Journal of Statistical Physics, Vol. 112, Nos. 1/2, July 2003 (© 2003)

Elimination of Fast Chaotic Degrees of Freedom: On the Accuracy of the Born Approximation

Wolfram Just,[1,2] Katrin Gelfert,[3] Nilüfer Baba,[3] Anja Riegert,[3] and Holger Kantz[3]

Received July 29, 2002; accepted February 11, 2003

We apply standard projection operator techniques known from nonequilibrium statistical mechanics to eliminate fast chaotic degrees of freedom in a low-dimensional dynamical system. Through the usual perturbative approach we end up in second order with a stochastic system where the fast chaotic degrees of freedom are modelled by Gaussian white noise. The accuracy of the perturbation expansion is analysed in detail by the discussion of an exactly solvable model.

KEY WORDS: Elimination of fast variables; projection operator techniques; Fokker–Planck equation.

1. INTRODUCTION

Elimination of fast degrees of freedom is one of the classical and central issues of nonequilibrium statistical physics. There exist numerous concepts to model the influence of a deterministic thermodynamic heat bath by effective stochastic forces or noise, resulting either in Langevin equations or stochastic differential equations, in Fokker–Planck equations, in Master equations or Boltzmann equations and so on. The thermodynamic properties of the heat bath, in particular its thermodynamic limit, plays a double role. On the one hand the limit guarantees the decay of correlations of bath variables. On the other hand, by a variant of the law of large numbers the

[1] Institute of Physics, Chemnitz University of Technology, D-09107 Chemnitz, Germany; e-mail: wolfram.just@physik.tu-chemnitz.de

[2] Permanent address: Department of Mathematics, Queen Mary, University of London, Mile End Road, London E1 4NS, United Kingdom.

[3] Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, D-01187 Dresden, Germany.

0022-4715/03/0700-0277/0

statistics of the heat bath and in particular of the stochastic forces becomes Gaussian. Furthermore, fluctuation dissipation relations display on the level of the slow variables the Hamiltonian structure of the underlying microscopic dynamics.

Problems which are related with the dynamics on different time scales are known for centuries, e.g., in celestial mechanics. Thus time scale separation and elimination of fast variables are a central issue in quite different contexts,[1] e.g., for the investigation of instabilities in physical, chemical or biological systems.[2] It is a common feature of all these examples that one typically has a finite number of fast degrees of freedom and time scale separation comes through the values of the parameters of the system. Depending on the type of the fast motion one can group elimination schemes which are qualitatively different.

If the fast motion is purely relaxatory then elimination of the fast degrees of freedom is usually called adiabatic elimination.[3] In formal terms the full system has a slow invariant manifold and one obtains a deterministic effective equation of motion for the slow variables. This case is very well understood even from a rigorous mathematical point of view (cf., e.g., refs. 4 and 5). When the motion of the fast variable is periodic in time then the elimination is performed by averaging the fast degrees of motion. Again one obtains a deterministic effective equation of motion for the slow part of the dynamics.[4, 6] Here we consider the third possibility where the slow degrees of freedom are coupled to a finite number of fast chaotic modes and where no manifold reduction can be applied.

Such problems arise for instance in climate research[7] or molecular dynamics. For numerical purposes in long time simulations, but also for principle reasons, the elimination of the fast modes is often highly desirable. Already a finite number of fast chaotic modes share with thermodynamic heat baths the decay of correlations so that the reasoning sketched above indicates that fast chaotic degrees of freedom can be modelled by suitable stochastic processes. Of course one should not expect a Gaussian statistics from the very beginning since nothing comparable to the law of large numbers is available in a system with a finite number of degrees of freedom. On the other hand, a simple temporal average resulting in an effective deterministic description of the slow dynamics is a rather coarse approximation, if fast chaotic modes are considered, as we will show. In fact as one of our main results we will derive a Fokker–Planck equation for the slow variables, Eq. (13), where drift and diffusion constants represent the features of the fast degrees of freedom and which can be calculated by time averages, Eqs. (14)–(16).

We would like to mention approaches which have been made for systems which are driven by a fast stochastic motion. Note that for such a system a diffusion approximation of the slow motion is known already for several decades. It is even proven rigorously that on a finite time interval the fast stochastic motion can be reduced to a kind of observational Gaussian white noise whenever time scale separation is pronounced.[8] Furthermore a kind of Langevin dynamics has been derived.[9] These approaches resemble response theory as used, e.g., by physicists for ages for the discussion of stochastic systems.[10] Since the mathematical rigorous concept is constrained to finite time intervals, the analysis is in general neither able to cope with noise induced transitions nor to describe accurately the stationary behaviour of the dynamics. Nevertheless there exist certain model systems where one can relax such a constraint and where fast chaotic degrees of freedom can be mapped to a Gaussian white noise due to certain scaling properties.[11]

In this paper we consider general model systems with two time scales, where the slow degrees of freedom $\mathbf{x}$ are coupled to fast variables $\mathbf{y}$

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= \mathbf{f}(\mathbf{x}, \mathbf{y}) \\ \frac{d\mathbf{y}}{dt} &= \frac{1}{\varepsilon}\, \mathbf{g}(\mathbf{x}, \mathbf{y}). \end{aligned} \tag{1}$$

