the original multivariate distributions of $\mathbf{x}(t)$ are obtained as follows: Let $\mathbf{g}$ denote the inverse function: $\mathbf{g}(\mathbf{f}(\mathbf{x}, t), t) = \mathbf{x}$, $\mathbf{f}(\mathbf{g}(\mathbf{x}, t), t) = \mathbf{x}$, which we assume to exist.* Then with the Jacobian, $\| \cdot \|$, of the inverse transformation we obtain

$$R(\mathbf{x}t|\mathbf{y}s) = \frac{\hat{R}(\mathbf{u}=\mathbf{f}(\mathbf{x},t), t|\mathbf{v}=\mathbf{f}(\mathbf{y},s), s)}{\left\| \frac{\partial \mathbf{g}}{\partial \mathbf{u}} \right\|_{\mathbf{u}=\mathbf{f}(\mathbf{x},t)}} \quad t > s, \tag{2.23}$$

and for the multivariate probability $p^{(n)}(t_1, \ldots, t_n)$, $t_1 \leqq \cdots \leqq t_n$, of the Markov process, $\mathbf{x}(t)$:

$$p^{(n)}(\mathbf{x}_1, t_1; \ldots; \mathbf{x}_n, t_n) = \prod_{i=2}^{n} \frac{\hat{R}(\mathbf{u}_i, t_i | \mathbf{u}_{i-1}, t_{i-1})}{\left\| \frac{\partial \mathbf{g}}{\partial \mathbf{u}} \right\|_{\mathbf{u}_i = \mathbf{f}(\mathbf{x}_i, t_i)}} \cdot \frac{\hat{p}(\mathbf{u}_1, \mathbf{t}_1)}{\left\| \frac{\partial \mathbf{g}}{\partial \mathbf{u}} \right\|_{\mathbf{u}_1 = \mathbf{f}(\mathbf{x}_1, t_1)}}. \tag{2.24}$$

In particular, for a one dimensional Fokker-Planck process with an (Ito)-stochastic differential equation [14]

$$dx(t) = a(x,t)\,dt + B(x,t)\,dw(t) \tag{2.25}$$

we obtain for any $f(x,t)$:strict monotonic, $f \in C^2$ and $\dot{f} = \partial f(x,t)/\partial t$ assumed to exist, for the process $y(t) = f(x,t), t)$ by use of the (Ito)-differentiation rule [14]:

$$dy(t) = \bar{a}(y,t)\,dt + \bar{B}(y,t)\,dw(t) \tag{2.26}$$

with

$$\bar{a}(y,t) = \dot{f}(y,t)|_{y=g(x,t)} + \left.\frac{\partial f(y,t)}{\partial y}\right|_{y=g(x,t)} \cdot a(y=g(x,t),t) + \frac{1}{2}\left.\frac{\partial^2 f(y,t)}{\partial y^2}\right|_{y=g(x,t)} \cdot B^2(y=g(x,t),t) \tag{2.27}$$

$$\bar{B}(y,t) = \left.\frac{\partial f(y,t)}{\partial y}\right|_{y=g(x,t)} \cdot B(y=g(x,t),t). \tag{2.28}$$

For example the choice, $f(x,t) = \int_a^x \frac{1}{B(y,t)} dy$ yields $\bar{B}(y,t) = 1$; and for time-homogeneous Fokker-Planck-processes, $x(t)$, the choice

$$f(x) = C_1 + C_2 \int_0^x \exp\left\{ -\int_0^z \frac{2a(y)}{B^2(y)}\,dy \right\} dz \tag{2.29}$$

yields $\bar{a}(y,t) = 0$. In this case we obtain

$$\bar{B}(y = f(x)) = B(x) \exp -\int_0^x \frac{2a(y)}{B^2(y)}\,dy \tag{2.30}$$

yielding for the stationary probability $p_{st}(y) = c/\bar{B}^2(y)$ or for the original process $x(t)$ an old known result [20]

$$p_{st}(x) = \frac{c}{B^2(x)} \exp 2 \int_0^x \frac{a(y)}{B^2(y)}\,dy. \tag{2.31}$$

Especially, in all cases where $\int dy/\bar{B}^2(y)$ is not bounded, the process $x(t)$ does not possess a stationary probability (system disperses to infinity: $p_{st}(x) \equiv 0$). The above transformations simplify also the Onsager-Machlup functional in the path integrals for one-dimensional processes [21, 22].

The main purpose is now to find a condition under which there exists a nonlinear transformation so that the transformed process $y(t)$ reduces to the following general Gauss-Markov process [23]:

$$dy(t) = \alpha(t)\,dt + \beta(t)\,y\,dt + \gamma(t)\,dw(t). \tag{2.32}$$

For this process, the conditional probability is explicitly known in terms of the three functions $\alpha(t)$, $\beta(t)$, $\gamma(t)$ [23]. Putting in (2.27–2.28) $y = f(x,t)$ we obtain in shorthand notation which from (2.27) is self-explanatory

$$\dot{f} + af' + \tfrac{1}{2}B^2 f'' = \alpha(t) + \beta(t) f \tag{2.33}$$

$$f' = \frac{\gamma(t)}{B}. \tag{2.34}$$

By differentiation of (2.33) with respect to x

$$\dot{\gamma} + \gamma B\left[-\frac{\dot{B}}{B^2} + \left(\frac{a}{B}\right)' - \frac{1}{2}B'' \right] = \beta(t)\,\gamma(t), \tag{2.35}$$

we obtain for the condition that there exists a transformation f such that $y(t)$ fulfills (2.32):

$$\frac{d}{dx}\left\{ B\left[\frac{\dot{B}}{B^2} - \left(\frac{a}{B}\right)' + \frac{1}{2}B'' \right] \right\} = 0. \tag{2.36}$$

When $x(t)$ *time-homogeneous* we can look for the most general function $f(x)$ such that y is a time-homogeneous Gauss-Markov process. Then a straightforward calculation gives from (2.33–2.34)

* More general we have with $y = g(x)$ and the *real* roots, $y = a = g(a_1) = \cdots = g(a_n)$, for the probability p_y: $p_y(a) = \sum_{i=1}^{n} p_x(a_i) \Big/ \left\| \frac{\partial g}{\partial x} \right\|_{x=a_i}$.

$$f(x)=\gamma\int_0^x \frac{1}{B(y)}\,dy+\text{const.} \tag{2.37}$$

As one simple example for this case we mention the Fokker Planck operator

$$\Gamma_{\mathrm{FP}}=+\tfrac{1}{2}V(\sin x\cos x+b\sin x)$$
$$+\tfrac{1}{2}V^2(\cos x+b)^2. \tag{2.38}$$

On the other hand, if one looks for a transformation yielding a in general *time-inhomogeneous* process $y(t)$, where $\beta(t)=0$ we obtain

$$\frac{\dot{\bar{B}}}{\bar{B}}=B\left[1/2B''-\left(\frac{a}{B}\right)'\right]=c. \tag{2.39}$$

Hence $f(x,t)$ can be chosen to be

$$f(x,t)=\exp ct\int_0^x \frac{dy}{B(y)}, \tag{2.40}$$

and for the drift we find

$$\bar{a}(t)=\text{const}\,\bar{B}(t)=\text{const}\exp ct. \tag{2.41}$$

Similar conditions can be derived for Fokker-Planck systems, where the conditional probability is known in function of a series of orthogonal polynomials [6, 24].

