Journal of Statistical Physics, Vol. 25, No. 1, 1981

Some Properties of Stochastic Equations for Coupled Chemical Reactions Far from Equilibrium

C. Blomberg[1]

Received May 23, 1978; revised May 29, 1979 and November 15, 1979

Solutions of master equations for coupled chemical reactions far from equilibrium with one varying molecule species are studied and used for getting information about nonlinear Fokker–Planck equations and slow time-dependent processes such as extinction to an absorbing state and transition between several steady states. The Fokker–Planck equation solution is compared to that of the master equation in a relative sense and it is shown that they agree quite well in some important situations but that in general the cases can deviate considerably, when, e.g., accounting for the mutual importance of two probability maxima.

KEY WORDS: Fluctuations; Fokker–Planck equations; transition rates; nonequilibrium thermodynamics.

1. INTRODUCTION

Open systems, driven far from equilibrium by actual or model chemical reactions, are much studied in recent times with the particular aim to understand new kinds of (dissipative) structures without any correspondence in equilibrium systems. Basically, these are described by chemical rate equations for average concentrations, i.e., deterministic equations. In the last years, a number of works$^{(1-10)}$ have considered general stochastic equations for such systems in order to get a more complete picture of the behavior. In this work, we shall take up some further questions in that context.

In principle, as in ordinary statistical mechanics, molecule concentrations are macroscopic variables and for these, fluctuations will occur

---

[1] Department of Theoretical Physics, Royal Institute of Technology, S-10044 Stockholm, Sweden.

because of irregularities in molecule encounters. The probability functions describing these are, however, very sharply centered around the deterministic (most probable) values of the concentrations, and are virtually zero outside an extremely small region close to the deterministic values. This feature is considered by Oppenheim *et al.*(2) and more rigorous statements are given by Kurtz.(11) Ventsel and Freidlin(12) have considered the mathematical structure for the probability far away from the most probable values. From their work, it is clear that the logarithm of the probability is an extensive variable, a fact that also has been explored in other of the cited works.

In the simplest case, there will occur one unique stationary stable solution of the deterministic equation which all trajectories approach in the course of time. There, all important physical features are given by such solutions and local fluctuations close to the deterministic values. The latter are most easily provided by the "linear" Fokker–Planck equation,(13) and, perhaps, most conveniently studied by making some transformation of the probability function. For instance, the Poisson representation by Gardiner and Chaturvedi(10) seems to be very powerful for providing general moments and correlations.

The situation is, however, different in a more complex case with additional structures where a detailed description of the global fluctuations is necessary in order to account for the mutual relevance of these structures. Moreover, in such cases it is quite possible that the deterministic and stochastic descriptions give rise to conceptually different interpretations.

This is particular the case when there are many stationary solutions of the deterministic equation. This situation is studied in several works(2–4, 14) and one usually has several points that are stable against local changes. In a stochastic picture, generally described by a Markoff process, there will be a unique stationary probability function where such points occur as local maxima. Of these, one will dominate (except for certain phase coexistence situations) and, for a large system, together with a narrow fluctuation region, it will contain essentially all the probability. For any finite system, such a largest maximum is always reached after some time, which, however, can be extremely long. In order to decide which of several locally stable points corresponds to the highest probability maximum, it is necessary to consider the relative change of the probability function along a path joining the corresponding maxima.

For this reason, global fluctuations become important and the most relevant quantity of study should be the logarithm of the probability, which is an extensive entity. This corresponds to a free energy (or a similar type of function) in equilibrium statistical mechanics, and the point of largest probability corresponds, of course, to the state of lowest free energy. For treating problems of this type, it is important to find methods for getting the probability function and its logarithm also for large fluctuations, and in that context one should compare different means of description. These are among the basic aims of this paper. As a basic starting point, we adopt here a common view, and use master equations for describing the changes of molecule numbers. It seems reasonable that the basic description uses a discrete representation of the molecule numbers but we shall not go further into questions of which equation should be the most relevant one as a starting point. For simplicity, we will be mainly concerned with one varying molecule species. For this, we consider in some detail a type of method that seems originally to have been introduced by Kubo *et al.*(5) and later formulated for this kind of equation by Görtz and Walls.(6)

As said above, we shall also consider alternative equations, in particular, the continuous representation given by the so-called nonlinear Fokker–Planck equation. This equation has a large potential interest since it may provide means for obtaining solutions and approximations also in fairly complex situations. However, van Kampen(13) has shown that this equation in a strict sense is a consistent approximation of the master equation only when describing small fluctuations around deterministic solutions. In spite of this, Horsthemke and Brenig,(7) by using arguments by Kurtz, claim that the nonlinear Fokker–Planck equation is a correct asymptotic representation of the master equation. It seems that this controversy is caused by different viewpoints, and different meanings of the concept of asymptotic representation. It seems that the Horsthemke–Brenig assertion is true for an absolute comparison: the absolute difference between the solution of the nonlinear Fokker–Planck equation and that of the master equation is negligible for a large system provided they get the most probable point at the same value. Of course, as stated above, these solutions are both close to zero outside a local fluctuation region, and thus the latter should be the only region of relevance for the statement. This is also clear from the fact that the Kurtz theorems do only consider the descent of the probability in the fluctuation region. The situation should then be satisfactory for the simple case discussed above of one single stationary point, but a case of several locally stable states is not correctly described, as also stated in another paper by Horsthemke *et al.*(4)

An agreement in an absolute sense as that of Horsthemke and Brenig does not imply that the solutions are close in a relative sense, which would mean that their quotients would be close to 1 or that the logarithms are close to each other. In that sense, the van Kampen arguments are valid and the nonlinear Fokker–Planck equation is not in general a consistent approximation to the master equation. As stated above, the global relative variation is important for the mutual relevance of several structures in the probability. Thus, even if the nonlinear Fokker–Planck equation in a certain absolute sense is an asymptotic approximation, it cannot necessarily be used to judge which of several possible, locally stable points would correspond to the dominating maximum of the stationary solution of the master equation.

We therefore put the main emphasis here upon the relative variations, which are studied through the logarithm of the probability function in general cases, and we will consider the question in what extent the nonlinear Fokker–Planck equation is a good approximation to the master equation in a relative sense. As will be demonstrated by explicit examples, the agreement can in general cases be very bad, but for a large group of important problems, the solutions agree very well in a wide parameter range. In particular, this is so for most proposed models of chemical reactions with one varying molecule species.

In cases like these, transformations of the probability function do not provide an appropriate approach since the logarithm of the probability is not directly available. The method by Kubo *et al.*(5) and Görtz and Walls(6) is more suitable, and also the Fokker–Planck equation yields direct, simple expressions for the logarithm of the probability. (Note that the Fokker–Planck equation of Gardiner and Chaturvedi(10) is an equation for the transformed probability and has quite another meaning than the one here.)

Besides studying the stationary features, we shall also discuss some time aspects. In the true stationary probability distribution, that always is reached after infinite times, one stable point survives and others are suppressed. However, the local stability of other points will mean that a probability distribution around such a point can be kept for long times as it is unstable only because of unprobable, large fluctuations. Such a probability corresponds to a metastable, decaying state, which has been studied in some works(2, 15) as well as in the general mathematical framework by Ventsel and Freidlin.(12) We also consider such cases here, and try to give more precise methods to describe the decaying times. The latter can also be regarded as transition times in a case of several probability maxima. We also study the situation of an absorbing barrier, where the only stationary state is the completely absorbed situation, but where there may be a well-defined long-lived metastable state. In certain cases, the decaying times are enormously long, which means that they can be considered as truly stable states. Still, these times can be quite relevant for moderately large systems, for instance for describing chemical reactions in living cells where the number of reacting molecules may be of the order 10^3–10^4 or even less.

2. THE BASIC MASTER EQUATION AND ITS STATIONARY SOLUTION

The basic starting point for our development will be a master equation of birth and death type for the probability function of the number of molecules of one varying species:

$$\frac{\partial P(n,t)}{\partial t} = \sum_{r=-R}^{R} \left[W_r(n+r)P(n+r,t) - W_r(n)P(n,t) \right] \tag{1}$$

$P(n,t)$ is the probability function of having n molecules at time t. The $W_r(n)$ are the probability intensities for changing the molecule number n by $-r$ in a single reaction step. r can be positive or negative, but $r=0$ is always excluded. R is the maximum change of molecule number, and is a small number (<10). If $R=1$ (i.e., the molecule numbers just change by one unit in each reaction step), we have a *one-step process*.

It is essential here that we have a large parameter, V, related to the total number of molecules, or the extensions of the system. In the limit of infinite V, the variable $x = n/V$ and the functions $w_r(x) = W_r(n)/V$ remain finite. We further assume that, for each value of r, the $w_r(x)$ in the limit turn to differentiable functions of x (e.g., polynomials).

The number n shall always be positive or zero, meaning that there will be a first equation of (1) for $n=0$, and that certain terms in the first equations are lacking. This is achieved if we put $W_r(n)$ equal to zero if $n<0$ or $n<r$ (the latter condition means that the molecule number can never decrease by more than the number of molecules in the system). The first R equations with this convention can be regarded as a boundary condition.

We shall not consider here an upper limit of n. In principle, such a limit exists, but in all cases considered $P(n)$ will go rapidly to zero for large n, and anyhow be virtually zero at a natural upper limit.

We get another type of equation if the first n equations (1) are summed with the conventions for the lowest orders mentioned above:

$$\sum_{m=0}^{n} \frac{\partial P(m,t)}{\partial t} = \sum_{r=1}^{R} \left\{ \left[\sum_{p=r}^{R} W_p(n+r) \right] P(n+r,t) - \left[\sum_{p=r}^{R} W_{-p}(n-r+1) \right] P(n-r+1,t) \right\} \tag{2}$$

In this case, the first $(R-1)$ equations will lack certain terms as discussed above.

This type of master equation is of the same form as used in the references mentioned in the Introduction. The asymptotic character of the stochastic process is the same as treated by rigorous mathematical methods by Kurtz(11) and Ventsel and Freidlin.(12) As we will be concerned here about some general features of the solutions of (1), we shall not impose more restrictions upon the W_r. However, the features of chemical reactions will provide certain restrictions. For example, for positive r (molecule number is decreasing), $W_r(n)$ is proportional to a binomial factor $\binom{n}{r}$, and is often (but not necessarily always) given by a low-order polynomial in n. As we shall see, such properties will not influence the general conclusions, but they may be important in a qualitative way for certain special conclusions.