The small parameter $0 < \varepsilon \ll 1$ mediates the separation of time scales, where $\mathbf{f}$ and $\mathbf{g}$ are assumed to be of the order of unity. It is our goal to approximate the motion of the slow variables $\mathbf{x}$ by an effective equation of motion where the fast variables $\mathbf{y}$ are replaced by a suitable stochastic process. This problem seems to be relatively simple if the slow degrees of freedom have no back-coupling to the fast ones, i.e., when $\mathbf{g}$ does not depend on $\mathbf{x}$. Then $\mathbf{y}$ is a fairly complicated stochastic process where the distribution is determined by the invariant measure of the fast dynamics. Of course there remains the highly nontrivial task to study whether one can approximate such a stochastic process in the limit of small ε by simpler processes, e.g., a white noise process. The situation is less evident when the slow degrees of freedom couple to the fast ones. Then the properties of an effective stochastic force may depend on the slow variable itself (cf. the numerical simulation in ref. 12). However, our formalism for the derivation of a stochastic model for the slow degrees of freedom will be independent of whether or not $\mathbf{g}$ depends on $\mathbf{x}$, and the only difference will be the particular dependence of the diffusion term on $\mathbf{x}$.

It is not a matter of principle, but the elimination of fast degrees of freedom is technically simpler to perform if one considers the time evolution of probability density functions. The equation of motion which corresponds to the system (1) reads

$$\frac{\partial \rho_t}{\partial t} = -\mathscr{L}\rho_t(\mathbf{x}, \mathbf{y}) = -\left(\frac{1}{\varepsilon}\mathscr{L}_0 + \mathscr{L}_1\right)\rho_t(\mathbf{x}, \mathbf{y}) \tag{2}$$

where the generator is given by

$$\begin{aligned} \mathscr{L}_0\rho(\mathbf{x}, \mathbf{y}) &= \sum_\nu \frac{\partial}{\partial y_\nu} g_\nu(\mathbf{x}, \mathbf{y})\, \rho(\mathbf{x}, \mathbf{y}) \\ \mathscr{L}_1\rho(\mathbf{x}, \mathbf{y}) &= \sum_\mu \frac{\partial}{\partial x_\mu} f_\mu(\mathbf{x}, \mathbf{y})\, \rho(\mathbf{x}, \mathbf{y}). \end{aligned} \tag{3}$$

Equation (2) displays a natural splitting of the generator which will be used to set up a perturbation scheme. In addition, the reduction to the slow degrees of freedom can be performed on the level of densities in a straightforward way by considering the reduced density

$$\bar{\rho}_t(\mathbf{x}) = \int d\mathbf{y}\, \rho_t(\mathbf{x}, \mathbf{y}) =: \mathrm{Tr}_\mathbf{y}[\rho_t]. \tag{4}$$

We have introduced the abbreviation $\mathrm{Tr}_\mathbf{y}$ to indicate the integral with respect to the fast variables.

It is the essential step to derive a closed evolution equation for the reduced density from the full equation of motion (2). The desired elimination can be performed with standard projection operator techniques which are well known in the context of nonequilibrium statistical physics.[13] To keep our paper self contained we review in Section 2 the main features of such an approach and perform the formal perturbation expansion for the model (1). In second order perturbation expansion we will derive a Fokker–Planck equation where the diffusion is given in terms of properties of the fast chaotic dynamics. In fact, the crucial part of the approach consists in the proper choice of the projection operator. In contrast to one of the previous approaches[12] here we directly end up with effective drift and diffusion coefficients which are given in terms of time averages of the fast dynamics.

Our perturbation expansion is still a formal procedure and we cannot prove its convergence. However, properties of the perturbation expansion can be understood when comparing with exactly solvable model systems. For that purpose we discuss in Section 3 a linear system coupled to a fast

Ornstein–Uhlenbeck process. Such an analysis clearly reveals the accuracy of the perturbation expansion and the relevance of the renormalization of the effective drift.

2. ELIMINATION OF FAST VARIABLES

The construction of the reduced density (4) may be considered as a formal projection of the full density. If we model the fast degrees of freedom by a fixed distribution $\rho_{ad}(\mathbf{y} \mid \mathbf{x})$ then the reduction can be achieved by the projection operator

$$\mathscr{P}\rho_t(\mathbf{x}, \mathbf{y}) = \rho_{ad}(\mathbf{y} \mid \mathbf{x}) \operatorname{Tr}_{\mathbf{y}} \rho_t = \rho_{ad}(\mathbf{y} \mid \mathbf{x}) \bar{\rho}_t(\mathbf{x}). \tag{5}$$

The normalization of the density ρ_{ad} ensures that $\mathscr{P}$ is idempotent.

Employing standard projection operator techniques[13] a formally exact and closed equation of motion for the reduced density can be derived

$$\frac{\partial \bar{\rho}_t}{\partial t} = -\operatorname{Tr}_{\mathbf{y}}[\mathscr{L}\rho_{ad}]\,\bar{\rho}_t + \int_0^t dt' \operatorname{Tr}_{\mathbf{y}}[\mathscr{L} \exp(-\mathscr{Q}\mathscr{L}t')\,\mathscr{Q}\mathscr{L}\rho_{ad}]\,\bar{\rho}_{t-t'} \tag{6}$$

where we have used the abbreviation $\mathscr{Q} = \mathscr{I} - \mathscr{P}$ to indicate the complementary projection operator. The usefulness of such an expression and in particular the properties of perturbation expansions depends crucially on the choice of the projection. In fact, there are no rules which yield a unique or at least optimal projection operator for a given problem. The choice (5) of the reduced density $\bar{\rho}_t(\mathbf{x})$ for $\mathscr{P}$ is very natural since we look for an evolution equation for the marginal distribution in the slow variable. In a previous work[12] we have chosen (5) with the conditional stationary density. Although the resulting perturbation expansion has nice formal properties and can be written down to infinite order (cf. ref. 14) the obtained drift and diffusion coefficients cannot be linked easily to dynamical properties of the fast equation of motion without making additional assumptions. Hence we follow here a different strategy and choose a different type of projection operator. Evidence for the choice of $\rho_{ad}(\mathbf{y} \mid \mathbf{x})$ for $\mathscr{P}$ is provided by the convergence of stochastic kernels (cf. ref. 15) yielding the zeroth order approximation in (13).

