Adiabatic elimination in stochastic systems. I. Formulation of methods and application to few-variable systems

C. W. Gardiner
Physics Department, University of Waikato, Hamilton, New Zealand
(Received 17 August 1983)

Adiabatic elimination of stochastic variables is formulated in a systematic and largely rigorous manner, and applied to a variety of examples. Ambiguities in naive methods are resolved, and it is shown that the resultant stochastic differential equation is not always to be interpreted as an Ito or as a Stratonovich stochastic differential equation. A class of nonlinear examples is also treated in which naive methods fail completely.

I. INTRODUCTION

There is a wide range of situations in theoretical physics in which a well-defined separation of time scales arises, where a class of variables varies on a time scale which is characteristically very much more rapid than the time scale of the remainder of the variables. When there is dissipation, it is possible for the fast variables to relax to a quasistationary state, in which the values of the fast variables follow the slow variables; as Haken would say, the slow variables "slave" the fast variables.

The correct formulation of this problem in stochastic systems is the subject of this paper. By a "correct" formulation, I mean a method by which fast variables may be eliminated from the equations of motion in some well-defined limit. In deterministic systems, the limit is normally quite simple, it is when one time scale is very much smaller than another. However, in stochastic systems, there is quite a variety of limits available, and results will depend on what limits are taken. The choice of a limit appropriate to a particular physical situation is a physical question, not a mathematical one—I do not wish to attack that question here, but merely to indicate a method by which one can clearly formulate limits, and derive results.

My aim in this paper is rather different from that of most recent works on adiabatic elimination, which are listed under Ref. 1. These works concentrate on developing systematic perturbation series which can be, in principle, computed to arbitrary accuracy. These series are normally applied to the problem of Brownian motion in a potential, and normally produce similar results. I do not wish, in general, to go beyond a lowest-order calculation, but I do wish to undertake adiabatic elimination in a wide class of problems, and to show that the correct treatment of noise in such elimination requires a modicum of care. That is, I suggest that before developing a full perturbation theory it is wise to look at the richness of the phenomena which arise even out of the lowest-order terms. In any perturbation theory there is always a measure of complexity associated with higher-order terms—it is a rare person who would go further than second order. As I show in Sec. II F, higher-order corrections yield non-Markov processes with initial-value dependencies, and therefore are very much less useful. However, in Ref. 2 I have shown how higher-order corrections can be computed by my methods, and have computed the corrections for the case of Brownian motion in a potential. Schuss[3] has also applied related methods to Brownian motion in his book.

This paper, then, proceeds as follows. The elimination method is developed in Sec. II: it is a reformulation of the projector methods introduced originally by Zwanzig[4] and Nakajima[5] and since developed by a wide variety of authors in numerous contexts, usually without any systematic analysis of the limits involved. Van Kampen[6] comments on these and, in fact, uses cumulant expansion techniques to elucidate the limits involved, and Chaturvedi and Shibata[7] have adapted projector techniques in such a way as to show their relation to those of van Kampen.[6]

Section II develops the method, but applies it only to a very simple problem: the white-noise limit of a differential equation driven by a nonwhite noise. This is a well-known problem, and has been definitively solved by Papanicolaou and Kohler.[8] The method used here is less rigorous but can be adapted easily to more complex problems.

In Sec. III a treatment of some simple "linear" stochastic adiabatic eliminations is given. In these, the variable to be eliminated occurs linearly in the equations, and the naive elimination procedure commonly employed is given by setting the time derivative of the fast variable equal to zero. This naive method gives a stochastic differential equation whose interpretation is ambiguous. We show that this ambiguity is a very real problem: Correct procedures always give a Fokker-Planck equation which is a possible interpretation of the naive method, but, depending on the particular problem being investigated, the result may be that of the Ito, Stratonovich, or indeed even other interpretations. Thus it is simply not true to say that the Stratonovich interpretation of a stochastic differential equation is the physical interpretation: A correct interpretation can only be given when the detailed mechanisms underlying the system are known.

Section IV treats "nonlinear" adiabatic elimination, in which the eliminated variable occurs nonlinearly in the equations. There is no naive elimination procedure available in this case but, subject to certain conditions, it can be shown that adiabatic elimination can proceed, and sensible answers developed.

II. THE ELIMINATION METHOD

We formulate our method on a well-studied problem, namely, the white-noise limit of a system with a fluctuating term. The most thorough treatment of this problem is that of Papanicolaou and Kohler.[8]

The system being studied can be written in the form of a differential equation

$$\frac{dx}{dt}=a(x)+\gamma b(x)\alpha_0(t/\gamma^2) \tag{2.1}$$

in which $\alpha_0(t)$ has zero mean, and is the solution of a stochastic differential equation

$$d\alpha_0=A(\alpha_0)dt+\sqrt{2B(\alpha_0)}dW(t)\ . \tag{2.2}$$

The limit $\gamma\to\infty$ results in an effective noise term in (2.1) whose correlation time goes to zero like γ^{-2}, and whose amplitude increases like γ. Thus the fluctuations increase and become very rapid. We can show that in this limit (2.1) becomes equivalent to the Stratonovich stochastic differential equation

$$(S)\ \ dx=a(x)dt+\sqrt{2D}\,b(x)dW(t)\ , \tag{2.3}$$

where

$$D=\int_0^\infty dt\langle\alpha_0(t)\alpha_0(0)\rangle_s\ . \tag{2.4}$$

Equation (2.3) is equivalent to the Ito stochastic differential equation

$$dx=[a(x)-Db(x)b'(x)]dt+\sqrt{2D}\,b(x)dW(t)\ . \tag{2.5}$$

The problem to be solved is now best restated in terms of the corresponding probability density function $p(x,\alpha,t)$, which obeys the Fokker-Planck equation