It is clearly demonstrated in the cited works that the probability function is almost entirely concentrated to n-values very close to what is given by the deterministic equation (see in particular Refs. 2, 10, and 12):

$$\frac{dn(t)}{dt} = -\sum rW_r(n) \equiv -\alpha_1(n) \tag{3}$$

Next, we will look for stationary solutions of (1) or (2) when the left-hand sides are zero. For a one-step process, the solution becomes almost trivial since Eq. (2) simply gives

$$0 = W_1(n+1)P_s(n+1) - W_{-1}(n)P_s(n) \tag{4}$$

which means that detailed balance is fulfilled. We can then write

$$P_s(n) = \exp z(n/V)P_s(n-1) \tag{5}$$

where

$$\exp z(x) = \frac{W_{-1}(xV-1)}{W_1(xV)} = \frac{w_{-1}(x-1/V)}{w_1(x)} \tag{6}$$

according to the earlier definitions of x and w_r. Note that z, in general, is not close to zero, a fact that is important for the asymptotic properties. If some regularity of w_r is assumed, as was done above, we can separate contributions to z of different order in V:

$$z(x) = z_0(x) + \frac{1}{V}z_1(x) + o\left(\frac{1}{V^2}\right) \tag{7}$$

Then, relation (5) gives the following expression for P_s:

$$\begin{aligned} P_s(n) &= \text{const}\exp\left[\sum_{m=1}^{n} z(m/V)\right] \\ &\approx \text{const}\exp\left[V\int_0^x z_0(\xi)\,d\xi\right]\exp\left[\int_0^x z_1(\xi)\,d\xi - (1/2)z_0(x)\right] \end{aligned} \tag{8}$$

For getting this in a correct asymptotic form, the ordinary Euler–McLaurin summation formula is used. With the assumptions made, it should be clear that this provides a correct asymptotic form for P_s and that the limit of

$(\ln P_s)/V$ is a well-behaved function of x. That limit will obtain extremal values when $z_0(x) = 0$, i.e., when $w_1(x) = w_{-1}(x)$, which is the condition for the deterministic equation (3) to have a stationary point.

In a general case, one should expect the same asymptotic form for the probability function. This is considered in some detail by Kubo *et al.*$^{(5)}$ and is shown rigorously for stochastic processes with these asymptotic features by Ventsel and Freidlin$^{(12)}$ (although their description of the process is different from the one here). Because of this asymptotic property, one should expect a relatively simple behavior in $\ln P$, and thus put the emphasis upon functions determining this logarithm. This was considered in the cited work by Kubo *et al.* and later reformulated by Görtz and Walls.$^{(6)}$ In their method, one expects a relation of type (5) to be valid, and that z becomes a simple function such that (7) and (8) are still meaningful. If one uses (5) in Eq. (1), and considers the lowest order of the asymptotic contribution, one gets an equation for z_0 of the form

$$\sum_{\substack{r=-R \\ r\neq 0}}^{R} w_r(x)\{\exp[rz_0(x)] - 1\} = 0 \tag{9}$$

If, instead Eq. (2) had been used, the following equation is obtained:

$$\sum_{p=R}^{2H-1}\left[\sum_{r=p+1-R}^{R} w_r(x)\right]\exp[pz_0(x)] - \sum_{p=0}^{R-1}\left[\sum_{R-p}^{R} w_{-r}(x)\right]\exp[-pz_0(x)] = 0 \tag{10}$$

As an equation in $y = \exp z_0$, (9) is an equation of degree $(2R + 1)$, and always has a solution $y = 1$. This solution is absent in (11), which is one degree lower. Otherwise, the equations are equivalent. For certain general considerations, (9) is more symmetric, and can therefore be preferable. In principle, every z_0 solution of (9) or (10) can give rise to a contribution to P_s of the form (8). The following features which have not been shown before are demonstrated in the Appendix:

(a) Besides the solution $z_0 = 0$ of (9), there is one unique, real solution of (9) or (10) which is the only one that can correspond to a meaningful physical function.

(b) There are $R - 1$ complex z_0 solutions with real parts larger than z_0 and zero.

(c) There are $R - 1$ complex z_0 values with real parts less than z_0 and zero.

(In the last cases, we do not count those that are obtained by adding $2\pi i$ to other solutions.) We shall refer to these as solutions of type a, b, and c, respectively. As Eq. (10) is a low-order equation in y, and there will never be any solution approaching the physical one, this solution (type a) is a well-behaved function of x. It should be clear that the corresponding contribution of form (8) really fulfills the basic master equation. This will also be the case for corresponding contributions from the other types of solutions. Naturally, the solution shall be uniquely determined by the master equation and boundary conditions, which in this case means that the solution shall be extended down to the lowest order in order to fulfill the first R equations of type (1) or $(R-1)$ equations of type (2). Further, it is necessary that P go to zero for large values of n. The latter property immediately forbids the occurrence of contributions of type-b solutions with large real parts of the z_0, yielding ever-increasing contributions. Further, such a contribution would lead to P functions that alternate in sign. However, the smaller type-c solutions cannot be outruled and they are, in fact, necessary when the probability function is extended to lowest order. Then, one should have a linear contribution:

$$P_s = \sum A_i p^{(i)}(n) \tag{11}$$

This sum includes $R-1$ contributions from the type-c solutions and as there are $R-1$ equations of the lowest orders of type (2) with some coefficients put equal to zero, the corresponding A_i coefficients in (11) will be completely determined. How this can be done is shown by explicit examples at the end of this section. As the real part of these solutions is always negative, the corresponding contribution in (11) will go rapidly to zero and one is left with the physical contribution. (As the solution of the master equation should be unique and our obtained solution fulfills the correct asymptotic features, it must be the correct solution of the problem.) We also mention that because of the solutions of type b, any direct numerical calculation of the P_s from (1) or (2) by a recursion procedure is bound to fail as there will always be a small contribution of these solutions, originating, e.g., from rounding errors, which eventually will dominate, and lead to a completely wrong behavior. It can easily be checked that any starting values will soon yield a wildly oscillating function corresponding to the solution with largest real part of z_0.

Once we have calculated the lowest-order z_0 from (9) or (10), it is possible to get a correction term, z_1, defined in (7). We put

$$w_r\left(x + \frac{p}{V}\right) = w_r^{(0)}(x) + \frac{1}{V} w_r^{(1)}(x) + \frac{p}{V} w_r^{(0)\prime}(x) + o\left(\frac{1}{V}\right) \tag{12}$$

and

$$z\left(x + \frac{p}{V}\right) = z_0(x) + \frac{1}{V} z_1(x) + \frac{p}{V} z_0'(x) + o\left(\frac{1}{V}\right) \tag{13}$$

Relation (5) yields:

$$\begin{aligned} P(n+r) &= \exp\left[\sum_{p=1}^{r} z\left(x+\frac{p}{V}\right)\right]P(n) \\ P(n-r) &= \exp\left[-\sum_{p=0}^{r-1} z\left(x-\frac{p}{V}\right)\right]P(n) \end{aligned} \tag{14}$$

If (13) is used in (14), then these expressions together with (12) in Eq. (9) give the following when contributions in different order in V are separated:

$$z_1 = -\frac{\sum_{r=-R}^{R}\left\{w_r^{(1)}(e^{rz_0}-1)+\left[rw_r^{(0)\prime}+\frac{r(r+1)}{2}w_r^{(0)}z_0'\right]e^{rz_0}\right\}}{\sum_{r=-R}^{R} rw_r^{(0)}e^{rz_0}} \tag{15}$$

We note that, at a stationary point x_s where $\sum rw_r^{(0)}=0$ and $z_0=0$, the denominator vanishes. However, then also the upper part vanishes and the expression goes to a well-defined limit. This follows from the expression for $z_0(x)$ when x is close to x_s. If (9) is expanded in powers of z_0, one gets

$$\sum_{p=1}^{\infty}\frac{1}{p!}\alpha_p(x)z_0^{p-1}=0 \tag{16}$$

where α_p is the pth moment of $w_r^{(0)}$:

$$\alpha_p^{(x)}=\sum_{r=-R}^{R} r^p w_r^{(0)}(x) \tag{17}$$

At $x=x_s$, $\alpha_1(x_s)=0$. The first terms of (16) give

$$z_0(x)=-2\frac{\alpha_1(x)}{\alpha_2(x)}-\frac{4}{3}\frac{\alpha_3(x)\alpha_1^2(x)}{\alpha_2^3(x)}+O\left(\alpha_1^3\right) \tag{18}$$

(18) can be regarded as a power series in α_1. In the limit when α_1 goes to zero, (18) and (15) yield

$$z_1=-\frac{\alpha_2'-\alpha_1'+2\alpha_1^{(1)}}{\alpha_2}-\frac{2}{3}\frac{\alpha_3\alpha_1'}{\alpha_2^2}+O(\alpha_1) \tag{19}$$

$\alpha_1^{(1)}=\sum rw_r^{(1)}$. Note that z_1 always is of order 1. Through (8), its contribution to $\ln P$ is also of order 1, which is much smaller than the corresponding contribution from z_0. In the important region where α_1 and z_0 are of order $1/\sqrt{V}$, the contribution of z_0 is of order $\sqrt{V}$. The behavior of $\ln P_s$ is then always dominated by z_0, given by Eqs. (9) or (10). (This is by no means self-evident. The corresponding correction term to the root $z_0=1$ of (9) is of order $\sqrt{V}$ in that region.)

In the following examples we show the features of the described method for obtaining the stationary probability function in two models.

Example A. In the first example, we consider a case where all $w_r(n)$ are of order unity when $n/V \ll 1$, and there can be put as n-independent constants. We put $R = 3$ and $w_3 = 4$, $w_2 = 2$, $w_1 = 1$, $w_{-1} = 26$, $w_{-2} = 13$, $w_{-3} = 14$. The Eq. (10) for $y = e^{z_0}$ becomes

$$4y^5 + 6y^4 + 7y^3 - 53y^3 - 27y - 14 = 0$$

with roots $y_1 = 2$ (the physical one), $y_{2,3} = -3/2 \pm i(\sqrt{19}/2)$ (the unacceptable ones of type b), and $y_{4,5} = -1/4 \pm i(\sqrt{3}/4)$ (the ones of type c). The two first equations of (1) are

$$4P(3) + 2P(2) + P(1) - 53P(0) = 0$$

$$4P(4) + 2P(3) + P(2) - 54P(1) + 26P(0) = 0$$

We put $P(n) = \text{const}\, 2^n + (a + ib)y_4^n + (a - ib)y_5^n$ and determine a and b from these equations. In the higher-order equations this expression is automatically a solution. One gets

$$P(n) = \text{const}\left(2^n - \frac{11}{53}\frac{1}{2^n}\cos\frac{2\pi n}{3} - \frac{73}{159\sqrt{3}}\frac{1}{2^n}\sin\frac{2\pi n}{3}\right)$$

This is the complete solution which avoids y_2 and y_3. The two last terms will soon be negligible compared to the first one.