III. On Asymptotic Representations of Markov Master Equations

In a lot of interesting cases the foregoing method cannot be applied. This especially holds for multivariable systems where in addition perturbation techniques for the dynamic behaviour are poorly developed. Fortunately, the transition probabilities per unit time, $W(\mathbf{X}'\to\mathbf{X};t)\equiv W(\mathbf{X},\mathbf{X}';t)$, fulfill often in terms of a *dimensionless* parameter Ω* a scaling law of the form

$$W(\mathbf{X},\mathbf{X}';t)=f(\Omega)\,\Omega\left[\phi_0\left(\frac{\mathbf{X}'}{\Omega},\mathbf{X}-\mathbf{X}';t\right)\right.$$
$$\left.+\Omega^{-1}\phi_1\left(\frac{\mathbf{X}'}{\Omega},\mathbf{X}-\mathbf{X}';t\right)+\cdots\right]. \tag{3.1}$$

The factor $f(\Omega)$ is innocuous because it can be absorbed in a redefinition of the time variable. (In the following $f(\Omega)$ is set equal to 1). If we denote by $\mathbf{x}_\varepsilon(t)=\mathbf{X}(t)/\Omega$ the stochastic process of the "intensive" macrovariables we obtain with $\varepsilon=\Omega^{-1}$ for the master equation

$$\varepsilon\frac{\partial P(\mathbf{x}_\varepsilon,t)}{\partial t}=\int\{\phi_0(\mathbf{x}_\varepsilon-\varepsilon\mathbf{Y},\mathbf{Y};t)$$
$$+\varepsilon\phi_1(\mathbf{x}_\varepsilon-\varepsilon\mathbf{Y},\mathbf{Y};t)+\cdots\}P(\mathbf{x}_\varepsilon-\varepsilon\mathbf{Y},t)\,d\mathbf{Y}$$
$$-P(\mathbf{x}_\varepsilon,t)\int\{\phi_0(\mathbf{x}_\varepsilon,\mathbf{Y};t)+\varepsilon\phi_1(\mathbf{x}_\varepsilon,\mathbf{Y};t)+\cdots\}\,d\mathbf{Y}. \tag{3.2}$$

Each of the functions ϕ_j has a Taylor expansion with respect to the first argument. Hence we obtain the Kramers-Moyal expansion

$$\frac{\partial P(\mathbf{x}_\varepsilon,t)}{\partial t}=\sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\varepsilon^{n-1}\left(\frac{\partial}{\partial\mathbf{x}_\varepsilon}\right)^n c_n(\mathbf{x}_\varepsilon,t)\,P(x_\varepsilon,t), \tag{3.3}$$

where

$$c_n(\mathbf{x}_\varepsilon,t)$$
$$=\int(\mathbf{Y})^n[\phi_0(\mathbf{x}_\varepsilon,\mathbf{Y};t)+\varepsilon\phi_1(\mathbf{x}_\varepsilon,\mathbf{Y};t)+\cdots]\,d\mathbf{Y}$$
$$=c_n^0(\mathbf{x}_\varepsilon,t)+\varepsilon c_n^1(\mathbf{x}_\varepsilon,t)+\varepsilon^2c_n^2(\mathbf{x}_\varepsilon,t)+\cdots. \tag{3.4}$$

Following van Kampen [9] one can derive formally a master equation for the fluctuations, $\zeta_\varepsilon(t)$, around the deterministic solution

$$\dot{\mathbf{x}}(t)=\mathbf{c}_1^0(\mathbf{x}(t),t),\qquad \mathbf{x}(0)=\mathbf{x}_0 \tag{3.5}$$
$$\Omega^{-\frac{1}{2}}\zeta_\varepsilon(t)=\mathbf{x}_\varepsilon(t)-\mathbf{x}(\mathbf{x}_0,t). \tag{3.6}$$

In the limit $\varepsilon\to 0$ this procedure yields formally the in general time-inhomogeneous Gauss-Markov master equation for the Markov process $\zeta_{\varepsilon=0}(t)\equiv\zeta(t)$ [9, 10]

$$\frac{\partial P(\zeta,t)}{\partial t}=-\nabla_\zeta[\nabla_x\mathbf{c}_1^0(\mathbf{x}(t),t)]\,\zeta P(\zeta,t)$$
$$+\tfrac{1}{2}\nabla_\zeta\nabla_\zeta:\ \mathbf{c}_2^0(\mathbf{x}(t),t)\,P(\zeta,t). \tag{3.7}$$

Although the problem of the validity and convergence concept of such an expansion was never investigated seriously in physical applications the above master equation has been applied successfully in many problems [see e.g. in 9, 10].

The above expansion method poses the following questions:

(1) Under what conditions is it possible to pass from (3.3) to the Gauss-Markov equation (3.7)?

(2) How does the sequence of processes $\xi_\varepsilon(t)$ converge to the process $\xi(t)$?

(3) What can be said about the convergence of the processes $\xi_\varepsilon(t)=\mathbf{f}(\zeta_\varepsilon(t))$?

(4) What when $\{\zeta_\varepsilon(t)\}$ does not converge?

A remarkable amount of work on the convergence of sequences of Markov processes to a limiting Markov process has been performed in the mathematical literature [25, 26, 27].* This work seems to have been overlooked by the physicists and partially by the

* We do not restrict here the meaning of the parameter Ω measuring the size of the system. It may also measure a kind of nonlinearity or characterizing the system behaviour for large times etc.

* Unfortunately, these works are not always written in the most comprehensible form for physicists

mathematicians itself. In the "Van Kampen case" the expansion procedure is justified by Kurtz's and Norman's work [27]. If the fluctuations around the deterministic path remain of the order $O(\Omega^{-\frac{1}{2}})$ at any time (*no* anomalous fluctuations with $\Omega^{-\frac{1}{2}}\zeta_\varepsilon(t)=O(1)$) and if the probability $P_{\zeta_\varepsilon}(\mathbf{u},0)$ converges to $P_\zeta(\mathbf{u},0)$, then the *finite dimensional probabilities* of $\{\zeta_\varepsilon(t)\}$ converge to the finite dimensional probabilities of the process $\zeta(t)$ (convergence in distribution). The mathematical conditions that the sequence of processes, $\zeta_\varepsilon(t)$, converges in distribution to $\zeta(t)$ have been studied in detail in terms of the initial fluctuation behaviour of $\xi_\varepsilon(0)$ and the transition function properties of $\mathbf{x}_\varepsilon(t)$ in the mathematical literature [27]. In particular, a correct and complete discussion represents the investigation by Norman [27b].