$$\frac{\partial p}{\partial t}=(\gamma^2L_1+\gamma L_2+L_3)p\ , \tag{2.6}$$

where

$$L_1=-\frac{\partial}{\partial\alpha}A(\alpha)+\frac{\partial^2}{\partial\alpha^2}B(\alpha)\ , \tag{2.7}$$

$$L_2=-\frac{\partial}{\partial x}b(x)\alpha\ , \tag{2.8}$$

$$L_3=-\frac{\partial}{\partial x}a(x)\ , \tag{2.9}$$

$$\alpha_0(t/\gamma^2)=\alpha(t)\ . \tag{2.10}$$

We want to show that the probability density function for x, $\hat{p}(x,t)$, defined by

$$\hat{p}(x,t)=\int d\alpha\,p(x,\alpha,t) \tag{2.11}$$

in the large γ limit, obeys the Fokker-Planck equation [equivalent to (2.3) and (2.5)]

$$\frac{\partial\hat{p}}{\partial t}=\left[-\frac{\partial}{\partial x}a(x)+D\frac{\partial}{\partial x}b(x)\frac{\partial}{\partial x}b(x)\right]\hat{p}\ . \tag{2.12}$$

A. Introduction of a projector

A projector is introduced in a relatively standard way. We assume a *unique* normalized stationary solution $p_s(\alpha)$ to the equation $L_1p_s=0$, and (as noted above) we *require*

$$\langle\alpha\rangle_s=\int d\alpha\,p_s(\alpha)=0\ . \tag{2.13}$$

A projector P is defined by

$$Pf(x,\alpha)=p_s(\alpha)\int d\alpha' f(x,\alpha')\ . \tag{2.14}$$

By definition, then, P projects any function $f(x,\alpha)$ onto the *null space* of L_1. There are then the following properties of P which (except for the last) are *essential* for the use of the method:

$$P^2=P\ , \tag{2.15a}$$

$$L_1P=PL_1=0\ , \tag{2.15b}$$

$$P=\lim_{t\to\infty}\exp(L_1t)\ , \tag{2.15c}$$

$$PL_2P=0\ , \tag{2.15d}$$

$$PL_3=L_3P\ . \tag{2.15e}$$

Properties (2.15a) and (2.15b) are trivial to show. The property (2.15c) follows if L_1 is a genuine Fokker-Planck operator, from the fact that $\exp(L_1t)f(x,\alpha)$ is a solution of the Fokker-Planck equation $\partial p/\partial t=L_1p$, which therefore approaches the stationary solution as $t\to\infty$. Equation (2.15d) is very important and follows from the requirment $\langle\alpha\rangle_s=0$ for

$$PL_2Pf(x,\alpha)=p_s(\alpha)\langle L_2\rangle_s\int d\alpha' f(\alpha',x)\ , \tag{2.16}$$

and the definition of L_2 [e.g., (2.8)] shows $\langle L_2\rangle_s=0$ if $\langle\alpha\rangle_s=0$. Property (2.15c) is trivial. The properties (2.15a)–(2.15d) will be used again and again in this paper, and are fundamental to the method.

B. Introduction of Laplace transform for asymptotic results

For the Laplace transform we use the notation

$$\tilde{f}(s)=\int_0^\infty dt\,e^{-st}f(t)\ . \tag{2.17}$$

The evolution equation (2.6) becomes

$$s\tilde{p}(s)-p(0)=(\gamma^2L_1+\gamma L_2+L_3)\tilde{p}(s)\ . \tag{2.18}$$

We now write

$$\tilde{v}(s)=P\tilde{p}(s)\ , \tag{2.19}$$

$$\tilde{w}(s)=(1-P)\tilde{p}(s) \tag{2.20}$$

so that $\tilde{v}(s)$ and $\tilde{w}(s)$ represent, respectively, components of $\tilde{p}(s)$ in the null space of L_1 and the complement of the null space of L_1. Using the properties (2.15) we readily deduce

$$s\tilde{v}(s)-v(0)=L_3\tilde{v}(s)+\gamma PL_2\tilde{w}(s) \tag{2.21}$$

and

$$s\tilde{w}(s)-w(0)=L_3\tilde{w}(s)+L_2\tilde{v}(s)+\gamma(1-P)L_2\tilde{w}(s)+\gamma^2L_1\tilde{w}(s)\ . \tag{2.22}$$

We solve (2.22) for $\tilde{w}(s)$, and substitute in (2.21) to find

$$s\tilde{v}(s)-v(0)=\{L_3+\gamma^2PL_2[s-\gamma^2L_1-\gamma(1-P)L_2-L_3]^{-1}L_2\}\tilde{v}(s)+\gamma PL_2[s-\gamma^2L_1-\gamma(1-P)L_2-L_3]^{-1}w(0)\ . \tag{2.23}$$

Equation (2.23) is the fundamental exact equation from which limits and asymptotic forms can be quite readily deduced—at least formally. More rigorous results have been given by Papanicolaou.[9]

C. The white-noise limit

We take the limit $\gamma\to\infty$ in (2.23) and readily obtain

$$s\tilde{v}(s)-v(0)=(L_3-PL_2L_1^{-1}L_2)\tilde{v}(s)\ . \tag{2.24}$$

The result shows a "drift" L_3, as well as a "noise" $-PL_2L_1{}^{-1}L_2$. This noise term is only well defined if $PL_2P=0$, because L_1^{-1} only exists in the complement of the null space of L_1, and $L_2\tilde{v}(s)$ is in this complement provided $PL_2\tilde{v}(s)=0$, which is satisfied, since

$$PL_2\tilde{v}(s)=PL_2P\tilde{p}(s)=0 \tag{2.25}$$

using (2.15d).