Example B. In realistic models of chemical reactions, it is generally valid that $W_r'(n)$ (the rate with which r out of n molecules simultaneously form another molecule) is proportional to $\binom{n}{r}$ (or to a higher expression), while $W_{-r}(n)$ at least for small n, can be put as constants. We consider $R = 2$ and put

$$w_2 = ax(x - 1/V), \qquad w_1 = bx, \qquad w_{-1} = c, \qquad w_{-2} = d$$

Eq. (10) is then

$$y^3ax^2 + y^2(ax^2 + bx) - y(c + d) - d = 0$$

The solutions are to lowest order in x

$$y_1 \approx \frac{\left[b^2 + 4a(c + d)\right]^{1/2} - b}{2a}\,\frac{1}{x} \qquad \text{(the physical one)}$$

$$y_2 \approx -\frac{\left[b^2 + 4a(c + d)\right]^{1/2} + b}{2a}\,\frac{1}{x} \qquad \text{(the one of type b)}$$

$$y_3 \approx -\frac{d}{c + d} \qquad \text{(the one of type c)}$$

The first equations are

$$P(2)\frac{2a}{V^2} + P(1)\frac{b}{V} - P(0)(c+d) = 0$$

$$P(3)\frac{6a}{V^2} + P(2)\frac{2b}{V} - P(1)\left(c + d + \frac{b}{V}\right) + P(0)c = 0$$

and for $n > 1$ $(n \ll V)$

$$P(n+2)\frac{(n+2)(n+1)a}{V^2} + P(n+1)\frac{(n+1)b}{V}$$
$$- P(n)\left[c + d + \frac{nb}{V} + \frac{n(n-1)a}{V^2}\right] + P(n-1)c + P(n-2)d = 0$$

To lowest order in $1/V$, one finds two solutions of all equations: $P^{(1)} = \text{const}(y_1^n V^n / n!)$ and $P^{(2)} = \text{const}(y_2^n V^n / n!)$. The first corresponds to the physical y_1, and is the acceptable solution. A contribution from $P^{(2)}$ must be avoided. In this case, y_3, the absolute value of which is always much lower than y_1, never does give a significant contribution.

3. THE FOKKER–PLANCK EQUATION

As $x = n/V$ turns into a continuous variable when V goes to infinity, it is tempting to look for an asymptotic differential equation of diffusion type for $P(n) \equiv P_V(x)$. Such a description is provided by the so-called nonlinear Fokker–Planck equation. (Certainly, this notation is not quite appropriate as it is a linear partial differential equation with coefficients that are nonlinear in the independent variables. However, since it is commonly known under this name, we shall use it here.) There has been some dispute during the last years about the relations between this equation and the master equation (1), and our aim here is to try to clarify the situation by looking at the basic arguments and by comparing, in a direct way, solutions of the two equations.

The nonlinear Fokker–Planck equation (which we refer to as FP) is

$$\frac{\partial P_V(x,t)}{\partial t} = \frac{\partial}{\partial x}\left[\alpha_1(x)P_V(x,t)\right] + \frac{1}{2V}\frac{\partial^2}{\partial x^2}\left[\alpha_2(x)P_V(x,t)\right] \qquad (20)$$

α_1 and α_2 are the first and second moments, respectively, defined in (3) and (17). This equation is obtained from the master equation (1), if one writes the arguments $n + r$ as $x + r/V$, and makes a Taylor expansion to the second order in the small terms (r/V). (See also below, where this procedure will be discussed.)

Now, van Kampen[13] has argued that this equation is not a consistent approximation of the master equation, while Horsthemke and Brenig,[7] by using arguments of Kurtz,^(10) have suggested that (20) indeed is a correct asymptotic representation of the master equation. These statements seem to be quite contradictory but, as already discussed in the Introduction, this is partly due to different views of comparing the solutions and they are not necessarily incompatible with each other. The Horsthemke–Brenig meaning of asymptotic representation should be interpreted in an absolute sense such that the corresponding probability functions shall be close to each other, i.e.,

$$|P^{\mathrm{ME}}(xV) - P_V^{\mathrm{FP}}(x)|$$

is small when V is large in a way that motivates the use of a diffusion type of equation. (The superscripts indicate the respective basic equations.) As pointed out several times, the probability function is essentially different from zero only in a small neighborhood of the most probable value, which means that the comparison is relevant just in that region as every function that is close to zero may fulfill the requirements outside that region. (Since we will here be mainly concerned with the relative comparisons, we will not go any deeper in questions about the complete meaning of the absolute comparison.)

Now, for probability functions of the type that occur here, the fact that the solutions of the two types of equations are close to each other in an absolute sense does not imply that they are close in a relative sense, which would mean that

$$P_V^{\mathrm{FP}}(x)/P^{\mathrm{ME}}(xV)$$

is close to 1. As discussed above, the limit of $(\ln P)/V$ shall exist for processes of this type and a necessary requisite for the relative agreement is that these limits be the same:

$$\left|\frac{1}{V}\left[\ln P_V^{\mathrm{FP}}(x) - \ln P^{\mathrm{ME}}(xV)\right]\right|_{V\to\infty} \to 0 \tag{21}$$

Again, this is a much stronger requirement than the absolute absolute agreement. In fact, (21) is never exactly fulfilled except at the extremal points of the probability.

This should have been clear from van Kampen's^(13) work, but as the point has not been completely settled, we shall briefly discuss some points here. In fact, the arguments of the first section as well as those in the cited works show that we have relations of the form

$$P_V(x + 1/V) = P(n + 1) = e^{z(n+1)}P_V(x)$$

where z in general is not close to zero (even in the limit). Thus, *$P_V(x + 1/V)$ is never close to $P_V(x)$ in a relative sense.* (In an absolute sense, both are very close to zero, except for x near extremal points, and this comparison is not relevant.) Then, again in a relative sense, $P_V(x+p/V)-P_V(x)$ is not close to $p[P_V(x+1/V)-P_V(x)]$, where p is a number of order unity. In this sense, a definition of a derivative as

$$P_V'(x)=\frac{P_V(x+1/V)-P_V(x)}{1/V} \tag{22}$$

is not appropriate and a differential equation as the FP that is based upon such a definition cannot be asymptotically correct. It can still provide a consistent equation in regions where z is close to zero, which is the case close to extremal points. [The relative sense for the differences above should mean that one rather compares expressions like $P(x+p/V)/P(x)-1$. In an absolute sense, all quantities are close to zero except close to extremal points. In that sense, (22) is meaningful as the expression is virtually zero.]

It is easy to see where the arguments for a nonlinear Fokker–Planck equation run into trouble. The equation is obtained from the Kramers–Moyal expansion based upon a Taylor expansion of the right-hand side of (1):

$$\begin{aligned}
&W_r(n+r)P(n+r)-W_r(n)P(n)\\
&\quad=V\left[w_r\left(x+\frac{r}{V}\right)P_V\left(x+\frac{r}{V}\right)-w_r(x)P_V(x)\right]\\
&\quad=V\sum_{m=1}^{\infty}\frac{1}{m!}\,\frac{r^m}{V^m}\,\frac{\partial^m}{\partial x^m}\left[w_r(x)P_V(x)\right]
\end{aligned} \tag{23}$$

This looks like a power series in $1/V$, which fact has been used for breaking the sum after two terms. However, it follows from the earlier arguments that for a derivative according to (22), $\partial P_V(x)/\partial x=O(VP_V)$, and that for a general derivative, $\partial^m P_V(x)/\partial x^m=O(V^m P_V)$. Thus the mth term in (23) is of order $VV^{-m}V^m P_V=VP_V$. *All terms in the Kramers–Moyal expansion are of the same order in* V. This is essentially the van Kampen argument, and shows that, in this sense, the FP is rather an *ad hoc* approximation to the master equation. [Note that the factor P_V in the orders above signifies the relative sense. Horsthemke and Brenig, whose work should have another interpretation, do not use the expansion (23), although some of their arguments seem to be related to this.]

As stated in the Introduction, if there is one maximum of the probability function and no other important structure, all important features follow from the behavior close to the maximum. In this case, any knowledge of the relative behavior of the probability function far from the maximum is certainly irrelevant. Then, the absolute comparison is sufficient and the functions given by the master equation and the Fokker–Planck equation describe both all-important features in the same way.

However, if there are several structures, e.g., two probability maxima, it is important for judging their mutual importance to follow the *relative variations* over intermediate regions. Then, if (21) is not at least approximately fulfilled in these intermediate regions, one cannot decide which of two maxima is the highest. This fact is clear also in the second work by Horsthemke *et al.*,(4) where multiple structures are concerned.

For this reason, it seems important to consider the validity of (21) far from probability maxima. As discussed by Horsthemke and Brenig, the FP always contains some important qualitative information of the probability function, and the appealing features of the FP makes it relevant to ask when the solution of the FP is a good approximation to that of the master equation also in a relative sense. We now discuss that problem, and our main criterion will be that (21) is approximately valid.

We shall here confine the discussion to the one-variable case and the stationary solution, that is given for the master equation by the methods of Section 1. For the FP of (20), the stationary solution is given by

$$P(x) = \text{const} \exp\left[-V\int_0^x 2\frac{\alpha_1(\xi)}{\alpha_2(\xi)}\,d\xi \right] \tag{24}$$

This is of the same form as (8) with $-2\alpha_1/\alpha_2$ in the place of z_0. In an expression as (23), these quantities shall be compared. From the expression (18), it is seen that this term is the same as the first-order contribution to z_0 close to a probability maximum. (It should be clear from the earlier discussion that the solutions should agree in that order.) The second term of (18)

$$-\frac{4\alpha_3\alpha_1^2}{3\alpha_2^3}$$

describes the first deviation of the FP solution when going away from the maximum. One should therefore expect the FP equation to yield a good approximation in a parameter range where α_3 and α_1'/α_2 are small (α_1 is small close to the maximum, and should be represented by its derivative.) In particular, this is the case for a one-step process where $\alpha_3 = \alpha_1$ and thus also vanishes at the maximum. In that case, the comparison can be made very precise. By using (6), we have

$$w_{-1}/w_1 = e^z$$

so that

$$-2\frac{\alpha_1}{\alpha_2} = 2\frac{w_{-1} - w_1}{w_{-1} + w_1} = 2\tanh\frac{z}{2} \tag{25}$$

Clearly, this is never equal to z unless $z = 0$, but they are approximately equal in a fairly large range (owing to the rapid convergence of series expansions for hyperbolic functions.) The relative deviation

$$\frac{z}{2\tanh(z/2)} - 1$$

is 0.001 for $z = 0.11$, 0.01 for $z = 0.35$, and 0.1 for $z = 1.11$. The relative difference between the logarithms of the probability functions, normalized according to the maxima, will be still better. The conclusion is that the FP expression yields a good approximation for the solution of the master equation of a one-variable one-step process as long as the absolute value of the parameter z is less than about 1. If z becomes much larger than 1, the agreement will disappear completely as (25) tends to cut off large z values. This is apparent in one of the examples below.

In this way, the comparison between the two types of equations can be made very directly, and we will show that in three examples.

It is important to note that, for a large number of models, z will never be large in important intermediate regions where the relative variation is essential. Then, the results from the master equation and the FP can agree quite well. In fact for a simple model of two probability maxima and a minimum in between, based upon chemical reactions, such as the Schlögl model (treated in Ref. 4 and our Section 6), w_1 and w_{-1} are low-order polynomials in x. For the extremal points, α_1 is zero. This fact and the one that w_1 and w_{-1} must be positive will in all models lead to the fact that z never is large, and commonly much smaller than 1. We come back to this feature in Section 6, and this is the reason that Horsthemke *et al.*$^{(4)}$ find an extremely good agreement between the two types of solutions. (Still, the expressions are not exactly the same and the qualitative discussion in that work cannot account for the agreement.)