However, note that this result would mean little in problems where we need the convergence of the probabilities of a functional $\xi_\varepsilon(t)=\mathbf{f}(\zeta_\varepsilon(t))$. It is then necessary to use a stronger concept of convergence. Let us denote by $\mathbf{y}_\varepsilon=\xi_\varepsilon(t)$ the random variable for fixed t. Then a *necessary* and *sufficient* condition for convergence of the probability of $\mathbf{y}_\varepsilon$ to the probability of $\mathbf{y}=f(\zeta(t))$ of all continuous f is that (weak convergence of P_{ζ_ε} to P_ζ):

$$\lim_{\Omega\to\infty}\int \mathbf{f}(\mathbf{u})\,P_{\zeta_\varepsilon}(\mathbf{u}=\zeta_\varepsilon(t))\,d\mathbf{u}=\int \mathbf{f}(\mathbf{u})\,P_\zeta(\zeta(t)=\mathbf{u})\,d\mathbf{u} \tag{3.8}$$

for *all* bounded functionals $\mathbf{f}$ which are continuous with respect to the metric ρ of the space of sample functions of $\zeta_\varepsilon(t)$ and $\zeta(t)$. To see that this is true, note that convergence of the sequence of probabilities of $\mathbf{y}_\varepsilon$ to the probability of $\mathbf{y}$ and the boundedness of $\mathbf{f}$ imply convergence of $\langle\mathbf{f}(\zeta_\varepsilon(t))\rangle$ to $\langle\mathbf{f}(\zeta(t))\rangle$ and hence validity of (3.8). On the other hand, (3.8) implies, that the sequence of characteristic functions of $\mathbf{y}_\varepsilon$ converges to the characteristic function of $\mathbf{y}$:

$$\begin{aligned}&\lim_{\Omega\to\infty}\langle\exp i\omega\mathbf{y}_\varepsilon\rangle\\&=\lim_{\Omega\to\infty}\int\exp i\omega\,\mathbf{f}(\mathbf{u})\,p_{\zeta_\varepsilon}(\mathbf{u}=\zeta_\varepsilon(t))\,d\mathbf{u}\\&=\int\exp i\,\omega\,\mathbf{f}(\mathbf{u})\,p_\zeta(\mathbf{u}=\zeta(t))\,d\mathbf{u}.\end{aligned} \tag{3.9}$$

We define [13, 26] that the process $\zeta_\varepsilon(t)$ converges *weakly* to the process $\zeta(t)$ if the finite dimensional probabilities of $\{\zeta_\varepsilon(t)\}$ converge weakly to the finite dimensional probabilities of $\zeta(t)$. This problem has been considered in more detail for sequences of Markov processes by Liggett [28], who shows that if the finite dimensional probabilities of a sequence of diffusion or birth and death processes converge to those of a *diffusion*, then the processes converge weakly! (nontrivial)

Hence, we obtain: Under the assumptions stated in Reference 27, the sequence of the processes $\zeta_\varepsilon(t)$ converges *even weakly* to a (in general time-inhomogeneous) Gauss-Markov process. Remark however, that this does not imply that the sample paths of the processes $\{\zeta_\varepsilon(t)\}$ converge to those of the Gauss-Markov process $\zeta(t)$!

The success with the "Van Kampen" expansion in (3.7) in many cases raised some people to claim that the use of a "nonlinear" Fokker-Planck equation, i.e. with nonlinear drift and diffusion coefficients, is inconsistent. This is false because it is known for a long time that in systems with multiple steady states [12, 29–31], problems with diffusion in periodic potentials as they occur in the description of superionic conductors or for the description of the current fluctuations in a one-dimensional superconductor, in systems undergoing a limit cycle transition or even a strange attractor transition, that the expansion method based on the smallness of the fluctuations ($\zeta_\varepsilon=O(\Omega^0)$) around the deterministic path breaks down completely or is valid only on a restricted short time range. On the other hand, the use of a nonlinear Fokker-Planck equation obtained from a truncated Kramers-Moyal expansion yields very good results [1, 4, 11, 12, 30, 31]. This fact was most clearly demonstrated in the two recent works by Horsthemke et al. [11, 12] for systems with absorbing states and systems with multiple stationary states. They also give an important discussion between the *differences* of the two expansion methods which will not be repeated here. Likewise, by use of an extensive complete numerical eigenvalue and eigenvector analysis satisfactory agreement, between the results of the exact master equation of the nonlinear birth and death type and the nonlinear Fokker-Planck equation for the stochastics of the charge fluctuations $N(t)$ in a bistable tunneldiode (Esaki diode [1, 32]) for both; the *stationary* and *time-dependent* behaviour (autocorrelation function) was found [31]. Figure 1 shows the ratio R of the two stationary probabilities

$$R(N)=\frac{P_{\mathrm{st}}^{\mathrm{NFP}}(N)}{P_{\mathrm{st}}^{\mathrm{exact}}(N)} \tag{3.10}$$

for three different values of the pump parameter λ, the supply current through a vacuumdiode, in a small system with parameter value $\Omega=1(N\simeq 100$ at 500 mV, which corresponds to a junction area of the Esaki diode: $A_{\Omega=1}=10^{-11}\,\mathrm{cm}^2$). We note that the region around the main maximum is well approximated; also the case with two maxima of nearly equal height. Deviations occur in the region where N is *not* centered around the peaks with maximal probability ($p_{\mathrm{st}}(N)\approx 1$). In this region, ($N\notin\{N|N\approx N_{\mathrm{Max}}\pm\langle\delta N^2\rangle^{\frac{1}{2}}\}$), we have however $p_{\mathrm{st}}(N)\ll 1$ so that the errors for all stochastic quantities which are most

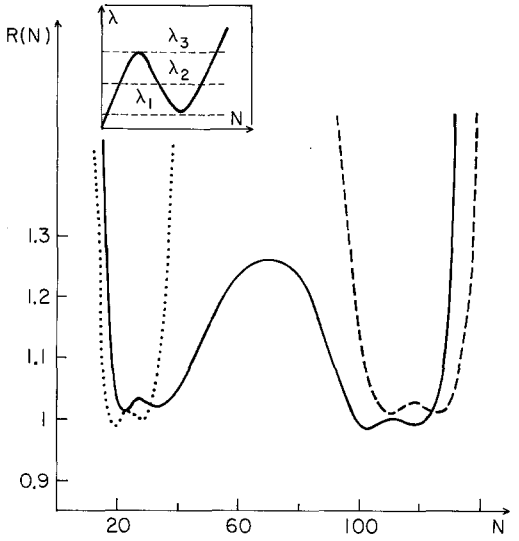

Fig. 1. Comparison (ratio) between the stationary probability $P_{st}(N)$ for the charge fluctuations in a bistable Esaki diode of the nonlinear Fokker-Planck approximation and the exact solution. The ratio R is shown for three different values of the pump parameter; λ_1: dotted, λ_2: solid, λ_3: dashed, indicated in the small figure for the nonlinear current-charge Esaki diode characteristic

sensitive to the peaks of maximal probability as $\langle N\rangle$, $\langle N^2\rangle$, etc. are of minor influence. How can this behaviour be understood on the basis of semi-group methods for the different Markoff processes? Besides the processes $\mathbf{x}_\varepsilon(t)$ we introduce the processes $\mathbf{x}_\varepsilon^{\mathrm{NFP}}(t)$ obtained by truncating off the Kramers-Moyal expansion (3.3) consistently in lowest order of Ω [11, 12, 31]:

$$\Gamma^{\mathrm{NFP}} = -\nabla\cdot[\mathbf{c}_1^0(\mathbf{x}_\varepsilon(t),t)+\varepsilon\,\mathbf{c}_1^1(\mathbf{x}_\varepsilon(t),t)] + \tfrac{1}{2}\varepsilon\nabla\nabla\colon\ \mathbf{c}_2^0(x_\varepsilon(t),t) \tag{3.11}$$

where for the following

$$\mathbf{c}_2^0(\mathbf{x},t)=\mathbf{B}_0(\mathbf{x},t)\,\mathbf{B}_0^+(\mathbf{x},t). \tag{3.12}$$