The term $PL_2L_1^{-1}L\tilde{v}(s)$ can be evaluated, because we note

$$\int_0^\infty dt\exp(L_1t)=-L_1^{-1}(1-P) \tag{2.26}$$

[by straightforward integration, and the use of (2.15c)]. Thus

$$\begin{aligned}PL_2L_1^{-1}L_2\tilde{v}(s)&=PL_2L_1^{-1}(1-P)L_2P\tilde{v}(s)\\&=-p_s(\alpha)\int d\alpha'\left[-\frac{\partial}{\partial x}b(x)\alpha'\right]\int_0^\infty dt\exp(L_1t)\left[-\frac{\partial}{\partial x}b(x)\alpha'\right]p_s(\alpha')\tilde{v}(s)\ .\end{aligned} \tag{2.27}$$

Note that $\exp(L_1t)\alpha'p_s(\alpha')$ is the solution of $\partial f/\partial t=L_1f$ with initial condition $f(\alpha')=\alpha'p_s(\alpha')$: This is simply

$$\int d\alpha''p(\alpha',t\mid\alpha'',0)\alpha''p_s(\alpha'')\ , \tag{2.28}$$

where $p(\alpha',t\mid\alpha'',0)$ is the conditional probability density, so that

$$PL_2L_1^{-1}L_2\tilde{v}(s)=-p_s(\alpha)\left[\frac{\partial}{\partial x}b(x)\frac{\partial}{\partial x}b(x)\right]\times\int_0^\infty dt\langle\alpha(t)\alpha(0)\rangle_s\tilde{v}(s)\ , \tag{2.29}$$

where

$$\langle\alpha(t)\alpha(0)\rangle=\int d\alpha'd\alpha''\alpha'\alpha''p(\alpha',t\mid\alpha'',0)p_s(\alpha'') \tag{2.30}$$

is the stationary autocorrelation function of $\alpha(t)$. Inverting the Laplace transform, we find $\hat{p}(x,t)$ obeys Eq. (2.12).

D. Markov property of the resulting equation

In the $\gamma\to\infty$ limit, for any choice of L_1, L_2, and L_3 [satisfying the conditions (2.15)] the initial value of $w(0)$ does not appear. This means that (2.24) gives rise to a Markov process whose distribution function is $\hat{p}(x,t)$. [Papanicolaou and Kohler[8] show that the Markov property arises even if $\alpha(t)$ does not form a Markov process, but satisfies only a weaker condition of being "strongly mixing." However, this stronger result does not generalize to the more interesting problems which will be considered later, since it does not provide an equation of motion for $\alpha(t)$.]

In this limit

$$\tilde{w}(s)\sim-\gamma^{-1}L_1^{-1}L_2\tilde{v}(s)\to0 \tag{2.31}$$

so that

$$p(x,\alpha,t)\sim v(x,\alpha,t)\sim p_s(\alpha)\hat{p}(x,t)\ , \tag{2.32}$$

where $\hat{p}(x,t)$ satisfies Eq. (2.12), which is a Fokker-Planck equation, involving only the x variable. This means that we can define a reduced conditional probability, $\hat{p}(x,t\mid x',t')$ which is a solution to the Fokker-Planck equation (2.12) subject to the initial condition

$$\hat{p}(x,t'\mid x',t')=\delta(x-x')\ . \tag{2.33}$$

In what way is this reduced conditional probability related to the full conditional probability $p(x,\alpha,t\mid x',\alpha',t')$? We first note that there are many solutions to Eq. (2.6), corresponding to different initial conditions. The result (2.24) gives an approximate solution

$$\hat{p}(x,t\mid x',t')p_s(\alpha) \tag{2.34}$$

which becomes

$$\delta(x-x')p_s(\alpha)\ \text{ when } t=t' \tag{2.35}$$

and the initial condition (2.35) is the same as that for

$$\int d\alpha'p(x,\alpha,t\mid x',\alpha',t')p_s(\alpha')\ , \tag{2.36}$$

which is the conditional probability for x and α, at time t, given x' and that α' is distributed over the stationary distribution $p_s(\alpha')$ at time t'. This means that $\hat{p}(x,t\mid x',t')$ is obtained from $p(x,\alpha,t\mid x',\alpha',t')$ by summing over the final value α and taking the stationary average over the initial value α', i.e.,

$$\int\int d\alpha\,d\alpha'\hat{p}(x,\alpha,t\mid x',\alpha',t')p_s(\alpha')\to\hat{p}(x,t\mid x',t')\ \text{ as }\gamma\to\infty\ . \tag{2.37}$$

E. Choice of the projector

The projection method used here has a strong resemblance to that used by many authors, which was originally developed by Żwanzig[4] in a statistical mechanics context, and has since been widely used in quantum optics.[10] In contrast to these more physical applications, in this method the projector is not arbitrary. All of the conditions (2.15) must be satisfied, for the following reasons.

(a) If condition (2.15b) is not satisfied, the equations (2.21) and (2.22) are altered in such a way that no $\gamma \to \infty$ limit exists.

(b) Condition (2.15c) follows automatically from (2.15b) if $p_s(\alpha)$ is unique, in other words, there is only one solution for P which satisfies (2.15b) if $p_s(\alpha)$ is unique.

(c) The condition (2.15d) arises as an essential prerequisite for the limit $\gamma \to \infty$ to exist, for it says that L_2 acting in the null space of L_1 gives a result not in the null space of L_1. Thus $L_1^{-1}L_2P$ exists, and this means that (2.24) exists.