For a general type of process (i.e., not a one-step process), α_3 does not vanish at the probability maximum and the agreement between the solutions is less controlled. For certain models, the FP can still be a good approximation, but in other cases it may well be possible that the linear Fokker–Planck equation (were α_2 is put equal to a constant and α_1 equal to a constant derivative times the deviation from the extremal point) is better than the nonlinear equation.

To show these features, we consider three examples. The first is a simple one-step process, where we show that the agreement between the solutions of the two equations is very good in a large range. The two other examples are constructed to show cases where the agreement is bad. The second example includes a step where three molecules change and in that case there is no control of the solution of the FP as compared to that of the master equation. We note that this is a model that contains only one maximum and therefore the disagreement would be less relevant according to the discussion above. However, in order to get information about the relative agreement, the information in an intermediate region would be lost in the FP. The third example contains two probability maxima and shows that, in a general case, the master equation and the FP can yield completely different results. In that case, w_{-1} and w_1 are given by exponential functions, which makes it possible to get a greater variation of z. This model is not given by a chemical reaction scheme, but, as said above, the general arguments are not limited to such cases. In fact, only a demonstration of this case can fully motivate the good agreement in the Schlögl model.

Example C. The reaction

$$X \underset{k_2}{\overset{k_1}{\leftrightarrows}} A$$

with the concentration of A constant is one of the simplest ones to describe. We put $w_1(x) = k_1 x = x$, $w_{-1}(x) = k_2 a = 1$. $x = n_x/V$ is the concentration of X, a that of A. The complete solution to the master equation according to (8) is simple:

$$P(n_x) = \frac{V^{n_x}}{n_x!} e^{-V} \tag{26}$$

z_0 is $-\ln x$, $-2\alpha_1/\alpha_2 = 2(1-x)/(1+x)$. If the integrals in $\ln P$ are taken from the maximum $x = 1$, we get

$$\int_1^x z_0\,d\xi = -x\ln x + x - 1, \qquad \int_1^x 2\frac{\alpha_1}{\alpha_2}\,d\xi = 4\ln\frac{x+1}{2} - 2(x-1) \tag{27}$$

These functions are shown in Fig. 1. The relative difference is about 2% when $x = 2$ or 0.5, and even when x is as different from 1 as 5 or 0.15, the relative difference is not more than 10%.

Example D. Consider a more complicated scheme:

$$3X \underset{k_2}{\overset{k_1}{\leftrightarrows}} A$$

$$B \underset{k_3}{\rightarrow} X$$

where A and B are held constant. We put

$$w_3 = k_1 n_x^3/V = 6x^3$$
$$w_{-3} = k_2 a/V = 1$$
$$w_{-1} = k_3 b/V = 15$$

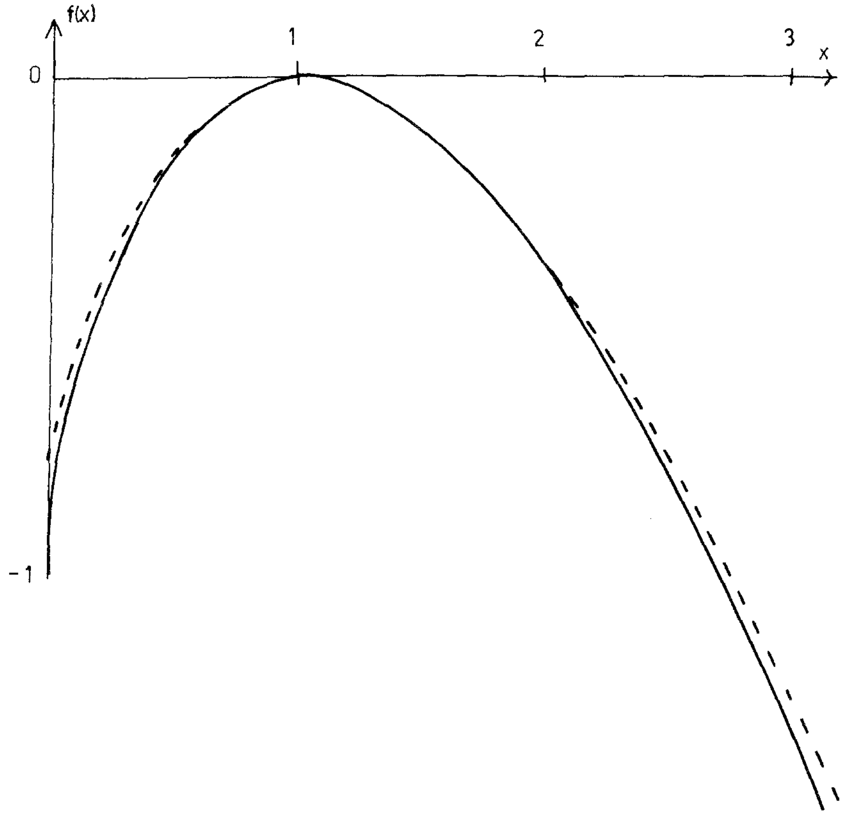

Fig. 1. Comparison between the logarithms of the stationary probability function for the master equation (——) and the nonlinear Fokker–Planck equation (- - - -) for the model in example C, formula (27).

The method of Section 2 is used, and Eq. (10) for $y = e^z$ is

$$(y^5 + y^4 + y^3)6x^3 - 16y^2 - y - 1 = 0 \tag{28}$$

This shall then be compared to the FP expression

$$z_{\mathrm{FP}} = -2\alpha_1/\alpha_2 = 6(1 - x^3)/(9x^3 + 4) \tag{29}$$

and is also compared to the linearized expression

$$z_1 = 2(1 - x)\alpha_1'(1)/\alpha_2(1) = 18(1 - x)/13 \tag{30}$$

The expressions are shown in Fig. 2. It is seen that z_{FP} of (29) is never a good approximation and that close to $x = 1$, (30) is indeed better. For $x \gtrsim 0.8$, the correct z of (30) is always closer to z_1 of (30) than to z_{FP} of (29).

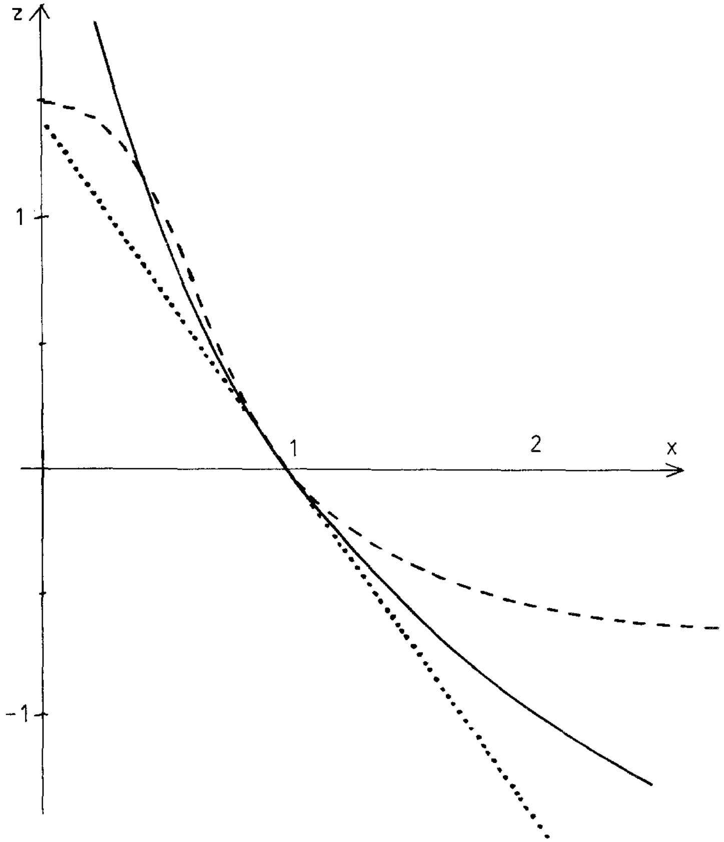

Fig. 2. Comparison between the different expressions of example D. (——) is obtained from the master equations, Eq. (28), (- - - -) from the nonlinear Fokker–Planck equation, Eq. (29), and (· · ·) from the linear Fokker–Planck equation, Eq. (30).

Example E. A one-step process is considered with

$$\begin{aligned} w_{-1}(x) &= \exp\left[50e^{-4(x-2)^2} + \frac{29x^2+75}{2000}\right] \\ w_1(x) &= \exp\left[50xe^{-4(x-2)^2} + \frac{x^3+103}{2000}\right] \end{aligned} \tag{31}$$

The corresponding z_0 is

$$z_0 = \ln\frac{w_{-1}}{w_1} = (1-x)\left[50e^{-4(x-2)^2} + \frac{(x-3)(x-25)}{2000}\right] \tag{32}$$

For these, the master equation solution of (8) and the Fokker–Planck solution of (24) are calculated by numerical integration. We show the

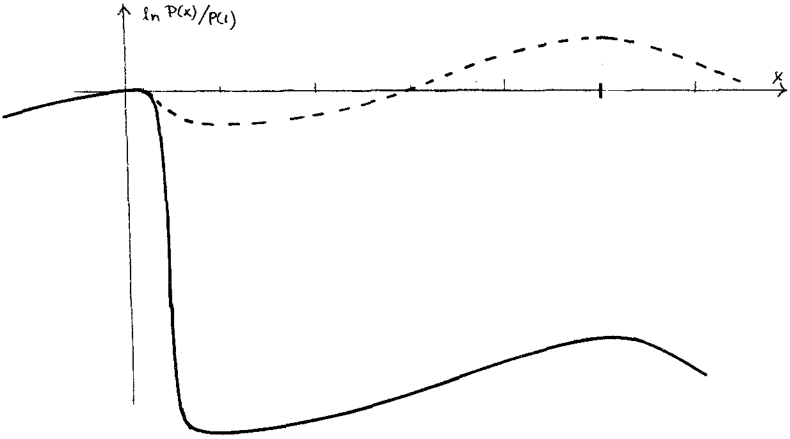

Fig. 3. The solution of the master equation (——) and the nonlinear Fokker–Planck equation (- - - -) for the model of example E. The figure shows $\ln[P(x)/P(1)]$ for the two cases and the main emphasis should be put on the quite different appearance of the first peak at $x = 1$.

expressions

$$(1/V)\left[\ln P_V(x) - \ln P_V(1)\right] \tag{33}$$

in Fig. 3. ($x = 1$ is the first peak of P_V.)