Before stating the main theorem of this section, we note that for a Markov process with the generator $\Gamma(t)$ and the transition function $R(\mathbf{x}r|\mathbf{y}t)$, $r>t$

$$\mathscr{T}(s)f(t)=\langle f(t+s)|\mathscr{F}_t\rangle=\langle f(\mathbf{x}(t+s))|\mathbf{x}(t)\rangle \qquad s>0$$
$$=\int f(\mathbf{u})\,R(\mathbf{u},t+s|\mathbf{x}(t),t)\,d\mathbf{u}, \tag{3.13}$$

where with a self-explaining notation

$$R(\mathbf{x},r|\mathbf{y},t)=\langle\mathbf{x}|\mathbf{y}\rangle+\int_t^r\langle\mathbf{x}|\Gamma(\tau)|\mathbf{y}\rangle\,d\tau + \tfrac{1}{2}\int_t^r\int_t^r\langle\mathbf{x}\,|\Gamma(\tau)\,\Gamma(\mu)|\mathbf{y}\rangle\,d\tau\,d\mu+\cdots. \tag{3.14}$$

A time ordering is implied in (3.14).

As a by-product from the theorem in Section 4 we obtain: A sequence of Markov processes $\mathbf{x}_\varepsilon(t)$ with the semi-group $\mathscr{T}_{x_\varepsilon}(s)$ converges in distribution to the Markov process $\mathbf{x}(t)$ with semi-group $\mathscr{T}_x(s)$ if $\lim_{\varepsilon\to 0} p_{x_\varepsilon}(\mathbf{u},0)=p_x(\mathbf{u},0)$ and

$$\lim_{\varepsilon\to 0}\langle|\mathscr{T}_{x_\varepsilon}(s)f(\mathbf{x}_\varepsilon(t))-\mathscr{T}_x(s)f(\mathbf{x}_\varepsilon(t))|\rangle=0,\ t+s<\infty \tag{3.15}$$

for all $t, s\geqq 0$ and every $f\in C_0$, the space of continuous functions vanishing at the natural boundaries. Moreover, under some mild mathematical restrictions (uniform convergence of the semi-groups [33]) the sequence of processes $\mathbf{x}_\varepsilon(t)$ converges *weakly* to the process $\mathbf{x}(t)$! If we consider the process $\mathbf{x}(t)$ where

$$\Gamma(t)=-\nabla\mathbf{c}_1^0(\mathbf{x}(t),t) \tag{3.16}$$

one can see that both the processes $\mathbf{x}_\varepsilon(t)$, Equation (3.3), and $x_\varepsilon^{\mathrm{NFP}}(t)$, Equation (3.11), converge in the above mentioned sense to the process $\mathbf{x}(t)$. This follows immediately from (3.15) by using (3.13–3.14). The term with c_1^0 cancels and all other terms are at least of order $O(\Omega^{-1})$ assuming that all the moments $c_n(\mathbf{x},t)$ are of order unity as usual and the derivatives of $p_{x\varepsilon}(t)$ with respect to the "intensive" variable $\mathbf{x}_\varepsilon$ are of order unity since always $\eta_\varepsilon(t)=\Omega^{-\frac{1}{2}}\zeta_\varepsilon(t)\lesssim O(1)$ [29d] (For the details of an exact proof procedure see the similar case in Section 4.) The solution of the master equation with the generator in (3.16) can be obtained in terms of the characteristic equations:

$$(d/dt)\,\mathbf{x}+\mathbf{c}_1^0(\mathbf{x}(t),t),\qquad \mathbf{x}(0)=\mathbf{x}_0 \tag{3.17}$$
$$(d/dt)\,P+P\nabla\cdot\mathbf{c}_1^0(\mathbf{x}(t),t)=0. \tag{3.18}$$

With the solution $\mathbf{x}(t)=\mathbf{x}(\mathbf{x}_0,t)$ of (3.17) we obtain

$$P(\mathbf{x},t)=P(\mathbf{x}_0,0)\left\|\frac{\partial\mathbf{x}(\mathbf{x}_0,t)}{\partial\mathbf{x}_0}\right\|_{\mathbf{x}_0},\qquad t<\infty$$
$$=P(\mathbf{x}_0,0)\exp-\int_0^t\nabla\cdot\mathbf{c}_1^0\mathbf{x}(s),s)\,ds. \tag{3.19}$$

Whence, both the processes $\{\mathbf{x}_\varepsilon(t)\}$ and $\{\mathbf{x}_\varepsilon^{\mathrm{NFP}}(t)\}$ converge weakly to the process determined by the deterministic motion where the fluctuation behaviour enters only by the randomness contained in the initial probability for $\mathbf{x}_0$. This proves on a *qualitative* basis the validity of the nonlinear Fokker-Planck equation.

A more detailed treatment is obtained by the study of the sequence of the processes $\bar{\mathbf{x}}_\varepsilon(t)$, describing the fluctuations around the regular part of $\mathbf{x}_\varepsilon(t)$, defined through the (Ito)-stochastic differential equation

$$\varepsilon^{\frac{1}{2}}d\bar{\mathbf{x}}_\varepsilon(t)=d\mathbf{x}_\varepsilon(t)-d\mathbf{y}(t) \tag{3.20}$$

with

$$\mathbf{y}(t)=\mathbf{x}_\varepsilon(0)+\int_0^t \{c_1^0(\mathbf{x}_\varepsilon,s)+\varepsilon c_1^1(\mathbf{x}_\varepsilon,s)\}\,ds. \tag{3.21}$$

The actual value of the random variable $\bar{\mathbf{x}}_\varepsilon(t)$ is determined by use of (3.21) in terms of the starting value $\mathbf{x}_\varepsilon(0)$, time t and the sample function characterized by a certain **label.** The master equation for $\mathbf{x}_\varepsilon(t)$ becomes then with (3.20–3.21) and (3.3)

$$\begin{aligned}&\frac{\partial P(\bar{\mathbf{x}}_\varepsilon,t)}{\partial t}\\&=-\varepsilon^{\frac{1}{2}}\frac{\partial}{\partial\bar{\mathbf{x}}_\varepsilon}[c_1(\mathbf{x}_\varepsilon,t)-c_1^0(\mathbf{x}_\varepsilon,t)-\varepsilon c_1^1(\mathbf{x}_\varepsilon,t)]P(\bar{\mathbf{x}}_\varepsilon,t),\\&+\sum_{n=2}^{\infty}\frac{(-1)^n}{n!}\varepsilon^{(n-2)/2}\left(\frac{\partial}{\partial\bar{\mathbf{x}}_\varepsilon}\right)^n c_n(\mathbf{x}_\varepsilon,t)\,P(\bar{\mathbf{x}}_\varepsilon,t).\end{aligned} \tag{3.22}$$