F. Alternative dependence on γ

Slightly less elegant asymptotic results arise if we consider the equation

$$\partial_t p = (\gamma^2 L_1 + \gamma^r L_2 + L_3)p \ , \quad 2 > r > 0 \ . \tag{2.38}$$

The reduced equation $\hat{p}$ is

$$\partial_t \hat{p} = (L_3 - \gamma^{2r-2} P L_2 L_1^{-1} L_2)\hat{p} \ . \tag{2.39}$$

If $r = 0$, the noise term $PL_2L_1^{-1}L_2$ is of the same order of magnitude as the term $w(0)$, whose neglect is only justifiable if the $PL_2L_1^{-1}L_2$ term is also neglected. By *not* neglecting either term, we obtain the equation

$$s\tilde{v}(s) - [v(0) + \gamma^{-2} P L_2 L_1^{-1} w(0)] = [L_3 - \gamma^{-2} P L_2 L_1^{-1} L_2]\tilde{v}(s) \ . \tag{2.40}$$

This is of the same form as the equation for any r, except that the initial condition on $v(t)$ is no longer $v(0)$, but instead the term in the square brackets on the left-hand side of (2.40). The process is no longer a Markov process in $\tilde{v}(s)$ to this order, since $v(t)$ is not determined solely by a knowledge of $v(0)$, but requires knowledge of $w(0)$ as well.

G. Non-Markov property of higher-order terms

Any perturbation expansion of (2.23) to higher order will yield terms of the same order of magnitude as the terms in $w(0)$, which are of order γ^{-1}. Thus the equation involves the full knowledge of initial conditions, and no longer represents a Markov process to this order. Useful results can be obtained in such perturbation expansions [e.g., see Refs. 1, 2, and 7] but care must be taken not to assume the Markov property.

III. ADIABATIC ELIMINATION OF VARIABLES WHICH APPEAR LINEARLY

The example of Sec. II contains all the symbolic formalism in terms of projectors, etc., to be used in a complete theory of adiabatic elimination. However, in the example chosen, the motion of $\alpha(t)$ is quite independent of x, and this is a rather special case.

In this section we treat a class of systems in which $\alpha(t)$ occurs only linearly, and in which a naive elimination procedure is commonly used, which consists of a setting $d\alpha/dt$ equal to zero. Thus in this section the eliminated variable and the remaining variable both influence each others motion. The motion of α is fast compared to that of x, so that $\alpha(t)$ is effectively a stationary process which depends on x. This is the classical meaning of the concept of an "adiabatic approximation."

We now present three models, in which we find that the resulting stochastic differential equation, though closely related to a "naive adiabatic elimination," may, in turn, be a Stratonovich, an Ito, or yet another interpretation of the white-noise stochastic differential equation.

A. Haken's slaving model (Ref. 11)

The deterministic version of the model is a pair of coupled equations, which may be written

$$\dot{x} = -\epsilon x - a x \alpha \ , \tag{3.1}$$

$$\dot{\alpha} = -\kappa\alpha + b x^2 \ . \tag{3.2}$$

One assumes that if κ is sufficiently large, α will relax rapidly to a quasistationary value given by setting $\dot{\alpha} = 0$ in (3.2) (this procedure is quite easy to justify rigorously). Thus, we may replace α by the stationary situation of (3.2) to obtain

$$\alpha = \frac{b}{\kappa} x^2 \ , \tag{3.3}$$

$$\dot{x} = -\epsilon x - \frac{ab}{\kappa} x^3 \ . \tag{3.4}$$

The result of the elimination of α is a single differential equation for x.

In a stochastic version of this model there are various possibilities available. The usual condition for the validity of adiabatic elimination in a deterministic system is

$$\epsilon \ll \kappa \ . \tag{3.5}$$

In a stochastic version, all other parameters come into play as well, and the condition (3.5) is able to be realized in different ways, with characteristically different answers.

Let us write stochastic versions of (3.1) and (3.2) by simply adding on noise terms:

$$dx = -(\epsilon x + a x \alpha)dt + C\, dW_1(t) \ , \tag{3.6}$$

$$d\alpha = (-\kappa\alpha + b x^2)dt + D\, dW_2(t) \ . \tag{3.7}$$

We assume here, for simplicity, that C and D are constants, and $W_1(t)$ and $W_2(t)$ are independent of each other.

The Fokker-Planck equation is

$$\frac{\partial p}{\partial t} = \left[\frac{\partial}{\partial x}(\epsilon x + a x \alpha) + \tfrac{1}{2} C^2 \frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial \alpha}(\kappa\alpha - b x^2) + \tfrac{1}{2} D^2 \frac{\partial^2}{\partial \alpha^2} \right] p \ . \tag{3.8}$$

We wish to eliminate α. It is convenient to define a new variable β by

$$\beta=\alpha-\frac{b}{\kappa}x^2 \tag{3.9}$$

so that, for fixed x, the quantity β has zero mean. In terms of this variable, we can write a Fokker-Planck equation

$$\frac{\partial p}{\partial t}=(L_1^0+L_2^0+L_3^0)p\ , \tag{3.10}$$

$$L_1^0=\frac{\partial}{\partial\beta}\kappa\beta+\frac{D^2}{2}\frac{\partial^2}{\partial\beta^2}\ , \tag{3.11}$$

$$L_2^0=\frac{\partial}{\partial x}a\beta x-\frac{2bx}{\kappa}\frac{\partial}{\partial\beta}\left[\epsilon x+\frac{ab}{\kappa}x^3+ax\beta\right]$$
$$-C^2\left[\frac{bx}{\kappa}\frac{\partial^2}{\partial x\partial\beta}+\frac{\partial^2}{\partial x\partial\beta}\frac{bx}{\kappa}\right]+\frac{2b^2x^2C^2}{\kappa^2}\frac{\partial^2}{\partial\beta^2}\ , \tag{3.12}$$

$$L_3^0=\frac{\partial}{\partial x}\left[\epsilon x+\frac{ab}{\kappa}x^3\right]+\frac{C^2}{2}\frac{\partial^2}{\partial x^2}\ . \tag{3.13}$$

In terms of these variables, the limit $\epsilon\to 0$ is not interesting, since we simply get the same system with $\epsilon=0$—no *elimination* is possible, since L_1 is not multiplied by a large parameter.