The probability function has two maxima, one of which occurs at $x = 1$, and the other very close to $x = 25$. For the master equation solution, the first peak is the dominating one. It is fairly narrow and very steep. The second maximum is much broader. Because the FP expression cuts off large z values, it cannot show a very steep change of the probability. The first peak is therefore greatly reduced in the FP expression and the second peak will dominate the expression. In the limit of large V, P^{ME} and P^{FP} will show peaks with essentially all of the probability at quite different places.

We remark finally that the w_{-1} and w_1 are chosen to give a simple expression of z_0. From general grounds, one cannot argue that this choice is peculiar in any way. Rather, the polynomials of chemical reactions are special choices since a positively definite polynomial of low order gets severe restrictions on the variation possibilities.

4. EIGENVALUE FORMALISM FOR THE TIME DEPENDENCE

In the last sections we shall consider time-dependent solutions of the master equation (1). An appropriate and systematic way to treat the problem is to write the probability function as a sum of eigensolutions:

$$P(n,t) = P_s(n) + \sum_i P_i(n) e^{-V\gamma_i t} \tag{34}$$

P_s is the stationary solution, treated in the earlier sections. (We omit cases that lack normalizable stationary probability functions.) P_i is the eigenfunction of the equation:

$$-V\gamma_i P_i(n) = \sum_r \left[W_r(n+r)P_i(n+r) - W_r(n)P_i(n) \right] \tag{35}$$

γ_i is an eigenvalue with the meaning of inverse relaxation time. The factor V is used for proper scaling. We will mostly be interested in small γ values, corresponding to slow decay of metastable states.

We have a requirement that the sum of all probabilities is 1 at all times. As P_s is what remains after long times, it must be valid that

$$1 = \sum_n P(n,t) = \sum_n P_s(n) \quad \text{and} \quad \sum_n P_i(n) = 0 \qquad \text{for all } i \tag{36}$$

In this case, the contributions $P_i(n)$ are not positively definite, which, however, must be the case for the total probability $P(n,t)$. The condition (36) together with the master equation is sufficient to determine the eigenvalues.

An ansatz of the form (5) can still be used, and the resulting equation for z_0 [cf. (9)] is

$$\sum_{r=-R}^{R} w_r(x)\left[\exp rz_0(x) - 1\right] + \gamma = 0 \tag{37}$$

In this case, there is no simple equation corresponding to (10) and the solution corresponding to $z_0 = 0$ in (9) (call it type d) cannot be eliminated here. In fact, that solution will become mixed with the physical z_0 (of type a) in certain regions, where this method loses its meaning.

This will in particular occur close to extremal points of P_s, where $\alpha_1(x) = 0$. At such a point, $\sum w_r(\exp rz_0 - 1)$ has a double root, $z_0 = 0$. Close to this point, the sum can be approximated by the second derivative expression $\alpha_2 z_0^2/2$, and (37) becomes

$$(1/2)\alpha_2 z_0^2 + \gamma = 0$$

The eigenvalues γ must always be positive in an equation of type (35), and therefore this equation cannot have any real solutions. Thus z_0 cannot be determined in this way in such a region. Fortunately, a region close to a stationary point of this kind can be treated by a number of other methods.

In particular, in this region, the linear Fokker–Planck equation is justified. For the eigenvalue problem, one has

$$-\gamma P(x) = \alpha_1'(x_s)\,\frac{d}{dx}\left[(x - x_s)P(x)\right] + \frac{\alpha_2(x_s)}{2V}\,\frac{d^2P(x)}{dx^2} \qquad (38)$$

The important problem is to follow $P(x)$ across x_s. This can be done by analytic methods and the equation is already met with in a description of Brownian motion.(15) The solution of (38) can be written in terms of parabolic cylinder functions, U and V.(20) We put

$$\xi = \left(\frac{2\alpha_1' V}{\alpha_2}\right)^{1/2}(x - x_s) \quad \text{and} \quad A = \gamma/\alpha_1'(x_s) \qquad (39)$$

The solutions of (38) are

$$e^{-\xi^2/4}U(-1/2 - A, \xi) \quad \text{and} \quad e^{-\xi^2/4}V(-1/2 - A, \xi)$$

The asymptotic expressions for large, positive ξ ($\xi \gtrsim 3$) are

$$\begin{aligned} e^{-\xi^2/4}U(-1/2 - A, \xi) &\sim e^{-\xi^2/2}\xi^{A} \\ e^{-\xi^2/4}V(-1/2 - A, \xi) &\sim (2/\pi)^{1/2}\xi^{-1-A} \end{aligned} \qquad (40)$$

The first corresponds to the contribution of the physical z_0 solution, the second one to the type-d contribution. When $\xi < 0$, the expressions are mixed:

$$\begin{aligned} e^{-\xi^2/4}U\left(-\tfrac{1}{2} - A, \xi\right) &\sim \cos A\pi e^{-\xi^2/2}|\xi|^{A} + \frac{(2\pi)^{1/2}}{\Gamma(-A)}|\xi|^{-1-A} \\ e^{-\xi^2/4}V\left(-\tfrac{1}{2} - A, \xi\right) &\sim \cos A\pi\left(\frac{2}{\pi}\right)^{1/2}|\xi|^{-1-A} + \frac{\sin^2\pi A}{\pi}\Gamma(-A)e^{-\xi^2/2}|\xi|^{A} \end{aligned} \qquad (41)$$

These expressions are valid for small γ values. Thus even if we start (for $\xi > 0$) with a pure U function, there will be a contribution of an unwanted kind in the last term of (41) for negative ξ, If γ is small, the coefficient is small, but it will always dominate at large enough ξ. Clearly, the contribution disappears if $1/\Gamma(-A) = 0$, i.e., A is an integer, in which case, we get solutions of Hermite function type. If the negative ξ values correspond to an intermediate region between probability maxima, contributions from the last term in U in (41) can be allowed. They should, however, not occur in regions below the lowest maxima or above the highest one.

The eigenvalue formalism for decaying states has been used by Tomita *et al.*(15) for treating the Fokker–Planck equation, and the crossing problem encountered here is treated for that equation in a subsequent paper by Leimar *et al.*(17)

5. A MODEL OF AN ABSORBING STATE

The typical reaction yielding an absorbing state contains two kind of processes(8):

$$X + X \underset{k_2}{\overset{k_1}{\leftrightarrows}} A + X$$

$$X \underset{k_3}{\rightarrow} B$$

The number of A molecules is kept constant. One has

$$\begin{aligned} W_1(n_x) &= k_1 n_x(n_x - 1) + k_3 n_x \\ W_{-1}(n_x) &= k_2 n_A n_x \end{aligned} \tag{42}$$

Without any loss in generality, we simply write

$$W_1(n_x) = n_x^2/V, \qquad W_{-1}(n_x) = n_x$$

The only stationary possibility is a state without any X molecules, i.e., when $P(0) = 1$, and all other $P(n) = 0$. We now look for a decaying state of the form of (34) with $P(n,t) = \delta_{n,0} + P_1(n)e^{-V\gamma t}$ + higher-order terms. As will be seen, there is always a very small γ eigenvalue. The equation for $P_1(0)$ is

$$-\gamma P_1(0) = \frac{1}{V} P_1(1) \tag{43}$$

As done for Eq. (2), we can sum Eq. (35) to order n, and obtain

$$-\gamma \sum_{m=0}^{n} P_1(m) = \frac{n}{V} P_1(n+1) - \frac{n^2}{V^2} P_1(n) \tag{44}$$

Further, (36) shall be valid:

$$P_1(0) + \sum_{n=1}^{\infty} P_1(n) = 0 \tag{45}$$

This is equivalent by demanding that P_1 in (44) shall go to zero for large n. If it is not fulfilled, the left-hand side is not zero, and P_1 would not go to zero. Since $P(n,t)$ is positive, it is necessary that all $P_1(n)$ with n larger than zero be positive. $\sum P_1(n) = -P_1(0)$ is then positive, and will be much larger than $P_1(1)$. From (43) then follows that γ must be very small. In fact, for n from 1 and passing the maximum of P where $n = V$, the left-hand side of (44) is very small compared with the other terms, and can be neglected. Because of (45), the left-hand side goes to zero at large n, and it can be neglected in all equations but the one with $n = 0$, i.e., (43). (This procedure is completely confirmed by the result below.) By recursion, one finds

$$P_1(n) = \frac{V^{n-1}}{n \cdot n!} P_1(1) \tag{46}$$

Thus (see, e.g., Ref. 20)

$$\sum_{n=1}^{\infty} P_1(n) = \frac{P_1(1)}{V}\left[Ei(V) - \gamma - \ln V\right] \approx \frac{P_1(1)}{V^2} e^V \tag{47}$$

Ei is an exponential integral and the last expression is valid for $V \gg 1$. Equations (43) and (45) give

$$\gamma = -\frac{1}{V}\frac{P_1(1)}{P_1(0)} = \frac{P_1(1)}{V\sum_1^{\infty} P_1(n)} = Ve^{-V} \tag{48}$$

The relaxation time for this state is $1/V\gamma = e^V/V^2$. For times up to that order of magnitude, the probability function $P_1(n)$ with a maximum at $n = V$ will sustain. This time is extremely long even for moderately large systems. If $V = 100$, one gets $1/V\gamma = 2.7 \times 10^{39}$ time units. The time unit is given by molecular reaction times, and even if it is as small as 10^{-13} sec, one sees that $1/V\gamma$ is 8.5×10^{18} years! The obvious conclusion is that even in a relatively small system of about 100 molecules, although $P_1(n)$ is not stationary in a strict sense, it must be considered as truly stationary in a physical sense, and will not decay within any meaningful time scale.

This resolves the dilemma that the deterministic equation yielding $n = V$ as the only stable point and $n = 0$ as a stationary but unstable point seems to be in conflict with the stochastic approach, where the extinct situation is the only strict stationary state. The answer is (this has also been discussed in Ref. 3 in a somewhat different manner) that the deterministic equation shows stability with respect to small fluctuations and the extinct state ($n = 0$) does not allow any fluctuations at all. It is indeed stable, but any initial molecule number different from zero will most probably give rise to a $P_1(n)$ distribution as discussed, which in a sense is stationary and where $n = V$ is a stable maximum.

We also mention that, besides Ref. 3, probabilities of events of this kind are also estimated in the work by Wentsel and Freidlin.(12) By the particular model used here, the decaying time can be calculated more precisely.

6. TRANSITION BETWEEN STABLE STATES

We now consider cases where the deterministic equation allows several stable points [where $\alpha_1(x) = 0$]. The typical example is the Schlögl model(14) for coupled reactions:

$$X \underset{k_2}{\overset{k_1}{\rightleftarrows}} A, \qquad 3X \underset{k_4}{\overset{k_3}{\rightleftarrows}} 2X + B$$

The A and B concentrations are assumed to be constants. One has

$$W_1(n_x) = k_1 n_x + k_3 n_x (n_x - 1)(n_x - 2)$$
$$W_{-1}(n_x) = k_2 a + k_4 b n_x (n_x - 1)$$

This model has been studied in several works partly because it shows a kind of phase transition and a critical behavior (see Refs. 4 and 18).