Let us consider the invariant density of the fast equation of motion when the slow variables $\mathbf{x}$ are considered as fixed parameters. This density ρ_{ad} obeys

$$\mathscr{L}_0 \rho_{ad}(\mathbf{y} \mid \mathbf{x}) = 0. \tag{7}$$

Averages with respect to such a density can be written as long time averages for typical initial conditions

$$\langle h\rangle_{\mathrm{ad}}(\mathbf{x}) := \int d\mathbf{y}\, h(\mathbf{x},\mathbf{y})\, \rho_{\mathrm{ad}}(\mathbf{y}\,|\,\mathbf{x}) = \lim_{T\to\infty} \frac{1}{T}\int_0^T dt\, h(\mathbf{x}, \boldsymbol{\eta}[t/\varepsilon, \mathbf{y}; \mathbf{x}]) \tag{8}$$

where $\boldsymbol{\eta}[t, \mathbf{y}; \mathbf{x}]$ denotes a solution of the fast equation of motion with initial condition $\mathbf{y}$, i.e.,

$$\frac{\partial \boldsymbol{\eta}[t, \mathbf{y}; \mathbf{x}]}{\partial t} = \mathbf{g}(\mathbf{x}, \boldsymbol{\eta}[t, \mathbf{y}; \mathbf{x}]), \qquad \boldsymbol{\eta}[t=0, \mathbf{y}; \mathbf{x}] = \mathbf{y}. \tag{9}$$

Here we assume that the dynamics of the fast degrees of freedom is mixing so that in addition correlation functions decay sufficiently rapid.

Equation (7) ensures that the projection operator (5) obeys

$$\mathscr{P}\mathscr{L}_0 = \mathscr{L}_0\mathscr{P} = 0. \tag{10}$$

Such an algebraic condition is the crucial property for the formal perturbation expansion. It ensures that the memory kernel is of higher order with respect to the expansion parameter and that the projection operator projects onto a subspace which is stationary with respect to the $\mathscr{L}_0$-dynamics. Thus the projector indeed eliminates the fast dynamics. Applying the standard second order perturbation expansion Eq. (6) reduces to

$$\frac{\partial \bar{\rho}_t}{\partial t} = -\langle \mathscr{L}_1\rangle_{\mathrm{ad}}\, \bar{\rho}_t + \int_0^t dt' \langle \mathscr{L}_1 \exp(-\mathscr{L}_0 t'/\varepsilon)\, \mathscr{Q}\mathscr{L}_1\rangle_{\mathrm{ad}}\, \bar{\rho}_{t-t'} \tag{11}$$

where third and higher order terms of the right hand side have just been discarded. The systematic part, i.e., the first term of the right hand side, just yields the adiabatic average of the slow vector field (cf. Eq. (3)). Since the kernel of the integral is determined by the propagator of the fast degrees of freedom, its evaluation is quite straightforward taking the identity

$$\mathrm{Tr}_{\mathbf{y}}[h(\mathbf{x},\mathbf{y}) \exp(-\mathscr{L}_0 t/\varepsilon)\, \rho(\mathbf{x},\mathbf{y})] = \mathrm{Tr}_{\mathbf{y}}[h(\mathbf{x}, \boldsymbol{\eta}[t/\varepsilon, \mathbf{y}; \mathbf{x}])\, \rho(\mathbf{x},\mathbf{y})] \tag{12}$$

into account. The time dependence of the kernel is governed by the correlations of the fast system. If these correlations do not decay sufficiently fast, then the memory kernel in the Master equation (11) produces a contribution which increases in time. Thus one has to require that correlations of the fast system decay, e.g., exponentially, in order to avoid such secular contributions. It is exactly this feature where we need the chaotic properties of the underlying fast motion.

In addition we may employ a Markov approximation. After some quite straightforward algebra (cf. Appendix A) we end up with the Fokker–Planck equation

$$\frac{\partial \bar{\rho}_t}{\partial t} = -\sum_{\mu} \frac{\partial}{\partial x_\mu} D_\mu^{(1,\,\mathrm{eff})}(\mathbf{x})\, \bar{\rho}_t(\mathbf{x}) + \sum_{\mu,\lambda} \frac{\partial^2}{\partial x_\mu\, \partial x_\lambda} D_{\mu\lambda}^{(2,\,\mathrm{eff})}(\mathbf{x})\, \bar{\rho}_t(\mathbf{x}). \tag{13}$$