Using again in a straightforward way the above procedure, Equation (3.15), and the fact that for $\mathbf{x}_\varepsilon(0)\to\mathbf{z}_0$, in (3.21) $\mathbf{x}_\varepsilon(t)$ converges to $\mathbf{x}(z_0,t)$ for $\varepsilon\to 0$ ((3.19) and [27a]), we obtain that $\{\bar{\mathbf{x}}_\varepsilon(t)\}$ converges (weakly) to the Markov process $\bar{\mathbf{x}}(t)$ with the generator (see also theorem 3.1 of [27a])

$$\Gamma_{\bar{x}}=\tfrac{1}{2}\,\mathbf{c}_2^0(\mathbf{x}(\mathbf{z}_0,t)\,t)\colon\nabla_{\bar{x}}\nabla_{\bar{x}}. \tag{3.23}$$

Hence, $\mathbf{x}_\varepsilon(t)$ is asymptotically well approximated by the non-linear Fokker-Planck process $\mathbf{x}_\varepsilon^{\mathrm{NFP}}(t)$, since with (3.12):

$$\begin{aligned}&d\bar{\mathbf{x}}_\varepsilon^{\mathrm{NFP}}(t)=\varepsilon^{-\frac{1}{2}}[d\mathbf{x}_\varepsilon^{\mathrm{NFP}}(t)-c_1^0(\mathbf{x}_\varepsilon^{\mathrm{NFP}},t)\,dt-\varepsilon c_1^1(\mathbf{x}_\varepsilon^{\mathrm{NFP}},t)\,dt]\\&=\mathbf{B}_0(\mathbf{x}_\varepsilon^{\mathrm{NFP}},t)\,d\mathbf{w}(t).\end{aligned} \tag{3.24}$$

This all generalizes for multivariable systems the result of Kurtz, (Theorem 3.1 in 2nd Ref. of [27a]), to the time-inhomogeneous case and general scaling law (3.1) using a more elegant proof procedure.

But nothing is said about a *quantitative* difference in the probabilities of the processes $\mathbf{x}_\varepsilon(t)$ *and* $\mathbf{x}_\varepsilon^{\mathrm{NFP}}(t)$ for a *finite* value of the parameter Ω. Error estimates can be obtained in principle by use of (3.15) and setting $f(\mathbf{x}_\varepsilon)=\exp i\omega\,\mathbf{x}_\varepsilon(t)$ yielding the characteristic function $\langle\exp i\omega\,\mathbf{x}_\varepsilon(t+s)|\mathbf{x}_\varepsilon(t)\rangle$ for the conditional probability and investigating the terms of $O(\Omega^{-1})$ in more detail. As one can see from Figure 1 a kind of a L_1 or L_2-estimate for the difference in $P_{x_\varepsilon}(t)$ and $P_{x_\varepsilon^{\mathrm{NFP}}}(t)$, e.g.

$$L_1\colon\int|P_{x_\varepsilon}(\mathbf{u},t)-P_{x_\varepsilon^{\mathrm{NFP}}}(\mathbf{u},t)|\,d\mathbf{u}\approx\mathbf{0}. \tag{3.25}$$

yields about zero, since the regions of deviations, where $P(\mathbf{x},t)\ll P(\mathbf{u},t)$, ($u$: at main maxima), have nearly a vanishing weight.

Finally we mention that with the validity of the Van Kampen procedure (no anomalous fluctuations) both master equations for the processes $\mathbf{x}_\varepsilon^{\mathrm{NFP}}(t)$ and $\zeta(t)$ can be used. The sequence of processes

$$\Omega^{-\frac{1}{2}}\zeta_\varepsilon^{\mathrm{NFP}}(t)=\mathbf{x}_\varepsilon^{\mathrm{NFP}}(t)-\mathbf{x}(\mathbf{x}_0,t) \tag{3.25}$$

converges also *weakly* to the Gauss-Markov process given by (3.7). The difference in use of either (3.11) or (3.7) is that with use of (3.7) the limit, $\Omega\to\infty$, *has been performed* whereas with use of the master equation in (3.11), or in terms of the transformed process $\zeta_\varepsilon^{\mathrm{NFP}}(t)$:

$$\begin{aligned}&\frac{\partial P}{\partial t}(\zeta_\varepsilon^{\mathrm{NFP}},t)\\&=-\varepsilon^{-\frac{1}{2}}\frac{\partial}{\partial\zeta_\varepsilon^{\mathrm{NFP}}}[c_1^0(\mathbf{x}(t)+\varepsilon^{\frac{1}{2}}\zeta_\varepsilon^{\mathrm{NFP}}(t),t)\\&+\varepsilon c_1^1(\mathbf{x}(t)+\varepsilon^{\frac{1}{2}}\zeta_\varepsilon^{\mathrm{NFP}}(t),t)-c_1^0(\mathbf{x}(t),t)]\,P(\zeta_\varepsilon^{\mathrm{NFP}},t)\\&+\tfrac{1}{2}\frac{\partial^2}{(\partial\zeta_\varepsilon^{\mathrm{NFP}})^2}\,\mathbf{c}_2^0(\mathbf{x}(t)+\varepsilon^{\frac{1}{2}}\zeta_\varepsilon^{\mathrm{NFP}}(t)),t)\,P(\zeta_\varepsilon^{\mathrm{NFP}},t)\end{aligned} \tag{3.27}$$

still some Ω-dependence is retained which converges to zero as $\Omega\to\infty$.

IV. Limit Theorem for Non-Markov Processes – a Model for Macroscopic Irreversibility

Any macroscopic stochastic process with physical relevance has its origin in a microscopic reversible dynamics of a total closed system. On the microscopic level, the dynamics is described in terms of an unitary group of time-shifts. A first step towards a macroscopic irreversible dynamics consists in the replacement of the microscopic phase function $\rho(\mathbf{q},\mathbf{p};t)$ by some "coarse grained" probability $\rho^c(\mathbf{x}^c,t)$ by use of a projection into some subspace [5, 15]. Such a "reduction" destroys usually the semi-group property governing the propagation of the probability ρ. Moreover, the process $\mathbf{x}^c(t)$, which still contains in hidden form (memory effect) the whole microscopic information as long as the projectin procedure is performed without any approximations, is usually non-Markovian [15]. The process $\mathbf{x}^c(t)$ can be considered as a kind of a microscopic dynamics including irreversibility represented by the linear semigroup of conditional shifts (Eq. (2.1)). In a final step, macroscopic irreversibility on a restricted time scale, characterized for example by the linear Markovian semi-group for a Fokker-Planck-process, can be obtained with convergence of a sequence of non-Markovian processes $\mathbf{x}_\varepsilon^c(t)$, (we assume an "intensive" character for $\mathbf{x}_\varepsilon^c(t)$), to the Fokker-Planck process $\mathbf{x}(t)$. In summary, this approach seems to provide the "missing link" between microscopic reversible dynamics and macroscopic irreversible dynamics. From a mathematical point of view, convergence of a sequence of non-Mar-