Here, we shall consider the transition between maxima as described by the eigenvalue formalism described above. The same problem, when formulated by the nonlinear Fokker–Planck equation, has been considered in some works. This can be treated by a method due to Kramers.(19) Procaccia and Ross(9) use essentially Kramers' method for the transition time, and Tomita *et al.*(15) use an eigenvalue method, which leads to the same result. This problem is also considered by Leimar *et al.*(17) We shall here consider the master equation by an entirely different treatment which, again, leads to the same result. (It should be pointed out that stochastic descriptions of a bistable model are subject to much interest at the moment, and here are only mentioned a few references that are most relevant for the present work.)

We shall consider a general one-step model with given W_1 and W_{-1}. The equilibrium probability, P_s, is then directly given by (8). We assume that it has two maxima at points $n = n_1$ and n_2 ($n_1 < n_2$), which are referred to as stable points. Between the maxima, there is a minimum at $n = n_3$.

As in the model of the previous section, there is a unique, very small γ value (this is confirmed by the results). It contains all information about the transition between the stable points. For sufficiently long times, all contributions to $P(n,t)$ given by (34) except the stationary state and the most long-lived one are extinct. Then, we have

$$P(n,t) = P_s(n) + P_1(n)e^{-V\gamma t}$$

where P_1 fulfills

$$\sum_{n=0}^{\infty} P_1(n) = 0$$

We shall use an equation of the same type as (44) for P_1, which is given by summing the first n equations of (35):

$$w_1(n+1)P_1(n+1) - w_{-1}(n)P_1(n) = -\gamma \sum_{m=0}^{n} P_1(m) \tag{49}$$

For not too large systems, this eigenvalue problem can easily be solved by treating the equation by recursion methods. Figure 4 shows the functions

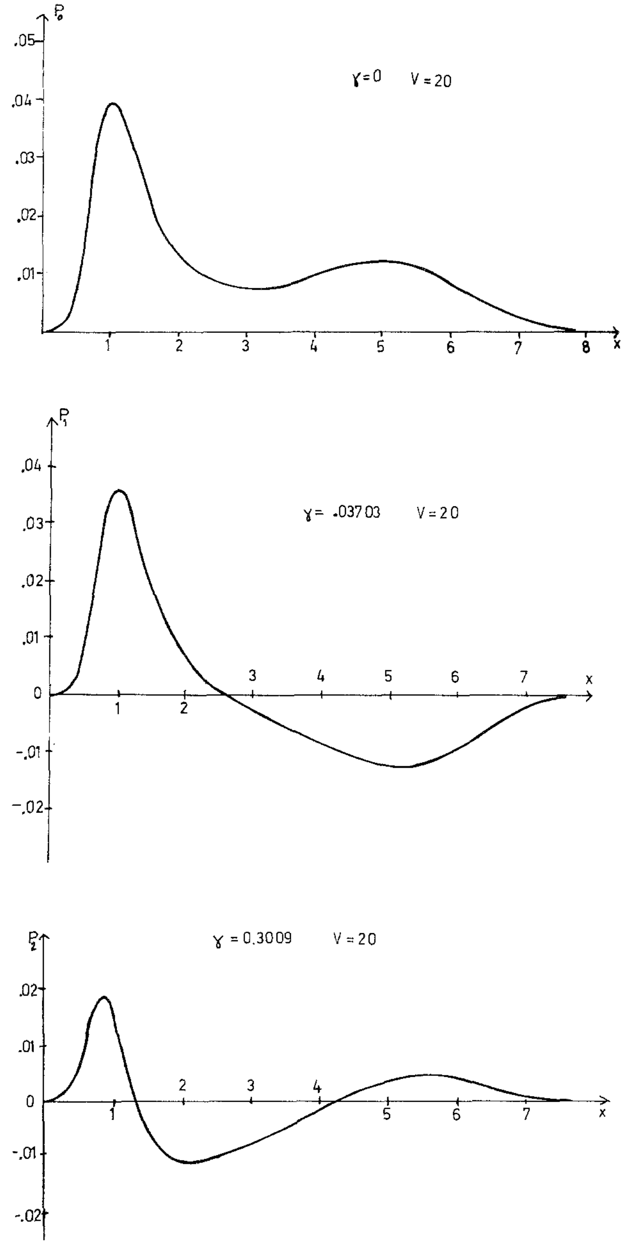

Fig. 4. The eigenfunctions for the probability function of the three lowest γ eigenvalues of the master equation for the Schlögl model. The constants are chosen such that $k_3 = V^{-2}$, $W_1(xV) = W_{-1}(xV+1)$ for $x = 1, 3$, and 5, and $V = 20$. The eigenvalues and functions are calculated directly from (49).

P_s, P_1, and the next contribution to (34), P_2, calculated in this manner for the Schlögl model.

We now consider (49) for large systems, and look for a γ value much smaller than the w functions. The right-hand side is then small compared to the other terms up to and about n_1. In that region, P_1 is almost proportional to P_s. When n varies between n_1 and n_2, P_1 shall change sign and the right-hand side will then yield an essential contribution. (Compare the earlier discussion. In the terminology of Section 4, this means that the type-d solution becomes important.) Later, $|P_1|$ increases and the right-hand side is again small compared to other terms for n around n_2 and for larger values, where the sum goes rapidly towards zero.

We next show how to calculate P_1 between n_1 and n_2. There should be one contribution proportional to P_s according to the discussion above. When γ is small, one can expect that another contribution is proportional to γ, and therefore write

$$P_1(n) = P_s(n) + \gamma P'(n) \tag{50}$$

We use the expression (8) for P_s. Thus with $e^z = w_{-1}/w_1$,

$$P_s(n) = C \exp \sum_{m=0}^{n} z(m)$$

where C is a constant. Equation (49) then becomes an equation for P'. If terms of order γ are neglected,

$$P'(n+1) = e^{z(n+1)}P'(n) - w_1^{-1}(n+1) \sum_{m_1=0}^{n} C \exp\left[\sum_{m=0}^{m_1} z(m)\right] \tag{51}$$

By a recursion procedure, one finds the solution as

$$\begin{aligned} P'(n) &= -\sum_{m_2=0}^{n} \sum_{m_1=0}^{m_2-1} \frac{C}{w_1(m_2)} \exp\left[\sum_{m=0}^{n} z(m) - \sum_{m=m_1}^{m_2} z(m)\right] \\ &= -P_0(n) \sum_{m_2} \sum_{m_1} \frac{1}{w_1(m_2)} \exp\left[-\sum_{m=m_1}^{m_2} z(m)\right] \end{aligned} \tag{52}$$

(The expression can be confirmed by direct insertion and proved by induction.) A typical z function is shown in Fig. 5 (this is calculated from the Schlögl model). It is negative between n_1 and n_3. Because of this, the largest terms of the double sum are those where the exponential contains as many positive terms as possible, and few negative ones, i.e., when $m_1 \approx n_1$ and $m_2 \approx n$ if $n < n_3$, or $m_2 \approx n_3$ if $n > n_3$. Close to n_1, we write

$$z(m) \approx -\frac{2\alpha_1'(n_1)}{\alpha_2(n_1)}(m - n_1) \equiv \frac{m - n_1}{\sigma_1}$$

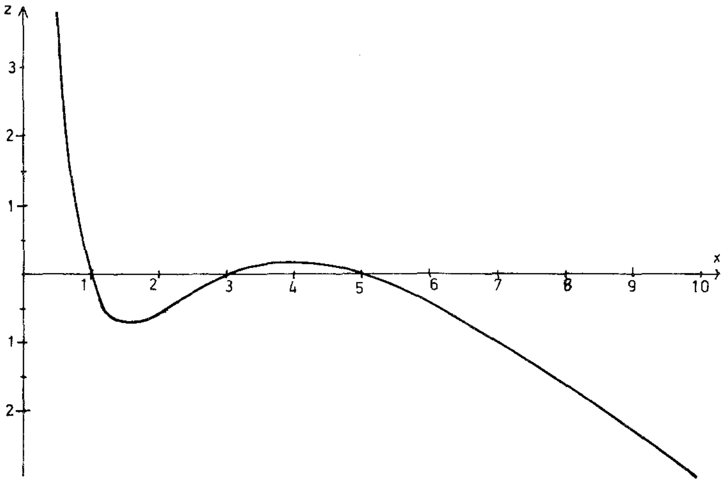

Fig. 5. The $z_0(x)$ of the Schlögl model with the same constants as in Fig. 4 (except that V is assumed to be very large).

which defines σ_1 (which is positive). Close to n_3, we can correspondingly write $z(m) = -(m - n_3)/\sigma_3$. We further put

$$S_{13} = \exp\left[-\sum_{n_1}^{n_3} z(m) \right] = \frac{P_s(n_3)}{P_s(n_1 - 1)} \approx \frac{P_s(n_3)}{P_s(n_1)}$$

When $m_1 \approx n_1$, $m_2 \approx n_3$

$$-\sum_{m_1}^{m_2} z(m) \approx -\sum_{n_1}^{n_3} z(m) - \frac{(m_1 - n_1)^2}{2\sigma_1} - \frac{(m_2 - n_3)^2}{2\sigma_3}$$

In (52), the m_1 and the m_2 sums can now be replaced by integrals, and we get for $n > n_3$

$$\begin{aligned} P'(n) &\approx -P_s(n) S_{13} w_1^{-1}(n_3) \int\int dx_1\, dx_2 \exp\left(-\frac{x_1^2}{2\sigma_1} - \frac{x_2^2}{2\sigma_3} \right) \\ &= -P_s(n) S_{13} w_1^{-1}(n_3) 2\pi (\sigma_1 \sigma_3)^{1/2} \end{aligned} \tag{53}$$

For large systems, the essential variation of the probability is restricted to narrow regions around the maximum points and the approximations behind this formula are justified. (For instance, the error introduced by extending the integration limits to infinity is negligible.) γ can now be determined by the requirement that $\sum_0^\infty P_1(n) = 0$.

The picture we now have about P_1 based upon (50) and the considerations above is the following. Up to and above the first probability maximum, n_1, the P' contribution in (50) is negligible compared to P_s.

Between n_1 and the minimum n_3, the P' contribution grows in importance according to (52), and above n_3 it is given by (53). As this expression is proportional to P_s, this decaying state again follows the stationary distribution with, however, a certain factor and changed sign compared to the first peak. For any sum of $P_s(n)$, the two peaks and the narrow regions around these dominate the sums. Following our formalism, we have the following:

The contribution to $\sum P_s(n)$ from the first peak is

$$\sum_1 = P_s(n_1)\int dx \exp\left(-\frac{x^2}{2\sigma_1}\right) = (2\pi\sigma_1)^{1/2}P_s(n_1)$$

The corresponding contribution from the second peak is

$$\sum_2 = (2\pi\sigma_2)^{1/2}P_s(n_2)$$

In $\sum P_1(n)$, these regions again dominate and the respective contributions around the peaks are proportional to these expressions.