The effective diffusion is given by the autocorrelation of the fluctuation of the slow vector field

$$D_{\mu\lambda}^{(2,\,\mathrm{eff})}(\mathbf{x}) = \int_0^\infty dt' \langle \delta_{\mathrm{ad}} f_\mu(\mathbf{x}, \boldsymbol{\eta}[t'/\varepsilon, \mathbf{y}; \mathbf{x}])\, \delta_{\mathrm{ad}} f_\lambda(\mathbf{x}, \mathbf{y}) \rangle_{\mathrm{ad}} \tag{14}$$

where

$$\delta_{\mathrm{ad}} \mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{f}(\mathbf{x}, \mathbf{y}) - \langle \mathbf{f} \rangle_{\mathrm{ad}}\, (\mathbf{x}) \tag{15}$$

denotes the static fluctuation. The effective drift consists of the adiabatic average of the slow vector field and a renormalization by chaotic fluctuations

$$D_\mu^{(1,\,\mathrm{eff})}(\mathbf{x}) = \langle f_\mu \rangle_{\mathrm{ad}}\, (\mathbf{x}) + \sum_\lambda \int_0^\infty dt' \langle f_\lambda(\mathbf{x}, \mathbf{y})\, \partial \delta_{\mathrm{ad}} f_\mu(\mathbf{x}, \boldsymbol{\eta}[t'/\varepsilon, \mathbf{y}; \mathbf{x}]) / \partial x_\lambda \rangle_{\mathrm{ad}}. \tag{16}$$

Since by assumption the correlation functions of the fast chaotic dynamics decay on a time scale proportional to ε, the integrals in Eqs. (14) and (16) converge, if we impose some regularity assumptions. The coefficients of the effective Fokker–Planck equation (13) are just adiabatic averages and can be computed as plain temporal averages with respect to the fast dynamics (cf. Eq. (8)).

Our perturbation expansion results to lowest nontrivial order in a Fokker–Planck equation. Equivalently one may describe the slow degrees of freedom by a stochastic differential equation with a Gaussian white noise (cf., e.g., ref. 16 for explicit expressions). From such a point of view the validity of the perturbation expansion is not obvious. Gaussian stochastic forces are unbounded in contrast to the real fast degrees of motion which typically have finite amplitude even in the limit of small ε. In certain situations the properties of exit time or tunnelling problems may depend on whether the forces driving the slow degrees of freedom have finite or infinite amplitude. If for instance the amplitude of the fast modes $\mathbf{y}$ is small one may apply linear response theory to the system (1) and the nontrivial aspects of exit time dynamics are suppressed completely. However the mathematical approaches mentioned in Section 1 state that regardless of such constraints the fast motion can be approximated by Gaussian stochastic forces on finite time intervals provided the time scale separation is pronounced (cf. the numerical simulations in ref. 12).

If we extend the formal expansion of Eq. (6) to higher orders then we obtain evolution equations which do not any longer display a Fokker–Planck structure but contain higher order derivatives. These contributions are not *a priori* small in a mathematical setting. Furthermore, in any finite order beyond the second, the density may loose its positivity so that no simple mapping of the density equation to a stochastic differential equation might be possible.

Thus the validity of the perturbation expansion on long time scales is a rather subtle subject and finally might involve the amplitude of the fast modes $\mathbf{y}$, too. To clarify this issue to some extent it helps to consider simple model systems that can be solved analytically. One should not expect an ultimate answer to all the questions just raised, but at least one gets some insight into the quality of the Born approximation (11).

3. EXACT SOLUTION OF A LINEAR MODEL

The perturbation expansion of the previous section is a formal procedure and it seems quite difficult to estimate its validity. Thus we are going to set up a simple model system where exact solutions are available and where the features of the expansion can be studied. There is unfortunately no simple time continuous chaotic model available which can serve for such a purpose. Geodesic flows on surfaces of negative curvature are technically cumbersome to handle and simple chaotic maps that can be generated by a kicked dynamics do not fit straightforwardly in our setup (cf., e.g., ref. 11 for a study of maps coupled to a dynamical system). In order to concentrate on the essential features of our perturbation expansion we follow a different strategy. We model the fast chaotic dynamics by a stochastic process and we consider a completely linear model so that analytical solutions can be obtained without great effort. The equations of motion read

$$\begin{aligned} \frac{dx}{dt} &= -\alpha x + \beta y, \qquad (\alpha > 0) \\ \frac{dy}{dt} &= \frac{1}{\varepsilon}(-\lambda y + \mu x) + \frac{\kappa}{\sqrt{\varepsilon}}\xi(t), \qquad (\lambda > 0) \end{aligned} \tag{17}$$

where ξ denotes a Gaussian white noise with correlation function $\langle \xi(t)\, \xi(t') \rangle = 2\delta(t-t')$, and the stochastic integrals are interpreted in the sense of

Stratonovich. Thus the fast chaotic mode is an Ornstein–Uhlenbeck process coupled to the slow coordinate. In fact models of this type are very well investigated, i.e., for the study of coloured noise.[17] Here we just focus on the elimination of the fast degree of freedom in lowest nontrivial order of the expansion parameter ε. We have to scale the noise amplitude with $1/\sqrt{\varepsilon}$ to guarantee that the corresponding Fokker–Planck operator can be split according to Eq. (2). The generator (3) of the dynamics reads

$$\begin{aligned}\mathscr{L}_0 \rho(x, y) &= \left(\frac{\partial}{\partial y}[-\lambda y+\mu x]-\frac{\partial^2}{\partial y^2}\kappa^2\right)\rho(x, y)\\ \mathscr{L}_1 \rho(x, y) &= \frac{\partial}{\partial x}[-\alpha x+\beta y]\,\rho(x, y).\end{aligned} \tag{18}$$

It splits again the motion into a fast and a slow part. The generator $\mathscr{L}_0$ already contains a diffusive part since the fast motion is constructed by a Gaussian stochastic process. Our general considerations of the previous section are however not influenced by such a property.