kov processes to a Markov process has been first considered in the pioneering work of Gihman [34] and Borokov [35]. Very recently, the results have been formulated in the language of semi-groups by Kurtz [33]. From his work we will use in the following a theorem which reads in a for physicists comprehensible form: Let $\{\mathbf{x}_\varepsilon(t)\}$ be a sequence of non-Markov processes and $\lim_{\varepsilon\to 0} p_{x_\varepsilon}(\mathbf{u},0)=p^x(\mathbf{u},0)$, where $\mathbf{x}(t)$ denotes the limiting Markov process with semi-group $\mathscr{T}_x$. Suppose

$$\lim_{\varepsilon\to 0}\langle|\langle f(\mathbf{x}_\varepsilon(t+s))|\mathscr{F}_t\rangle-\mathscr{T}_x(s)f(\mathbf{x}_\varepsilon(t))|\rangle=0,\qquad t+s<\infty \tag{4.1}$$

for all $t, s\geqq 0$ and every $f\in C_0$. Then the finite dimensional probabilities $p^{(n)}_{x_\varepsilon}$ of $\mathbf{x}_\varepsilon(t)$ converge to the finite dimensional probabilities $p^{(n)}_x$ of the Markov process $\mathbf{x}(t)$ with the semi-group $\mathscr{T}_x(s)$. (See Eqs. (2.7–2.8).) The special model we have in mind is the following: The characteristics of the process $\mathbf{x}_\varepsilon(t)$ is the existence of a set of states in the state space which on each visit knock out some amount of non-Markov behaviour (memory). With increasing parameter value Ω, ($\varepsilon(\Omega)\to 0$ for $\Omega\to\Omega_0$) these states are hit more and more faster so that the remaining history decays on a characteristic time τ_ε with $\tau_\varepsilon\to 0$ in the limit $\varepsilon\to 0$. In particular, we suppose for the conditional moments of $\Delta\mathbf{x}_\varepsilon(t)=\mathbf{x}_\varepsilon(t+\tau_\varepsilon)-\mathbf{x}_\varepsilon(t)$, $\tau_\varepsilon>0$*

$$\langle\Delta\mathbf{x}_\varepsilon(t)|\mathscr{F}_t\rangle=\tau_\varepsilon\,\mathbf{a}(\mathbf{x}_\varepsilon(t))+e^1_\varepsilon(t) \tag{4.2}$$

$$\langle(\Delta\mathbf{x}_\varepsilon(t))^2|\mathscr{F}_t\rangle=\tau_\varepsilon\,\mathbf{D}(\mathbf{x}_\varepsilon(t))+e^2_\varepsilon(t) \tag{4.3}$$

$$\langle|\Delta\mathbf{x}_\varepsilon(t)|^3|\mathscr{F}_t\rangle=e^3_\varepsilon(t). \tag{4.4}$$

Here the "error"-terms $\{e^1_\varepsilon, e^2_\varepsilon, e^3_\varepsilon\}$ are of the *order* $o(\tau_\varepsilon)$ *as* $\varepsilon\to 0$. For the following it is convenient to introduce the corresponding non-Markov chain approximation for $\mathbf{x}_\varepsilon(t)$ (discretization of the time-parameter t) in the sense that

$$\mathbf{x}_\varepsilon(t)=\mathbf{x}^\varepsilon_n \tag{4.5}$$

with

$$n=\text{Integer}\,[t/\tau_\varepsilon]. \tag{4.6}$$

Equations (4.2–4.4) can simply be rewritten; e.g. we have for (4.2)

$$\langle\Delta x^\varepsilon_n|\mathscr{F}_n\rangle=\tau_\varepsilon\,\mathbf{a}(\mathbf{x}^\varepsilon_n)+e^1_\varepsilon(n) \tag{4.7}$$

with $e^i_\varepsilon(n)$, $i=1,2,3$, in the form

$$\lim_{\varepsilon\to 0}\sum_{n<[t/\tau_\varepsilon]}\langle|e^i_\varepsilon(n)|\rangle=0. \tag{4.8}$$

* Note that the relations (4.2–4.4) hold for any sequence of non-Markovian processes as long as the "error"-terms ae not specified more. In the random variables, $e^i_\varepsilon(t)$, the dependence on the condition, $\mathscr{F}_t$, is suppressed

In this section we end up, by use of the above theorem showing that the *multivariate* probabilities $p^{(n)}_{x_\varepsilon}$ converge to those of the Fokker-Planck process, $p^{(n)}_x$, characterized by the condition [13, 14, 17]:

$$\int(\mathbf{y}-\mathbf{x})\,R(\mathbf{x}\to\mathbf{y};\tau)\,d\mathbf{y}=\tau\,\mathbf{a}(\mathbf{x})+e_1(\tau) \tag{4.9}$$

$$\int(\mathbf{y}-\mathbf{x})^2\,R(\mathbf{x}\to\mathbf{y};\tau)\,d\mathbf{y}=\tau\,\mathbf{D}(x)+e_2(\tau) \tag{4.10}$$

$$\int|\mathbf{y}-\mathbf{x}|^3\,R(\mathbf{x}\to\mathbf{y};\tau)\,d\mathbf{y}=e_3(\tau) \tag{4.11}$$

where $e_i(\tau)$, $i=1,2,3$, is of the order $o(\tau)$.
In a first step we show that for $f\in C_0$, $n>m$:

$$\langle|\langle f(\mathbf{x}_n)|\mathscr{F}_m\rangle-\mathscr{T}_x([n-m]\,\tau)f(\mathbf{x}_m)|\rangle$$
$$\leqq\sum_{i=m}^{n-1}\langle|\langle h_{i+1}(\mathbf{x}_{i+1})|\mathscr{F}_i\rangle-\mathscr{T}_x(\tau)\,h_{i+1}(\mathbf{x}_i)|\rangle, \tag{4.12}$$

where $h_i=\mathscr{T}_x([n-i]\,\tau)f$ (4.13)

and ε has been suppressed. By use of

$$f(\mathbf{x}_n)-\mathscr{T}_x([n-m]\,\tau)f(\mathbf{x}_m)=h_n(\mathbf{x}_n)-h_m(\mathbf{x}_m)$$
$$=\sum_{i=m}^{n-1}(h_{i+1}(\mathbf{x}_{i+1})-h_i(\mathbf{x}_i)) \tag{4.14}$$

we obtain with Equations (2.4–2.5):

$$\langle f(\mathbf{x}_n)|\mathscr{F}_m\rangle-\mathscr{T}_x([n-m]\,\tau)f(\mathbf{x}_m)$$
$$=\sum_{i=m}^{n-1}\langle h_{i+1}(\mathbf{x}_{i+1})-h_i(\mathbf{x}_i)|\mathscr{F}_m\rangle$$
$$=\sum_{i=m}^{n-1}\langle\langle h_{i+1}(\mathbf{x}_{i+1})|\mathscr{F}_i\rangle-h_i(\mathbf{x}_i)|\mathscr{F}_m\rangle. \tag{4.15}$$