For the first peak, $n = n_1$,

$$P_1(n_1) \approx P_s(n_1)$$

and the contribution to $\sum P_1(n)$ is $(2\pi\sigma_1)^{1/2}P_s(n_1)$. For the second peak, $n = n_2$, P_1 is given by (50) and (53):

$$P_1(n_2) \approx -P_s(n)\left[S_{13}2\pi(\sigma_1\sigma_3)^{1/2}w_1^{-1}(n_3)\gamma - 1\right]$$

The contribution to $\sum P_1(n)$ is this expression with $n = n_2$ multiplied by $(2\pi\sigma_2)^{1/2}$.

If $\sum_0^\infty P_1(n) = 0$, these two contributions must cancel, i.e.,

$$(2\pi\sigma_1)^{1/2}P_s(n_1) = (2\pi\sigma_2)^{1/2}P_s(n_2)\left[S_{13}2\pi(\sigma_1\sigma_3)^{1/2}w_1^{-1}(n_3)\gamma - 1\right]$$

Thus

$$\gamma = \frac{(\sigma_1/\sigma_2)^{1/2}S_{12} + 1}{2\pi(\sigma_1\sigma_3)^{1/2}S_{13}w_1^{-1}(n_3)} \tag{54}$$

where $S_{12} = P_s(n_1)/P_s(n_2)$ is the quotient between the potential maxima for the stationary distribution. It is generally (except at phase transition points) much smaller or larger than 1 (depending on which maximum dominates). Assuming that the first peak is the largest one, we get

$$\gamma = \frac{W_1(n_3)}{2\pi(\sigma_2\sigma_3)^{1/2}}S_{23} \tag{55}$$

S_{23} is equal to $P_s(n_2)/P_s(n_3)$, i.e., the quotient between the probability at the second (smaller) maximum and at the minimum. The inverse value represents the very small probability to get a large fluctuation to the value

n_3 if one originally had $n = n_2$. It is the appropriate generalization of the Boltzmann factor $e^{-\Delta E/kT}$.

$1/\gamma V$ represents a very large lifetime of a decaying state. If P_1 is the function defined according to (50) and (51), a general time-dependent function can be written as

$$P(n,t) = P_s(n) + \text{const}\, P_1(n) e^{-V\gamma t}$$

The constant can here be chosen according to the initial situation. If it is equal to -1, the above description shows that the first peak at $n = n_1$ is compensated for at $t = 0$, which means that the initial state is almost entirely distributed around $n = n_2$. If instead, the constant was chosen as $-P_s(n_2)/P_1(n_2)$ (which is a positive quantity), the peak at $n = n_2$ is eliminated at $t = 0$. A transition rate can be defined as the initial decay rate (i.e., the time derivative) of the probability around the initial point. The transition rate from a low state to the highest maximum is given by $V\gamma$.

As representing the transition rate, (55) is equivalent to the Kramers$^{(19)}$ result. His method can be directly used in this case if the nonlinear Fokker–Planck equation is appropriate. The result for the latter equation is the same as (55) if S_{23} represents the corresponding quotient between the probability function of the maximum and minimum. (This quotient in general differs for the two kinds of equations.)

We end this section with some qualitative remarks. First, we note that (55) seems to be good also for moderately large values of V. In particular, we have studied the situation of Fig. 4, where $V = 20$, and compared an exact γ value calculated directly from (49) with that of (54) using the features of the stationary solution in Fig. 4a. In this case, the first two figures (nonzero) agree.

A general feature, mentioned earlier, is seen from Fig. 5, namely, that $z(x)$ varies very little between the extremal points. In the case shown there, which should be typical for the Schlögl model, $|z|$ is never larger than 0.1. This has two important consequences. First, it is small enough to allow a good description by the nonlinear Fokker–Planck equation according to the earlier discussion. It seems to be a fair conclusion that the FP can be used with great confidence for describing the relative variation of the probability function around and between the stable points in the Schlögl model. This also explains the good agreement for phase transitions found by Horsthemke *et al.*$^{(4)}$ We emphasize that this conclusion, which is quantitative rather than qualitative, does not follow from their conclusions, nor does Kurtz's paper$^{(11)}$ tell anything about that situation.

The other consequence is that much larger systems are needed to get extreme relaxation times than in the model of Section 5. For the case of Fig. 4, the relaxation time is not very large even if the first stable point, n_1, is equal to 1000. Turner(3) has studied the approach to a stationary state by numerically solving the master equation for a similar model in a comparatively small system. When parameters were changed, the average value was eventually given by the dominant maximum, and no hysteresis was seen. This seems to be an effect of the smallness of the system, and of the low values of z. For a sufficiently large system (probably 1–2 orders of magnitude larger than was considered), the relaxation time $1/V\gamma$ is large enough to prevent transitions during meaningful times. Then, hysteresis would have been seen.

7. DISCUSSION AND EXTENSION OF THE RESULTS

We shall here make a summary of the most important features of this work, and also discuss some possible extensions. First, we have in some detail explored a method for obtaining solutions of the master equation, earlier introduced by Kubo *et al.*(5) and Görtz and Walls.(6) This method is well suited for providing the logarithm of the probability function in a general stationary situation. The asymptotic features come out in a very suitable way. The time-dependent situation, described in Section 4, can be treated by this method for points away from extremal points of the stationary distribution (the stationary points in a deterministic description). Close to such extremal points, the method fails and it should be complemented by other methods. It is quite possible to do that, and such problems are taken up in another paper by Leimar *et al.*(17)

One might hope that the method could be used for cases with several stochastic variables. There are large difficulties of treating the master equation in such cases, and the general form of the Görtz–Walls method, which leads to a highly nonlinear partial differential equation, does not seem to be very tractable. It is important to investigate this further. One possibility is to make an expansion around a probability maximum of the same kind as in Section 2 [formula (18)]. As discussed earlier in this work, higher-order terms yield information about deviations between the two types of main equations. The method by Gardiner and Chaturvedi(10) is probably also useful in such cases.

We next point out that the z function of the main method seems to provide a direct connection to thermodynamic quantities. For a system close to equilibrium and at constant temperature and pressure, the logarithm of the probability function is of course related to the free enthalpy, G. For a chemical reaction, the important quantity is the affinity, which, of course, is the change of G in the reaction. For a simple reaction such as

$$2B \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C$$

the affinity is

$$A = \mu(n_C) - 2\mu(n_B) = RT\ln\frac{k_2n_C}{k_1n_B^2}$$

The expression in the logarithm is the quotient of the total reaction velocity going to the left and that going to the right. Comparison with the formalism in Section 2, in particular formula (6), shows that the z function shall be identified with A/RT (or a multiple of it). In that case, z is the logarithm of a quotient between the total velocity by which n is decreased by 1 and that by which n is increased by 1. The essential difference to the equilibrium situation (where detailed balance is valid) is that the expression in (6) involves the sums of all velocities by which the species X is changed in all reactions. Of these, the most rapid one often dominates. Such a dominance is not found close to equilibrium because detailed balance is valid and all reactions are equally important for providing the fluctuations. In that case, each reaction will separately contribute to G by its affinity. For the system driven far from equilibrium, it seems that the natural generalization is a dynamic affinity involving the rates of all reactions. It seems quite reasonable that a rapid reaction dominates the situation in the driven case. In fact, a similar expression is used for the driven chemical potential difference over a biological membrane where ions are kept at different concentrations by active transport. The potential difference is described by the standard Goldman equation,(21) which contains a logarithm of a quotient between an outgoing ionic electric current and an ingoing one. Again, such an expression is motivated by the special situation, and the current of the ions that masses the membrane most easily (normally the potassium ions) dominates the potential.

When making the connections to thermodynamics, one should therefore consider such a kinetic affinity, and not use the sum of the affinities of the single reactions (at least not in a case as described here). The interpretation of the affinity is simple in formula (6). In a general case, it shall still be proportional to our z_0, given by the solution of (9). For several varying molecule species, the situation is still worse, but, at least formally, it can be introduced in the same way. We remark that an extension of thermodynamic relations based upon the stationary solution of the master equation is also given by Jähnig and Richter.(22) Their λ is the same as z in this work.

A significant part of this work has been devoted to the study of Fokker–Planck type of equations. We have argued that in order to describe the mutual importance of several structures of the probability function, it is necessary to consider relative variations. The important quantity is the logarithm of the probability function. Such a global picture of many structures is not covered by Kurtz's paper,(11) and represents a more general view of the situation than the asymptotic representation by Horsthemke *et al.*$^{(4,7)}$ In general cases, the nonlinear Fokker–Planck equation is not a good approximation for the logarithm of the probability function (and still less a correct asymptotic representation of it) outside a small fluctuation region around a probability maximum. As shown by example, the FP can completely fail to give the mutual importance of two probability maxima.

Still, in special but important cases, the FP and the master equation agree quite well over large regions. This is particularly the case for a one-step process where the reaction rates are given by low-order polynomials of the molecule numbers (they can likewise be described by quotients of such polynomials). Almost all models which have been used in these connections earlier are indeed of that type, which shows that the nonlinear Fokker–Planck equation should be able to be used with confidence in a number of problems. Little is known about the situation of several variables. It is quite possible that the arguments about a one-step process remain the same (for instance, third-order moments will again be the same as the first-order ones). If this is valid, the nonlinear Fokker–Planck equation is satisfactory, which would lead to a great advantage since it is a linear partial differential equation with possibilities for obtaining approximate solutions. A master equation does not seem to be tractable in any easy way for several variables.

A third important problem treated here is the calculation of transition times by the eigenvalue method in Section 4. In Section 6, a general formula for a global transition time between two stable states was derived for the master equation. For a large system, this time becomes so large that it is meaningless. The analysis of the transition is still important for a relatively small system. If the stable states involve a particle number of, say, about 1000, the transitions do not take extremely long times. This can be the situtation in a living cell, where the number of certain molecule species (as enzymes) may be of the order 100–1000 or still lower. Then, asymptotic forms are valid, multiple stable states may well occur, and the transition times will be quite meaningful.

This is also the case for local fluctuations in a larger system. In fact, the global picture given here (and in many other works) should be complemented to a local picture where diffusion effects in the space are taken into account (see, e.g., Ref. 1). For a transition between two stable states, the global transition time can be meaningless when it is as in Section 5. However, local fluctuations in a small region can provide a transition in a shorter time and form a nucleation for further transition. To treat this, it should be relevant to make a cell division of the total system, and consider diffusion between the cells. The size of the cells should be given by some

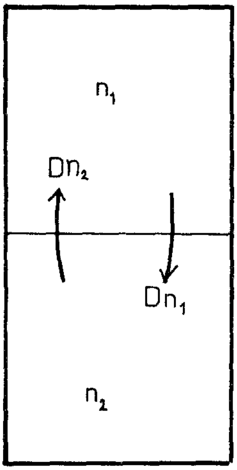

Fig. 6. The model with flow between two subsystems.

effective free mean path of the relevant molecules, and described by the picture in this work.