It is quite straightforward to determine the adiabatic density using Eq. (7):

$$\rho_{\mathrm{ad}}(y\,|\,x) = \sqrt{\frac{\lambda}{2\pi\kappa^2}}\exp\left[-\frac{\lambda}{2\kappa^2}\left(y-\frac{\mu}{\lambda}x\right)^2\right]. \tag{19}$$

Hence the adiabatic average (8) and the static fluctuation (15) of the slow vector field read

$$\begin{aligned}\langle f\rangle_{\mathrm{ad}}(x) &= -\left(\alpha-\beta\frac{\mu}{\lambda}\right)x\\ \delta_{\mathrm{ad}} f(x, y) &= \beta\delta_{\mathrm{ad}}\,y = \beta\left(y-\frac{\mu}{\lambda}x\right).\end{aligned} \tag{20}$$

If $\mathscr{L}_0^\dagger$ denotes the adjoint operator of $\mathscr{L}_0$ with respect to Tr_y we have $\mathscr{L}_0^\dagger\delta_{\mathrm{ad}}\,y = -\lambda\delta_{\mathrm{ad}}\,y$, and the time dependent fluctuations are easily obtained as[4]

$$\exp(-\mathscr{L}_0^\dagger t/\varepsilon)\,\delta_{\mathrm{ad}} f(x, y) = \exp(-\lambda t/\varepsilon)\,\beta\delta_{\mathrm{ad}}\,y = \delta_{\mathrm{ad}} f(x, \eta[t/\varepsilon, y; x]). \tag{21}$$

[4] In order to keep for our stochastic model (17) the same notation as in Section 2 we take Eq. (21) as the definition for $\delta_{\mathrm{ad}} f(x, \eta[t/\varepsilon, y; x])$.

Thus taking the definition (21) into account the evaluation of the effective drift and diffusion coefficients (16), (14) of second order perturbation theory is quite simple and we end up with

$$D^{(1,\,\mathrm{eff})}(x) = -\left(\alpha-\beta\frac{\mu}{\lambda}\right)\left(1-\varepsilon\beta\frac{\mu}{\lambda^2}\right)x =: -\alpha^{\mathrm{eff}}x$$
$$D^{(2,\,\mathrm{eff})}(x) = \epsilon\kappa^2\frac{\beta^2}{\lambda^2}. \tag{22}$$

The effective diffusion is determined by the time scale of the fluctuations and the coupling strength of the fast degrees of freedom to the slow ones. The coefficient α^{eff} of the effective drift is renormalised by fast degrees of freedom in two ways. The renormalization in the first factor is already known from centre manifold theory and just takes the transformation to the slow manifold into account (cf. Eq. (27)). The renormalization by the second factor comes through the fluctuations of the fast variables.

It is quite well known that the motion of the slow variable is governed by an effective Fokker–Planck equation only up to the second order of the perturbation expansion.[17] Thus even for such a simple linear system no plain Fokker–Planck structure shows up on an exact level. To estimate the accuracy of our perturbative result we compare it to the solution of the full system. Let us first consider the stationary solution of the reduced density. Taking the expressions (22) into account the effective Fokker–Planck equation (13) yields

$$\bar{\rho}_*(x) = \sqrt{\frac{\alpha^{\mathrm{eff}}\lambda^2}{2\pi\varepsilon\kappa^2\beta^2}}\exp\left(-\frac{\alpha^{\mathrm{eff}}\lambda^2}{2\varepsilon\kappa^2\beta^2}x^2\right). \tag{23}$$

Since the full system (cf. Eqs. (18)) is a two-dimensional linear Fokker–Planck system its stationary solution can be written down easily. For the reduced density we obtain

$$\bar{\rho}_*(x)|_{\mathrm{exact}} = \sqrt{\frac{\hat{\alpha}\lambda^2}{2\pi\varepsilon\kappa^2\beta^2}}\exp\left(-\frac{\hat{\alpha}\lambda^2}{2\varepsilon\kappa^2\beta^2}x^2\right) \tag{24}$$

where

$$\hat{\alpha} = \left(\alpha-\beta\frac{\mu}{\lambda}\right)\left(1+\varepsilon\frac{\alpha}{\lambda}\right). \tag{25}$$

Comparison shows that our perturbation expansion reproduces the stationary reduced distribution, i.e., all its cumulants, in the lowest nontrivial order of ε.

As a second quantifier for the accuracy of the perturbation expansion let us dwell on the relaxation rate, i.e., the first nontrivial eigenvalue of the Fokker–Planck operator. These eigenvalues are just determined by the drift of the Fokker–Planck equation (cf., e.g., ref. 16). Equation (13) together with the expressions (21) yields

$$\Lambda = \left(\alpha - \beta \frac{\mu}{\lambda}\right)\left(1 - \varepsilon\beta \frac{\mu}{\lambda^2}\right) \tag{26}$$

whereas the exact quantity is obtained from the diagonalization of the deterministic part of the two-dimensional system (17)

$$\Lambda|_{\text{exact}} = \frac{\alpha - \beta\mu/\lambda}{(1+\varepsilon\alpha/\lambda)/2 + \sqrt{(1-\varepsilon\alpha/\lambda)^2/4 + \varepsilon\beta\mu/\lambda^2}} = \Lambda + O(\varepsilon^2). \tag{27}$$

Here we observe in fact a coincidence including the first order in ε which is one order beyond the lowest nontrivial order in ε. In the one-dimensional case, i.e., in our approximation, both the width of the distribution and the relaxation rate are essentially determined by $D^{(1,\text{eff})}$. Since the width of the exact reduced density $\bar{\rho}_*(x)$ is not identical to the relaxation of the two-dimensional full problem, no one-dimensional projection can reproduce both results simultaneously. Our approximation reproduces the relaxation rate with better accuracy than the width of the distribution. Such a property depends crucially on the proper drift renormalization through the chaotic fluctuations.