Using $h_i=\mathscr{T}_x(\tau)\,h_{i+1}$ and taking absolute values and expectations on both sides of (4.15) we obtain the result in (4.12). Next we need an estimation for the terms on the right side of (4.12). First we note, that the Kolmogorov-backward equation for the time-homogeneous Fokker-Planck process gives the relation [14, 17]:

$$\frac{\partial}{\partial t}\langle f(\mathbf{x}(t))|\mathbf{y}\rangle=\Gamma^+ f(\mathbf{y}) \tag{4.16}$$

$$\Gamma^+ f(\mathbf{y})=\mathbf{a}(\mathbf{y})\nabla f(\mathbf{y})+\tfrac{1}{2}\mathbf{D}(\mathbf{y}):\nabla\nabla f(\mathbf{y}). \tag{4.17}$$

Using the Taylor series for $g\in C^2_0$:

$$g(\mathbf{y})=g(\mathbf{x})+(\mathbf{y}-\mathbf{x})\nabla g(\mathbf{x})+\tfrac{1}{2}(\mathbf{y}-\mathbf{x})^2:$$
$$\nabla\nabla g(\mathbf{x})+c|\mathbf{y}-\mathbf{x}|^3\,|g''|^L, \tag{4.18}$$

where with

$$g(\mathbf{y})-g(\mathbf{x})-(\mathbf{y}-\mathbf{x})\nabla g(x)$$
$$=(\mathbf{y}-\mathbf{x})^2:\int_0^1(1-\theta)\nabla\nabla g(\mathbf{x}+\theta(\mathbf{y}-\mathbf{x}))\,d\theta$$
$$=\tfrac{1}{2}(\mathbf{y}-\mathbf{x})^2:\nabla\nabla g(\mathbf{x})+(\mathbf{y}-\mathbf{x})^2:$$
$$\int_0^1(1-\theta)\,[\nabla\nabla g(\mathbf{x}+\theta(\mathbf{y}-\mathbf{x}))-\nabla\nabla g(\mathbf{x})]\,d\theta. \tag{4.19}$$

$|c|<1$, and $|g''|^L$ denotes the Lipschitz norm: $\sup_{x \neq y} |\nabla\nabla g(\mathbf{y}) - \nabla\nabla g(\mathbf{x})|/|\mathbf{y}-\mathbf{x}| < \infty$. For the following let

$$\|g\| = \sup|\nabla g| + \sup|\nabla\nabla g| + \sup|\nabla\nabla g(\mathbf{y}) - \nabla\nabla g(\mathbf{x})|/|\mathbf{y}-\mathbf{x}|. \tag{4.20}$$

Hence in view of Equations (4.16–4.17) we have

$$\mathscr{T}_x(\tau)\, g(\mathbf{x}_i) = g(\mathbf{x}_i) + \tau \Gamma^+ g(\mathbf{x}_i) + c\,\|g\| \sum_{j=1}^{3} |e_j(\tau)| \tag{4.21}$$

and from (4.2–4.4)

$$\langle g(\mathbf{x}_{i+1}) | \mathscr{F}_i \rangle = g(\mathbf{x}_i) + \tau \Gamma^+ g(\mathbf{x}_i) + c\,\|g\| \sum_{j=1}^{3} |e^j(i)|. \tag{4.22}$$

Consequently, with (4.21–4.22) we obtain

$$\langle |\langle h_{i+1}(x_{i+1}) | \mathscr{F}_i \rangle - \mathscr{T}_x(\tau)\, h_{i+1}(\mathbf{x}_i)| \rangle \leqq \|g\| \sum_{j=1}^{3} [\langle |e_j(\tau)| \rangle + \langle |e^j(i)| \rangle]. \tag{4.23}$$

Using $n = [(t+s)/\tau_\varepsilon]$ and $m = [t/\tau_\varepsilon]$ and some constant S we can write

$$\langle |\langle f(\mathbf{x}_n^\varepsilon) | \mathscr{F}_m \rangle - \mathscr{T}_x([n-m]\,\tau) f(\mathbf{x}_m^\varepsilon)| \rangle \leqq S \sum_{j=1}^{3} \left\{ (n-m)\langle |e_j(\tau)| \rangle + \sum_{i=m}^{n-1} \langle |e_\varepsilon^j(i)| \rangle \right\}. \tag{4.24}$$

But $(n-m)\cdot\tau \to s$ as $\varepsilon \to 0$ and C_0^2 dense in C_0 so that with the properties for $e_j(\tau)$ and $e_\varepsilon^j(i)$:

$$\lim_{\varepsilon \to 0} \langle |\langle f(\mathbf{x}_\varepsilon(t+s)) | \mathscr{F}_t \rangle - \mathscr{T}_x(s) f(\mathbf{x}_\varepsilon(\mathbf{t}))| \rangle = 0. \tag{4.25}$$

Hence, after some inevitable mathematics, we have shown that the multivariate probabilities $p_{x_\varepsilon}^{(n)}$ converge indeed to the multivariate probabilities $p_x^{(n)}$ of the Fokker-Planck process. Note that this does *not* imply necessarily *weak* convergence of $\mathbf{x}_\varepsilon(t)$ to $\mathbf{x}(t)$. – An approximative validity of this model justifies the broad use of (Markov)-Fokker-Planck processes for the description of macroscopic physical systems [1, 2, 4, 5, 11, 12, 29–32].

The method used in this section can also be applied without essential modifications to the investigation of the convergence to a Fokker-Planck process which is inhomogeneous in time.

V. Discussion

In this paper we presented various results for Markov processes and non-Markov processes by using the apparatus of semi-groups. The elegance of the method has been clearly demonstrated in the derivation of a master equation for a Markov process which undergoes both continuous and discontinuous jumps. This method is well based on mathematical grounds whereas the alternative approach based on generalized Langevin equations is subject to some pitfalls. First the properties for the stochastic forces (e.g., their cumulant averages) may not be always consistent with a descritpion for a Markov process. In addition one needs the information between the correlations of the macroscopic variables $\mathbf{x}(t)$ and the stochastic forces, a task which is normally not easy to untangle. The concept of nonlinear transformation of stochastic processes yields a broader class of master equations whose solutions can be written down explicitly.