In order for the limit $\epsilon\to 0$ to have the meaning deterministically that (3.4) is a valid *limiting form,* there must be a quantity A such that

$$\frac{ab}{\kappa}=\epsilon A \quad \text{as } \epsilon\to 0 \tag{3.14}$$

and for this limit to be recognizable deterministically, it must not be swamped by noise, so one must also have (for some B)

$$\frac{C^2}{2}=\epsilon B \quad \text{as } \epsilon\to 0 \tag{3.15}$$

which means, as $\epsilon\to 0$,

$$L_3^0\to\epsilon\left[\frac{\partial}{\partial x}(x+Ax^3)+B\frac{\partial^2}{\partial x^2}\right]\ . \tag{3.16}$$

However, there are two distinct possibilities for L_2^0. In order for L_1^0 to be independent of ϵ, we must have κ independent of ϵ, which is reasonable. Thus, the limit (3.14) must be achieved by the product ab being proportional to ϵ. We consider two possibilities.

(i) *The silent slave: a proportional to* ϵ. We first consider a case where we can write

$$a=\epsilon\tilde{a}\ . \tag{3.17}$$

We see that L_1^0 is independent of ϵ, while L_2^0 and L_3^0 are proportional to ϵ. If we rescale time by

$$\tau=\epsilon t \tag{3.18}$$

then

$$\frac{\partial p}{\partial\tau}=\left[\frac{1}{\epsilon}L_1+L_2+L_3\right]p\ , \tag{3.19}$$

where

$$L_1=L_1^0\ ,\quad L_2=L_2^0/\epsilon\ ,\quad L_3=L_3^0/\epsilon\ . \tag{3.20}$$

The equation (3.19) is now in a form similar to (2.6), with the exception that the coefficient of L_2 does not become infinite as $\epsilon\to 0$. The $PL_2L_1^{-1}L_2$ term is therefore negligible compared to terms arising from L_3, and the limiting equation is

$$\frac{\partial\hat{p}}{\partial t}=L_3\hat{p}=\left[\frac{\partial}{\partial x}(x+Ax^3)+B\frac{\partial^2}{\partial x^2}\right]\hat{p}\ . \tag{3.21}$$

This corresponds exactly to eliminating α adiabatically, ignoring the fluctuations in α, and simply setting the deterministic value in the x equation. I call it the "silent slave," since (in Haken's terminology) α is slaved by x, and makes no contribution to the noise in the x equation. This is the usual form of slaving, as considered by Haken.

(ii) *The noisy slave: a proportional to* $\epsilon^{1/2}$. We now consider a case where both a and b are proportional to $\epsilon^{1/2}$ we can write

$$a=\tilde{a}\epsilon^{1/2}\ ,\quad b=\tilde{b}\epsilon^{1/2}\ , \tag{3.22}$$

where

$$\tilde{a}\tilde{b}=\kappa A\ . \tag{3.23}$$

L_1^0 stays constant, L_3^0 is proportional to ϵ, and

$$L_2^0=\epsilon^{1/2}L_2+\cdots\ ,$$

where the ellipses stand for higher-order terms in ϵ and

$$L_2=\tilde{a}\beta\frac{\partial}{\partial x}x\ . \tag{3.24}$$

Thus the limiting equation is (following a procedure similar to that in Sec. III B)

$$\frac{\partial\hat{p}}{\partial\tau}=(L_3-PL_2L_1^{-1}L_2)\hat{p}\ . \tag{3.25}$$

The term $PL_2L_1^{-1}L_2$ can be worked out by noting that $L_1=L_1^0$ as defined by Eq. (3.11) corresponds to an Ornstein-Uhlenbeck process, for which the stationary autocorrelation function is given by

$$\langle\beta(t),\beta(t')\rangle=\frac{D^2}{2\kappa}\exp(-\kappa|t-t'|) \tag{3.26}$$

and following the procedure of Sec. II, we find

$$-PL_2L_1^{-1}L_2=\tilde{a}^2\frac{D^2}{2\kappa^2}\frac{\partial}{\partial x}\left[x\left[\frac{\partial}{\partial x}xp_s(\beta)\right]\right] \tag{3.27}$$

so

$$\frac{\partial\hat{p}}{\partial\tau}=\left\{\frac{\partial}{\partial x}\left[x\left[1-\frac{\tilde{a}^2D^2}{2\kappa^2}\right]+Ax^3\right]+\frac{\partial^2}{\partial x^2}\left[B+\frac{\tilde{a}^2D^2x^2}{2\kappa^2}\right]\right\}\hat{p}\ . \tag{3.28}$$

I call this the "noisy slave," since the slave makes its presence felt in the final equation by adding noise (and affecting the drift, though this appears only in the Ito form as written—as a Stratonovich form, there would be no extra drift).