4. CONCLUSION

We have described a scheme to model fast chaotic degrees of freedom by suitable stochastic forces. The approach is based on a description using densities and employs standard projection techniques which are known in statistical mechanics for decades. Contrary to previous work[12] the projection operator proposed here yields effective diffusion and drift renormalization which can be expressed in terms of temporal averages with respect to the fast chaotic motion.

The crucial step of the whole procedure is the application of the perturbation expansion with respect to the time scale separation and the restriction to a low order of the perturbation expansion. The formal expansion can be developed easily and we end up with a Fokker–Planck equation respectively a Langevin equation with Gaussian white noise. To check the accuracy of the expansion we have analysed a simple linear model which can be solved exactly. The perturbation expansion reproduces the stationary density in the lowest nontrivial order, and it reproduces temporal characteristics including the first order of ε. We should stress that the diffusion constant is of order ε, hence in the limit of perfect time scale separation the diffusion tends to zero.

In general one cannot expect that such formal expansions are uniformly valid in time. Fast chaotic degrees of freedom have typically a finite amplitude, i.e., such motion cannot induce transitions between different stationary states if the amplitude of chaotic oscillations is too small. On the other hand any arbitrary small Gaussian noise induces tunnelling phenomena on exponentially long time scales. Nevertheless the modelling of fast chaotic degrees of freedom may be suitable on finite time scales. But such subtle mechanisms turn a real proof of the convergence of the expansion into a real challenge and there is no ultimate answer available yet.

APPENDIX A. THE PERTURBATION EXPANSION

To keep our presentation entirely self contained we will describe the derivation of the Fokker–Planck equation in more detail. Subject of our investigation is the propagation of a statistical ensemble representing the distribution of initial states on some bounded domain Ω of the phase space. We take in the following the physicists' point of view and assume for simplicity that the ensemble can be described by a measure which is absolutely continuous with respect to the Lebesgue measure. Hence states can be described by densities and the Liouville-like equation (2) which is understood in the sense of distributions governs the motion of the system. For general dynamical systems especially the existence of invariant distributions which are absolutely continuous crucially depends on the underlying dynamics. We expect that most parts of our approach can be rewritten in terms of measures, but we do not give a rigorous account here.

Let us assume that on Ω the forward flow for Eq. (1) exists. Furthermore we suppose that the dynamics (9) in the fast variables possesses a smooth stationary distribution $\rho_{\text{ad}}(\cdot \,|\, \mathbf{x})$ for any frozen $\mathbf{x}$ and that the motion is exponentially mixing with a mixing rate which is uniformly bounded for all $\mathbf{x}$. Let us consider an ensemble initially distributed according to some density ρ for which $\rho(\mathbf{x}, \mathbf{y}) = \rho_{\text{ad}}(\mathbf{y} \,|\, \mathbf{x}) \, \mathrm{Tr}_{\mathbf{y}}[\rho(\mathbf{x}, \mathbf{y})]$.

We apply the Zwanzig projection method[13] to Eq. (2) and obtain

$$\frac{\partial \mathscr{P} \rho_t}{\partial t} = -\mathscr{P}\mathscr{L}\mathscr{P}\rho_t + \mathscr{P}\mathscr{L} \int_0^t e^{-\mathscr{Q}\mathscr{L}s} \mathscr{Q}\mathscr{L}\mathscr{P} e^{-\mathscr{L}(t-s)} \rho \, ds - \mathscr{P}\mathscr{L} e^{-\mathscr{Q}\mathscr{L}t} \mathscr{Q}\rho, \quad \rho_{t=0} = \rho, \tag{28}$$

which is the starting point for our approximations. The first term of the right hand side of (28) can be viewed as the self-interaction of the projected components. For the projector (5) it is the generator for the dynamics of the adiabatically averaged slow motion (see below). The second term embodies a memory term. The third term is determined only by the complementary component of the initial distribution and vanishes due to our assumption for the initial distribution. Note that if such an assumption is not satisfied additional contributions to the equation of motion for $\mathscr{P}\rho_t$ may occur. However such contributions will be discarded since they should not play any relevant role, e.g., for evaluating stationary distributions.

The Nakajima–Zwanzig equation (28) involves the dynamics which is determined by the abstract problem

$$\frac{\partial \sigma_t}{\partial t} = -\mathscr{Q}\mathscr{L}\sigma_t, \qquad \sigma_{t=0} = \sigma. \tag{29}$$

One can easily check that for a solution of (29) we always have $\partial\mathscr{P}\sigma_t/\partial t = 0$ and hence $\mathscr{P}\sigma_t = 0$ for all t provided that $\mathscr{P}\sigma = 0$. Thus (29) models the complementary dynamics. Although the Zwanzig projection method is a well-established technique we would like to remark that contrary to the Liouville-like equation (1) where solutions of associated initial value problems always exist it is a severe problem to show that Eq. (29) has a proper mathematical meaning. A mathematically rigorous setup of the Nakajima–Zwanzig equation is a nontrivial task. No general solution has been proposed so far to our best knowledge (cf. ref. 18 for the analysis of finite rank projection operators). In our opinion it is a challenge for further research from the mathematical point of view.