The results in Section III elucidate the validity and the relationships among different asymptotic representations of Markov processes. Further, we stress that for an initially time-homogeneous process $\mathbf{X}(t)$ the occurrence of in general time-inhomogeneous drift and diffusion coefficients in the master equation for the fluctuations around the time-dependent deterministic solution, Equation (3.7), yielding a two-parameter propagator description, does not imply a non-Markofficity for $\mathbf{x}_\varepsilon(t)$ as it is the underlying philosophy in a series of papers about chemical reactions by Keizer [36]. The result that the sequence of processes converges weakly to the corresponding asymptotic representation justifies the broad use of the asymptotic probabilities for the calculation of stochastic quantities. In a very recent letter [27] Van Kampen himself discusses the validity of the nonlinear Fokker-Planck approximation. He claims that the nonlinear Fokker-Planck equation can be used consistently only in the case where $c_1^0(\mathbf{x}, t)$ (see (3.4)) is zero and "that the oft quoted Kolmogomov derivation applies to this case *alone* because it is based on the Lindeberg condition and thereby precludes the possibility that $c_1^0(\mathbf{x}, t)$ differs from zero." Such statements are just not true because the Lindeberg condition guarantees for a Markov process continuous sample functions and whence the validity of any general nonlinear Fokker-Planck equation independent of any special scaling assumptions for the transition probabilities. Furthermore, his claim about the inconsistent use of the nonlinear Fokker-Planck equation beyond the Gaussian approximation is not justified. (See end of Section III and the situation with multiple stationary states.) Some people justify the incosistence of a nonlinear Fokker-Planck equation beyond the Gauss-Markov approximation with the impossibility to take both p_{st} and the stochastic forces in a Langevin equation Gaussian for a nonlinear system. This is true as long as one

restricts the nonlinear Fokker-Planck equations to those with a constant diffusion coefficient. With the possibility of a nonlinear diffusion coefficient the argument breaks down as can be seen easily from simple examples and again Section III.

The model in Section IV provides an understanding between the concepts of microscopic reversible dynamics and macroscopic irreversibility. In this context there are other topics related to the very definition of unstable particles; a problem which attracts presently a great deal of interest. The concept of the characteristic overall memory – decay time τ_e is closely related to the relaxation time of the memory kernels occuring in the exact master equation for the single event probability [5, 15]. But the relaxation time τ_e has much more restrictive properties. It also assures a short time decay for "higher"-memory kernels as they occur in the exact master equations for multivariate probabilities [15]. Unfortunately, the approach for macroscopic irreversibility lacks uniqueness in the sense that no unambiguous prescription for the operation "coarse graining" can be formulated. But on the other hand the construction of the (ambiguous) irreversible microprocess $\mathbf{x}_\varepsilon^c(t)$ is directed by a certain physical insight. A different point of view is taken in an interesting paper by Prigogine et al. [38]. They introduce on a microscopic level a quadratic Liapunov functional Ω with a non-increasing time derivative. However, the ambiguity enters here through the fact that for a given reversible microdynamics the construction of such a Liapunov functional is certainly not unique.

References

1. Haken, H.: Rev. Mod. Phys. **47**, 67 (1975)
2. Graham, R.: in Fluctuations, Instabilities and Phase Transitions (ed. Riste) New York: Plenum 1976
3. Normand, C., Pomeau, Y., Velarde, M.G.: Rev. Mod. Phys. **49**, 581 (1977)
4. Graham, R.: Springer Tracts in Modern Physics. **66**, 1 (1973)
5. Haake, F.: Springer Tracts in Modern Physics **66**, 98 (1973)
6. Goel, N.W., Richter-Dyn, N.: Stochastic Models in Biology, New York: Academic Press 1974
7. Görtz, R., Walls, D.F.: Z. Physik B**25**, 423 (1976)
8. Haag, G., Weidlich, W., Alber, P.: Z. Physik B**26**, 207 (1976)
9. Van Kampen, N.G.: Adv. Chem. Phys. **34**, 245 (1976)
10. Kubo, R., Matsuo, K., Kitahara, K.: J. Stat. Phys. **9**, 51 (1973)
11. Horsthemke, W., Brenig, L.: Z. Physik B**27**, 341 (1977)
12. Horsthemke, W., Malek-Mansour, M., Brenig, L.: Z. Physik B**28**, 135 (1977)
13. Gihman, I.I., Skorohod, A.V.: The Theory of Stochastic Processes, Vol. II. Berlin-New York: Springer 1975
14. Arnold, L.: Stochastic Differential Equations, New York: John Wiley 1974
15. Grabert, H., Talkner, P., Hänggi, P., Thomas, H.: Z. Physik B**29**, 273 (1978)
16. Hänggi, P., Thomas, H.: Z. Physik B**26**, 85 (1977)
17. Hänggi, P.: Helv. Phys. Acta: **51**, 2 (1978)
18. Sture, K., Nordholm, K.S., Zwanzig, R.: J. Stat. Phys. **11**, 143 (1974)
19. Hänggi, P., Rösel, F., Trautmann, D.: Z. Naturf.: in press
20. Kramers, H.A.: Physica **7**, 284 (1940)
21. Horsthemke, W., Bach, A.: Z. Physik B**22**, 189 (1975)
22. Graham, R.: Z. Physik B**26**, 281 (1977)
23. Hänggi, P., Thomas, H.: Z. Physik B**22**, 295 (1975)
24. Cuiker, R.I., Lakatos-Lindenberg, K., Shuler, K.E.: J. Stat. Phys. **9**, 137 (1973)
25. Gihman, I.I., Skorohod, A.V.: Introduction to the theory of random processes, Chapt. 9, Philadelphia: W.B. Saunders Co. 1969 (translation from Russian)
26. Stone, Ch.: Illinois J. Math. **7**, 638 (1963)
27. a) Kurtz, T.G.: J. Chem. Phys. **57**, 2976 (1972); J. Appl. Probability **8**, 344 (1971)
    b) Norman, F.M.: British J. Math. Stat. Psychology **21**, 141 (1968); Ann. Probability **2**, 1065 (1974)
28. Liggett, T.M.: Indiana Univ. Math. J. **20**, 265 (1970)
29. a) Janssen, H.K.: Z. Physik **270**, 67 (1974)
    b) Saito, Y.: J. Phys. Soc. Japan **41**, 388 (1976)
    c) Matsuo, K.: J. Stat. Phys. **16**, 169 (1977)
    d) Suzuki, M.: J. Stat. Phys. **16**, 477 (1977)
    e) Oppenheim, I.I., Shuler, K.E., Weiss, G.H.: Physica **88**A, 191 (1977)
    f) Nicolis, G., Lefever, R.: Physics Letters **62**A, 469 (1977)
    g) Procaccia, I., Ross, J.: J. Chem. Phys. **67**, 5565 (1977)
30. Matheson, I., Walls, D.F., Gardiner, C.W.: J. Stat. Phys. **12**, 21 (1975)
31. Hänggi, P.: Dynamical Behaviour for a Stochastic Model of a Bistable Tunnel-Diode; Thesis, (1977)
32. Landauer, R.: J. Appl. Phys. **33**, 2209 (1962)
33. Kurtz, T.G.: Ann. Probability **3**, 618 (1975)
34. Gihman, I.I.: Ukrainan Math. J. **21**, 263 (1969)
35. Borovkov, A.A.: Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **16**, 47 (1970)
36. a) Keizer, J.: J. Math. Phys. **18**, 1317 (1977)
    b) Keizer, J.: J. Chem. Phys. **63**, 5037 (1975)
    c) Keizer, J.: J. Chem. Phys. **63**, 398 (1975)
37. Van Kampen, N.G.: Phys. Lett. **62**A, 383 (1977)
38. Prigogine, I., Mayne, F., George, C., DeHaan, M.: Proc. Natl. Acad. Sci. **74**, 4152 (1977)

Peter Hänggi
Physics Department
Loomis Laboratory of Physics
Urbana, IL 61801 U.S.A.