(iii) *Comment on naive elimination.* We could write, for case (ii) given by (3.22), ordinary Langevin equations equivalent to (3.7),

$$\dot{x}=-\epsilon x+\tilde{a}\epsilon^{1/2}x\alpha+(2\epsilon B)^{1/2}\xi_1(t)\ , \tag{3.29}$$

$$\dot{\alpha}=-\kappa\alpha+\tilde{b}\epsilon^{1/2}x^2+D\xi_2(t)\ , \tag{3.30}$$

and eliminate α by substituting the "stationary solution" of (3.30) obtained by setting $\dot{\alpha}=0$,

$$\alpha=\left[\frac{\tilde{b}\epsilon^{1/2}}{\kappa}\right]x^2+\frac{D}{\kappa}\xi_2(t)$$

in Eq. (3.29), to obtain

$$\dot{x}=-\epsilon x+\epsilon\frac{\tilde{a}\tilde{b}}{\kappa}x^3+\epsilon^{1/2}\left[(2B^{1/2})\xi_1(t)+\frac{\tilde{a}D}{\kappa}x\xi_2(t)\right]\ . \tag{3.31}$$

This is obviously *related* to the Fokker-Planck equation (3.28), but the precise interpretation rule for the nonlinear noise $x\xi_2(t)$ is not given. The method given here indicates that in this case, each noise is to be interpreted as an independent *Stratonovich* noise. However, this is not always the case, as is shown in the next example.

B. A case which yields an Ito equation

We consider the large-γ limit of

$$\begin{aligned}dx&=p\,dt\ ,\\ dp&=-\gamma[p-f(x)]+\gamma g(x)dW(t)\ .\end{aligned} \tag{3.32}$$

There is no Ito-Stratonovich ambiguity in this case, since the p equation depends only on x. We define a new variable

$$\alpha=\gamma^{-1/2}[p-f(x)] \tag{3.33}$$

so that the Eqs. (3.32) become

$$\begin{aligned}dx&=\gamma^{1/2}\alpha\,dt+f(x)dt\ ,\\ d\alpha&=-\gamma\alpha+\gamma^{1/2}g(x)dW(t)\\ &\quad+\gamma^{-1/2}f'(x)[\gamma^{1/2}\alpha\,dt+f(x)dt]\end{aligned} \tag{3.34}$$

so we find

$$\begin{aligned}\gamma L_1&=\gamma\left[\frac{\partial}{\partial\alpha}+\tfrac{1}{2}g(x)^2\frac{\partial^2}{\partial\alpha^2}\right]\ ,\\ \gamma^{1/2}L_2(\gamma)&=-\gamma^{1/2}\frac{\partial}{\partial x}\alpha\\ &\quad-\frac{\partial}{\partial\alpha}[f'(x)\alpha+f(x)f'(x)\gamma^{-1/2}]\ ,\\ L_3&=-\frac{\partial}{\partial x}f(x)\ .\end{aligned} \tag{3.35}$$

We note the following.

(i) Lower-order terms in $L(\gamma)$ will not contribute to the asymptotic result for large γ.

(ii) Because L_1 depends on x, $L_3P\neq PL_3$. Hence instead of (2.23) we obtain [setting $w(0)=0$]

$$s\tilde{v}(s)-v(0)=PL_3\tilde{v}(s)+P[\gamma^{1/2}L_2(\gamma)+L_3][s-\gamma L_1-\gamma^{1/2}(1-P)L_2(\gamma)-(1-P)L_3]^{-1}\times[\gamma^{1/2}L_2(\gamma)+(1-P)L_3]\tilde{v}(s) \tag{3.36}$$

and in the large-γ limit, Eq. (3.36) approaches

$$s\tilde{v}(s)-v(0)=PL_3\tilde{v}(s)-PL_2L_1^{-1}L_2\tilde{v}(s)\ , \tag{3.37}$$

where

$$L_2=\lim_{\gamma\to\infty}L_2(\gamma)=-\alpha\frac{\partial}{\partial x}\ .$$

In this case, and in other situations in which L_1 depends on x, it is technically a little simpler to carry out the process of computation of $PL_2L_1^{-1}L_2$ using the *backward* or adjoint Fokker-Planck operators. These are described in Ref. 2. For the system described in (3.32) the backward operators are

$$\begin{aligned}L_1^*&=-\alpha\frac{\partial}{\partial\alpha}+\tfrac{1}{2}g(x)^2\frac{\partial^2}{\partial\alpha^2}\ ,\\ L_2^*&=\alpha\frac{\partial}{\partial x}\ ,\\ L_3^*&=f(x)\frac{\partial}{\partial x}\ ,\end{aligned} \tag{3.38}$$

and the projector becomes P^*, defined by

$$(P^*f)(\alpha,x)=\int d\alpha\,p_s(\alpha)f(\alpha,x)\ . \tag{3.39}$$

We denote the solutions of the backward equation by $q(\alpha,x)$, and define

$$\hat{q}(x)=\int d\alpha\,p_s(\alpha)q(\alpha,x)\ . \tag{3.40}$$

The equation for $\hat{q}(x)$ will be

$$\frac{\partial \hat{q}}{\partial t}=[P^*L_3^*-P^*L_2^*(L_1^*)^{-1}L_2^*]\hat{q} \tag{3.41}$$

and

$$-P^*L_2^*(L_1^*)^{-1}L_2^*\hat{q} = -\int d\alpha\, p_s(\alpha)\alpha\frac{\partial}{\partial x}\left[(L_1^*)^{-1}\alpha\frac{\partial}{\partial x}\hat{q}(x)\right] . \tag{3.42}$$

From (3.38) it is clear that α is the eigenfunction of L_1^* with eigenvalue -1, hence Eq. (3.42) equals

$$\int d\alpha\, p_s(\alpha)\alpha^2\frac{\partial^2}{\partial x^2}\hat{q}(x)=\frac{g(x)^2}{2}\frac{\partial^2}{\partial x^2}\hat{q}(x) \tag{3.43}$$

and hence

$$\frac{\partial \hat{q}}{\partial t}=f(x)\frac{\partial \hat{q}}{\partial x}+\tfrac{1}{2}g(x)^2\frac{\partial^2}{\partial x^2}\hat{q}(x) \tag{3.44}$$

which is the backward Fokker-Planck equation corresponding to the Ito stochastic differential equation

$$\dot{x}=f(x)+g(x)\xi(t) , \tag{3.45}$$

which would result from naive elimination methods. The basic difference between this example and the previous one is where the x dependence of the noise arises. In the previous example it arose because of an x multiplying the eliminated variable in the x equation, here the x dependence arises because of a dependence on x of the noise term in the equation for the eliminated variable p.