Here we just focus on the formal evaluation of Eq. (28) which is based on the essential property (10). We obtain for Eq. (28) using $\mathscr{Q}\mathscr{L} = \varepsilon^{-1}\mathscr{Q}\mathscr{L}_0 + \mathscr{Q}\mathscr{L}_1$ and the standard exponential identity

$$\frac{\partial}{\partial t}\mathscr{P}\rho_t = -\mathscr{P}\mathscr{L}_1\mathscr{P}\rho_t + \mathscr{P}\mathscr{L}_1\int_0^t e^{-\mathscr{Q}\mathscr{L}_0(t-s)/\varepsilon}\mathscr{Q}\mathscr{L}_1\mathscr{P}\rho_s\,ds$$

$$-\varepsilon^2\mathscr{P}\mathscr{L}_1\int_0^{t/\varepsilon}\int_0^{s/\varepsilon} e^{-\mathscr{Q}\mathscr{L}_0(s-\tau)}\mathscr{Q}\mathscr{L}_1 e^{-\mathscr{Q}(\mathscr{L}_0+\varepsilon\mathscr{L}_1)\tau}\mathscr{Q}\mathscr{L}_1\mathscr{P}\rho_{t-\varepsilon s}\,d\tau\,ds. \tag{30}$$

In this presentation we are aiming to derive an approximate equation of motion for $\bar{\rho}_t$ including corrections up to order ε only. Therefore we may truncate Eq. (30) by neglecting the third term on the right hand side which is formally of order ε^2. Whether such a Born approximation is in general valid on long time scales is still a very delicate problem. In fact such an approximation means that in the memory kernel the full propagator $\exp(-\mathscr{Q}\mathscr{L}t)$ can be replaced by the lowest order contribution $\exp(-\mathscr{Q}\mathscr{L}_0 t/\varepsilon)$. Evaluation of the remaining contributions in Eq. (30) is quite straightforward

$$\mathscr{P}\mathscr{L}_1\mathscr{P}\rho_t = \rho_{\mathrm{ad}}(\mathbf{y}\,|\,\mathbf{x}) \sum_\mu \frac{\partial}{\partial x_\mu} \langle f_\mu \rangle_{\mathrm{ad}}(\mathbf{x})\, \bar{\rho}_t(\mathbf{x}) \tag{31}$$

$$\begin{aligned}
&\mathscr{P}\mathscr{L}_1 e^{-\mathscr{Q}\mathscr{L}_0(t-s)/\varepsilon} \mathscr{Q}\mathscr{L}_1\mathscr{P}\rho_s \\
&\quad = \rho_{\mathrm{ad}}(\mathbf{y}\,|\,\mathbf{x})\, \mathrm{Tr}_{\mathbf{y}} \Bigg[\sum_\mu \frac{\partial}{\partial x_\mu} f_\mu(\mathbf{x}, \mathbf{y}) \Bigg(e^{-\mathscr{Q}\mathscr{L}_0(t-s)/\varepsilon} \sum_\lambda \frac{\partial}{\partial x_\lambda} f_\lambda(\mathbf{x}, \mathbf{y})\, \rho_{\mathrm{ad}}(\mathbf{y}\,|\,\mathbf{x}) \\
&\qquad - e^{-\mathscr{Q}\mathscr{L}_0(t-s)/\varepsilon} \rho_{\mathrm{ad}}(\mathbf{y}\,|\,\mathbf{x})\, \mathrm{Tr}_{\tilde{\mathbf{y}}} \Bigg[\sum_\lambda \frac{\partial}{\partial x_\lambda} f_\lambda(\mathbf{x}, \tilde{\mathbf{y}})\, \rho_{\mathrm{ad}}(\tilde{\mathbf{y}}\,|\,\mathbf{x}) \Bigg] \Bigg) \bar{\rho}_s(\mathbf{x}) \Bigg].
\end{aligned} \tag{32}$$

Since $\rho_{\mathrm{ad}}(\cdot\,|\,\mathbf{x})$ is stationary we have $\exp(-\mathscr{Q}\mathscr{L}_0 t)\, \rho_{\mathrm{ad}}(\mathbf{y}\,|\,\mathbf{x}) = \rho_{\mathrm{ad}}(\mathbf{y}\,|\,\mathbf{x})$ for any $t \geqslant 0$. For an integrable function ψ which vanishes at the boundaries of the domain Ω we have $\mathrm{Tr}_{\mathbf{y}}[\psi] = \mathrm{Tr}_{\mathbf{y}}[\exp(-\mathscr{Q}\mathscr{L}_0 t)\, \psi]$ for any $t \geqslant 0$. Therefore, Eq. (32) may be expressed in terms of the static fluctuation (15). Employing in addition the solution of the fast equation of motion (9) with initial condition $\mathbf{y}$ and taking into account the identity $\exp(-\mathscr{L}_0^\dagger \mathscr{Q}^\dagger t)\, \delta_{\mathrm{ad}} f_\mu(\mathbf{x}, \mathbf{y}) = \delta_{\mathrm{ad}} f_\mu(\mathbf{x}, \boldsymbol{\eta}[t/\varepsilon, \mathbf{y}; \mathbf{x}])$ the equation of motion (30) results in