We here make a preliminary discussion of this problem in order to show some important features. The situation is then simplified and the system of interest is divided in two subsystems as in Fig. 6, where the particle numbers may differ and a corresponding flow of particles arise. If x_1 and x_2 are the particle densities in the subsystems, the following equations are obtained for the average numbers:

$$\dot{x}_1 = D(x_2 - x_1) + \alpha_1(x_1)$$
$$\dot{x}_2 = D(x_1 - x_2) + \alpha_1(x_2)$$

α_1 is, as before, the first moment of the w. We have used the following particular choice of α_1:

$$\alpha_1(x) = (1 - x)(3 - x)(5 - x)$$

This is a typical expression that one obtains from the Schlögl model, and is used earlier in Figs. 3 and 4. Without diffusion ($D = 0$), $x = 1$ or 5 are stable states, and $x = 3$ is an unstable one corresponding to a minimum of the probability. We show solution curves and stationary points for the combined system in Fig. 7a for $D = 1$. In Fig. 7b, the stationary points are shown as functions of varying D. The following features are found.

When D is small, stationary points are found when x_1 and x_2 are close to the values 1, 3, or 5. If both are close to 1 or 5 (stable values without diffusion), the point is stable. If one of x_1 or x_2 is close to 3, one gets a saddle point corresponding to a point of minimum probability that must be

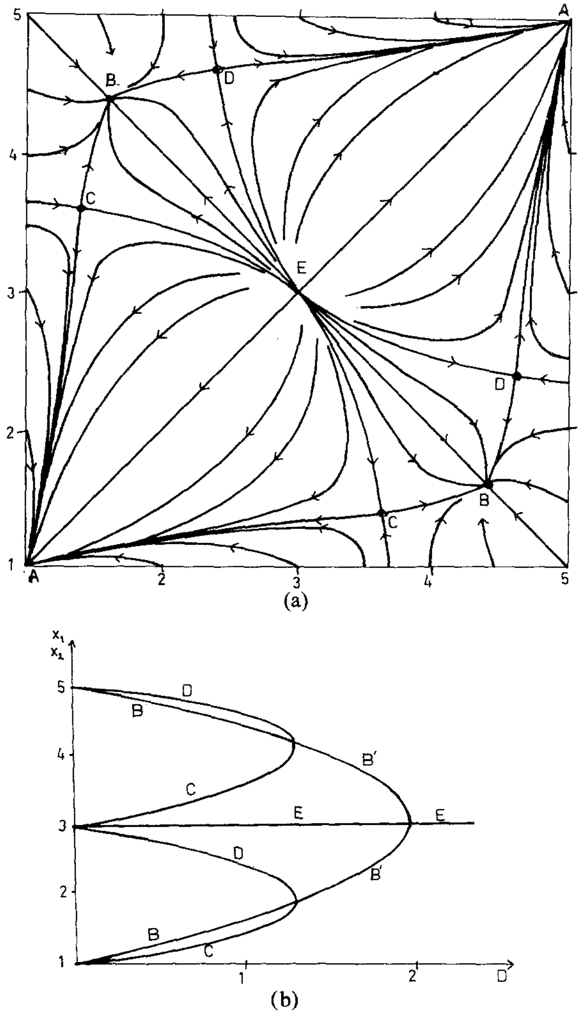

Fig. 7. The features by the diffusion model of two subsystems as discussed in the text. (a) Time evolution curves and stationary points with $D = 1$, and other quantities as in the text. A shows the homogeneous, stable points, B a stable, intermediate state, and C and D unstable saddle points, which are more favorable for the transition than the unstable E. (b) The stationary branches as function of the diffusion constant D. The letters have the same meaning as in (a). Branches of the same letter yield the x_1 and x_2 values of the same solutions.

crossed for a transition. In this case, one subsystem can make a transition between the stable states almost independently of the other subsystem. As an intermediate state, one obtains a stable distribution where the subsystems are in different states.

For larger D values, the diffusion becomes more and more important. The saddle points then approach the stable intermediate where they merge at a certain value of D. Thereafter, there is no stable distribution with different values of x_1 and x_2 but the earlier stable intermediate rather corresponds to a stable point. Again, this is a point of smallest probability that must be passed during a transition. At still higher values of D, all stationary distributions with different values of x_1 and x_2 vanish. The situation is then the same as the global one in Section 6 and the entire system acts as an entity.

This example may be very simple, but the general features are obvious, and will be found also in general cases. For small values of the diffusion constant, small regions of the systems act independently of each other, and stable intermediate states with inhomogeneous particle distributions arise. Such local features are lost for larger values of D, when the entire system behaves in a homogeneous way.

It is the hope that this problem can be treated by stochastic methods. As in Chapter 6, the main emphasis is to be put upon the saddle point, which must be passed and which determines the transition time. This will be further developed in later work.

APPENDIX

We demonstrate here some mathematical properties of equations (10) and (11), which we rewrite here as

$$\sum_{r=1}^{R}\left[w_r(y^r-1)+w_{-r}(y^{-r}-1)\right]=F_A=0 \tag{A}$$

$$\sum_{p=R}^{2R-1}\left(\sum_{r=p+1-R}^{R} w_r\right)y^p-\sum_{p=0}^{R-1}\left(\sum_{r=R-p}^{R} w_{-r}\right)y^p=F_B=0 \tag{B}$$

Equation (A) has a pole of order R at the origin, and a root $y = 1$. Equation (B) is the same as (A) except for this root and the pole.

We now have the following:

(1) Every equation of the form

$$F_{NM}(y)=\sum_{n=M}^{N} a_n y^n-\sum_{n=0}^{M-1} b_n y^n=0$$

where $M > 0$, and all a_n, b_n are positive, has one and only one positive root.

This can be proved by induction. Assume it is true for a polynomial of degree $N-1$. Then, it is true for F'_{NM}. It is clear that $F_{NM}(0) < 0$, that $F'_{NM}(0) < 0$, if $M > 1$, and that $F_{NM}(y) > 0$ for sufficiently large values of y. Thus there is at least one positive root and one positive y where F is minimum. The assumption of the induction proof tells that this minimum is the only possible extremal point of F_{NM} for positive y. Because the minimum must occur between $y = 0$ and a positive root, there cannot be more than one root. Further, it is clear that if $M = 0$, i.e., all b_n are absent, the equation cannot have any positive root. Thus an equation with $M = 1$ has no minimum, and there cannot be more than one root. This confirms the induction. Clearly, Eq. (B) is of this form. Thus this equation has a unique, positive root, y_0. We will now show the distribution of the remaining roots.

(2) Besides the roots y_0 and 1 [for (A)], there are $R - 1$ roots, complex or negative with absolute values less than $\min(y_0, 1)$, and $R - 1$ complex or negative roots with absolute values larger than $\max(y_0, 1)$.

This can be shown by consideration of argument variation for complex y in Eq. (A). First, we note that F_A attains large, positive values for very small and very large positive values of y. $F_A = 0$ has only two real roots at $y = 1$ and $y = y_0$, as shown above. It follows that F_A is negative between these y values.

Now, consider a circle in the complex y space with $y = \rho e^{i\varphi}$. The real part of $F_A(y)$ is then

$$\sum_{r=1}^{R} (w_r + w_{-r})(\rho^r \cos r\varphi - 1) = \operatorname{Re} F_A(\rho e^{i\varphi})$$

Clearly, for given ρ, this expression is largest when $\varphi = 0$ as all w_r are positive. If ρ is between 1 and y_0, this largest value is $F_A(\rho)$. According to what was said above, this value is negative. Thus the real part is always negative when φ is varying. When y goes around the corresponding circle in the complex plane, F_A cannot go around the origin (because then the real part must change sign). From the argument criterion then follows that there are as many zeros as poles inside the y circle. The pole at the origin of order R is counted R times, thus there are R zeros: the smallest of 1 and y_0 and $R - 1$ other, which can be negative or complex with absolute value smaller than ρ. Since this is true for all ρ between 1 and y_0, the $R - 1$ zeros must have absolute values smaller than $\min(1, y_0)$. There cannot be any zero with absolute value between 1 and y_0. Since (A) in total has $2R$ zeros, there must be $R - 1$ ones which are negative or complex with absolute values larger than $\max(1, y_0)$.

ACKNOWLEDGMENTS

I wish to thank Drs. O. Leimar, Stockholm, J. de la Rubia, Madrid, and J. Stucki, Bern, for remarks and discussions about the work; further Drs. L. Brenig, W. Horsthemke, and M. Malek-Mansour, Brussels, for correspondence and discussions about the mutual features of our works. Also, I owe my gratitude to Prof. M. Velarde and Dr. V. Fairén of the Autonomous University in Madrid for fruitful discussions and for their hospitality during my visit to their institute, at which time this work was initiated.

REFERENCES

1. G. Nicolis and I. Prigogine, *Self-organization in Nonequilibrium Systems* (Wiley, New York, 1977).
2. I. Oppenheimer, K. E. Shuler, and G. H. Weiss, *Physica (The Hague)* **88A**:191 (1977).
3. J. S. Turner, *Adv. Chem. Phys.* **29**:63 (1975).
4. W. Horsthemke, M. Malek-Mansour, and L. Brenig, *Z. Phys. B* **28**:135 (1977).
5. R. Kubo, K. Matsuo, and K. Kitahara, *J. Stat. Phys.* **9**:41 (1973).
6. R. Görtz and D. F. Walls, *Z. Phys. B* **25**:423 (1976).
7. W. Horsthemke and L. Brenig, *Z. Phys. B* **27**:341 (1977).
8. M. Malek-Mansour and G. Nicolis, *J. Stat. Phys.* **13**:197 (1975).
9. I. Procaccia and J. Ross, *J. Chem. Phys.* **67**:5565 (1977).
10. C. W. Gardiner and S. Chaturvedi, *J. Stat. Phys.* **17**:429 (1977).
11. T. G. Kurtz, *J. Appl. Prob.* **8**:344 (1971).
12. A. D. Ventsel and M. I. Freidlin, *Russ. Math. Surveys* **25**:1 (1970).
13. N. G. van Kampen, *Adv. Chem. Phys.* **34**:245 (1976).
14. F. Schlögl, *Z. Phys.* **253**:147 (1972).
15. H. Tomita, A. Ito, and H. Kidachi, *Prog. Theory. Phys.* **56**:786 (1976).
16. C. Blomberg, *Physica (The Hague)* **86A**:49 (1977).
17. O. Leimar, J. de la Rubia, and C. Blomberg, preprint (1979).
18. G. Nicolis and J. W. Turner, *Physica (The Hague)* **89A**:326 (1977).
19. H. A. Kramers, *Physica (The Hague)* **7**:284 (1940).
20. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
21. J. J. B. Jack, D. Noble, and R. W. Tsien, *Electric current flow in excitable cells* (Oxford University Press, Oxford, 1975).
22. F. Jähnig and P. H. Richter, *J. Chem. Phys.* **64**:4645 (1976).