C. Examples which give neither Ito nor Stratonovich equations

Using the equations

$$\begin{aligned} dx &= h(x)p\,dt ,\\ dp &= -\gamma\left[p-\frac{f(x)}{h(x)}\right]+\gamma g(x)dW(t) \end{aligned} \tag{3.46}$$

will similarly produce a backward Fokker-Planck equation

$$\frac{\partial \hat{q}}{\partial t}=f(x)\frac{\partial}{\partial x}\hat{q}(x)+\tfrac{1}{2}g(x)^2h(x)\frac{\partial}{\partial x}h(x)\frac{\partial}{\partial x}\hat{q}(x) \tag{3.47}$$

which is clearly neither the Ito nor the Stratonovich equation for the naive result,

$$\dot{x}=f(x)+g(x)h(x)\xi(t) , \tag{3.48}$$

but rather, a mixture of the two kinds of equation.

D. Summary

This section has demonstrated that naive adiabatic elimination is of strictly limited value—that even in systems where no Ito-Stratonovich ambiguity exists, one cannot automatically predict what kind of stochastic differential equation interpretation is to be used.

IV. NONLINEAR SITUATIONS

The examples in Sec. III were all similar in one respect: The eliminated variable (or a variable linearly related to it) corresponds to an Ornstein-Uhlenbeck process, and in every case, a naive adiabatic elimination corresponded quite closely to the correct result; the main difficulty lying in the resolution of ambiguities of the Ito-Stratonovich or similar types.

In genuinely nonlinear situations the situation is very much more complex. The result of a naive elimination can be quite inscrutable. For example, instead of Haken's model [(3.6) and (3.7)] we could consider the pair

$$dx=-(\epsilon x+ax\alpha)dt+C\,dW_1(t) , \tag{4.1}$$

$$d\alpha=(-\kappa\alpha-\mu\alpha^3+bx^2)dt+D\,dW_2(t) . \tag{4.2}$$

The stationary solution for α involves the solution of a cubic equation

$$\kappa\alpha+\mu\alpha^3=bx^2-D\xi_2(t) \tag{4.3}$$

whose solution for α is a nonlinear function of $\xi_2(t)$; a function whose meaning is not defined. Thus, naive adiabatic elimination is no longer feasible even for giving a rough idea of the eliminated equation, and the more abstract formulation in terms of operators is quite essential.

A. A class of nonlinear examples

We consider the pair of equations

$$\begin{aligned} dx &= \gamma b(x)\alpha\,dt ,\\ d\alpha &= -\gamma^2A(x,\alpha,\gamma)dt+\gamma\sqrt{2B(x,\alpha,\gamma)}\,dW(t) , \end{aligned} \tag{4.4}$$

and we assume the existence of the following limits and asymptotic expansions:

$$\begin{aligned} A(x,\alpha,\gamma) &\sim \sum_{n=0}^{\infty}A_n(x,\alpha)\gamma^{-n} ,\\ B(x,\alpha,\gamma) &\sim \sum_{n=0}^{\infty}B_n(x,\alpha)\gamma^{-n} , \end{aligned} \tag{4.5}$$

and these expansions imply that there is an asymptotic stationary distribution of α at fixed x given by

$$p_s(\alpha,x)=\lim_{\gamma\to\infty}p_s(\alpha,x,\gamma) , \tag{4.6}$$

$$p_s(\alpha,x)\propto B_0(x,\alpha)^{-1}\exp\int d\alpha[A_0(x,\alpha)/B_0(x,\alpha)] . \tag{4.7}$$

We assume that $A_0(x,\alpha)$ and $B_0(x,\alpha)$ are such that

$$\langle\alpha(x)\rangle_s=\int d\alpha\,\alpha p_s(\alpha,x)=0 \tag{4.8}$$

so that we deduce from (4.5) that, for finite γ

$$\langle\alpha(x,\gamma)\rangle_s=\int d\alpha\,\alpha p_s(\alpha,x,\gamma)\sim\alpha_0(x)\gamma^{-1} , \tag{4.9}$$

where $\alpha_0(x)$ can be determined from (4.5).

We define the new variables

$$\beta=\alpha-\frac{1}{\gamma}\alpha_0(x)\ ,\quad x_1=x\ , \tag{4.10}$$

in terms of which the Fokker-Planck operator becomes (the Jacobian is a constant, as usual) on changing x_1 back to x

$$L=-\frac{\partial}{\partial x}[\alpha_0(x)b(x)]-\gamma\beta\frac{\partial}{\partial x}b(x)+\frac{1}{\gamma}\alpha_0'(x)\alpha_0(x)b(x)\frac{\partial}{\partial\beta}+\frac{\partial}{\partial\beta}\{\beta[\alpha_0'(x)b(x)]\}$$
$$+\gamma^2\left\{\frac{\partial}{\partial\beta}\left[A\left[\frac{\alpha_0(x)}{\gamma}+\beta,x,\gamma\right]\right]+\frac{\partial^2}{\partial\beta^2}\left[B\left[\frac{\alpha_0(x)}{\gamma}+\beta,x,\gamma\right]\right]\right\} \tag{4.11}$$

and, by using the asymptotic expansions (4.5) we can write as

$$L=L_3+\gamma L_2(\gamma)+\gamma^2L_1 \tag{4.12}$$

with

$$L_3=-\frac{\partial}{\partial x}\alpha_0(x)b(x)\ , \tag{4.13}$$

$$L_1=\frac{\partial}{\partial\beta}A_0(\beta,x)+\frac{\partial^2}{\partial\beta^2}B_0(\beta,x)\ , \tag{4.14}$$

$$L_2(\gamma)=L_2+O(\gamma^{-1})\ , \tag{4.15}$$

$$L_2=-\beta\frac{\partial}{\partial x}b(x)-\frac{\partial}{\partial\beta}\left[\frac{\partial A_0(\beta,x)}{\partial\beta}\alpha_0(x)+A_1(\beta,x)\right]-\frac{\partial^2}{\partial\beta^2}\left[\frac{\partial B_0(\beta,x)}{\partial\beta}\alpha_0(x)+B_1(\beta,x)\right]\ . \tag{4.16}$$

We note that L_3 and L_1 do not commute, but this does not affect the limiting result, which takes the usual form

$$\frac{\partial\hat{p}}{\partial t}=(L_3-PL_2L_1^{-1}L_2)\hat{p}\ . \tag{4.17}$$

The evaluation of the $PL_2L_1^{-1}L_2$ term is straightforward, but messy. We note that the terms involving $\partial/\partial\beta$ vanish after being operated on by P. From the explicit form of $p_s(\alpha,x)$ one can write

$$\left[\frac{\partial}{\partial\beta}\left[\frac{\partial A_0(\beta,x)}{\partial\beta}\alpha_0(x)+A_1(\beta,x)\right]+\frac{\partial^2}{\partial\beta^2}\left[\frac{\partial B_0(\beta,x)}{\partial\beta}\alpha_0(x)+B_1(\beta,x)\right]\right]p_s(\beta,x)=G(\beta,x)p_s(\beta,x) \tag{4.18}$$

and one finds that

$$PL_2L_1^{-1}L_2\hat{p}=\left[\frac{\partial}{\partial x}b(x)D(x)\frac{\partial}{\partial x}+\frac{\partial}{\partial x}b(x)E(x)\right]\hat{p} \tag{4.19}$$

with

$$D(x)=\int_0^\infty dt\langle\beta(t),\beta(0)\,|\,x\rangle\ ,\quad E(x)=\int_0^\infty dt\langle\beta(t),G(\beta,x)\,|\,x\rangle\ , \tag{4.20}$$

where $\langle\cdots|x\rangle$ indicates an average over $p_s(\beta,x)$. This is a rather strong adiabatic elimination result, in which an arbitrary nonlinear elimination can be handled and a finite resulting noise dealt with.

V. CONCLUSIONS

In this paper we have developed a completely practical method of carrying out adiabatic elimination in a range of few variable stochastic systems, which is reasonably rigorous, is simple, and, above all, is necessary if the *right* answer is to be obtained, as is demonstrated by the three models of Sec. III, or if *any* answer is to be obtained, as is demonstrated in Sec. IV. The basic idea of the method is to compute limits of the terms in Eq. (2.23) as $\gamma\to\infty$. Obviously any physical situation is presented to us with the parameters defined: Physical insight must be used to determine which parameters are large, i.e., which variables are to be eliminated, and hence to compute what scaling and redefining of variable is necessary to bring the Fokker-Planck equations into a form suitable for application of the method.

In the following papers the method will be applied to the problem of coarse-graining in reaction diffusion systems, to the derivation of stochastic amplitude equations in hydrodynamic instabilities, and finally to derivation of blocking equations in renormalization-group theory. In all of these we will eliminate infinite numbers of variables, using basically the same methods as developed here.

ACKNOWLEDGMENTS

I have greatly benefited from conversations with Pierre Hohenberg, Robert Graham, Bernard Matkowsky, Ze'ev Schuss, and Moira Steyn-Ross. The work was supported by a visit to the Institute for Theoretical Physics at Santa Barbara, under National Science Foundation Grant No. PHY-77-27084, and by the U. S. Army through its European Research Office.

[1]U. M. Titulaer, Physica (Utrecht) 31A, 321 (1978); 100A, 251 (1980); J. L. Skinner and P. G. Woyles, *ibid.* 96A, 561 (1979); S. Chaturvedi and F. Shibata, Z. Phys. B 35, 297 (1979); M. San Miguel and J. M. Sancho, J. Stat. Phys. 22, 605 (1980); H. Risken, D. H. Vollmer, and M. Mörsch, Z. Phys. B 40, 343 (1981); F. Haake, *ibid.* 48, 31 (1982); U. M. Titulaer, *ibid.* 50, 71 (1983).

[2]C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, Heidelberg, New York, 1983).

[3]Z. Schuss, *Theory and Applications of Stochastic Differential Equations* (Wiley, New York, 1980).

[4]R. Zwanzig, J. Chem. Phys. 33, 1338 (1960).

[5]S. Nakajima, Prog. Theor. Phys. 20, 948 (1958).

[6]N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981), p. 397.

[7]S. Chaturvedi and F. Shibata, Z. Phys. B 35, 297 (1979).

[8]G. C. Papanicolaou and W. Kohler, Commun. Pure Appl. Math. 27, 641 (1974).

[9]G. C. Papanicolaou, in *Modern Modelling of Continuum Phenomena,* Vol. 16 of *Lecture Notes in Applied Maths* (American Mathematical Society, Providence, Rhode Island, 1977), p. 109.

[10]F. Haake, Springer Tracts in Mod. Phys. 66, 98 (1973).

[11]H. Haken, *Synergetics—an Introduction,* 2nd ed. (Springer, Heidelberg, Berlin, New York, 1978), Chap. 7.