$$\begin{aligned}
\frac{\partial \bar{\rho}_t(\mathbf{x})}{\partial t} &= -\sum_\mu \frac{\partial}{\partial x_\mu} \langle f_\mu \rangle_{\mathrm{ad}}(\mathbf{x})\, \bar{\rho}_t(\mathbf{x}) + \sum_{\mu,\lambda} \frac{\partial^2}{\partial x_\mu\, \partial x_\lambda} I^{(2)}_{\mu\lambda}(\varepsilon, t, \mathbf{x}) \\
&\quad -\sum_\mu \frac{\partial}{\partial x_\mu} I^{(1)}_\mu(\varepsilon, t, \mathbf{x})
\end{aligned} \tag{33}$$

where the two remaining integral terms have been abbreviated by

$$\begin{aligned}
&I^{(2)}_{\mu\lambda}(\varepsilon, t, \mathbf{x}) \\
&\quad := \int_0^t \mathrm{Tr}_{\mathbf{y}}[\delta_{\mathrm{ad}} f_\mu(\mathbf{x}, \boldsymbol{\eta}[(t-s)/\varepsilon, \mathbf{y}; \mathbf{x}])\, \delta_{\mathrm{ad}} f_\lambda(\mathbf{x}, \mathbf{y})\, \rho_{\mathrm{ad}}(\mathbf{y}\,|\,\mathbf{x})]\, \bar{\rho}_s(\mathbf{x})\, ds
\end{aligned} \tag{34}$$

and by

$$\begin{aligned}
&I^{(1)}_\mu(\varepsilon, t, \mathbf{x}) \\
&:= \int_0^t \mathrm{Tr}_{\mathbf{y}} \sum_\lambda \Bigg[\Bigg(\frac{\partial}{\partial x_\lambda} \delta_{\mathrm{ad}} f_\mu(\mathbf{x}, \boldsymbol{\eta}[(t-s)/\varepsilon, \mathbf{y}; \mathbf{x}]) \Bigg) f_\lambda(\mathbf{x}, \mathbf{y})\, \rho_{\mathrm{ad}}(\mathbf{y}\,|\,\mathbf{x}) \Bigg] \bar{\rho}_s(\mathbf{x})\, ds.
\end{aligned} \tag{35}$$

Now we employ a Markov approximation. The fast dynamics is supposed to be exponentially mixing. Hence correlation functions decay rapidly. Moreover $\mathrm{Tr}_{\mathbf{y}}[\delta_{\mathrm{ad}} f_\mu(\mathbf{x}, \mathbf{y})\, \rho_{\mathrm{ad}}(\mathbf{y} \mid \mathbf{x})] = 0$. Thus, the kernel in Eq. (34) decays on a time scale determined by the inverse mixing rate of the fast subsystem. Replacing $\bar{\rho}_s$ by $\bar{\rho}_t$ and extending the integration to infinity we can approximate the integral term (34) by $D^{(2,\,\mathrm{eff})}_{\mu\lambda}(\mathbf{x})\, \bar{\rho}_t(\mathbf{x})$. Analogously, investigating Eq. (35) we use the equality

$$\begin{aligned}
&\mathrm{Tr}_{\mathbf{y}}\left[\left(\frac{\partial}{\partial x_\lambda} \delta_{\mathrm{ad}} f_\mu(\mathbf{x}, \boldsymbol{\eta}[s, \mathbf{y}; \mathbf{x}])\right) f_\lambda(\mathbf{x}, \mathbf{y})\, \rho_{\mathrm{ad}}(\mathbf{y} \mid \mathbf{x})\right] \\
&= \frac{\partial}{\partial x_\lambda} \mathrm{Tr}_{\mathbf{y}}[\delta_{\mathrm{ad}} f_\mu(\mathbf{x}, \boldsymbol{\eta}[s, \mathbf{y}; \mathbf{x}])\, f_\lambda(\mathbf{x}, \mathbf{y})\, \rho_{\mathrm{ad}}(\mathbf{y} \mid \mathbf{x})] \\
&\quad - \mathrm{Tr}_{\mathbf{y}}\left[\delta_{\mathrm{ad}} f_\mu(\mathbf{x}, \boldsymbol{\eta}[s, \mathbf{y}; \mathbf{x}])\left(\frac{\partial f_\lambda(\mathbf{x}, \mathbf{y})}{\partial x_\lambda} + f_\lambda(\mathbf{x}, \mathbf{y}) \frac{\partial \ln \rho_{\mathrm{ad}}(\mathbf{y} \mid \mathbf{x})}{\partial x_\lambda}\right) \rho_{\mathrm{ad}}(\mathbf{y} \mid \mathbf{x})\right].
\end{aligned} \tag{36}$$

Because of mixing the right hand side tends to zero as $s \to \infty$. Thus, for all t exceeding the inverse mixing rate we can approximate the integral term (35) by $(D^{(1,\,\mathrm{eff})}_\mu(\mathbf{x}) - \langle \mathbf{f}_\mu \rangle_{\mathrm{ad}}(\mathbf{x}))\, \bar{\rho}_t(\mathbf{x})$. Putting all terms together, we obtain the approximate equation (13).

REFERENCES

1. N. G. van Kampen, Elimination of fast variables, *Phys. Rep.* **124**:69 (1985).
2. M. C. Cross and P. C. Hohenberg, Pattern formation outside of equilibrium, *Rev. Mod. Phys.* **65**:851 (1993).
3. H. Haken, *Synergetics* (Springer, Berlin, 1977).
4. J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer, New York, 1986).
5. C. Jones, Geometric singular perturbation theory, in *Dynamical Systems*, R. Johnson, ed. (Springer, Berlin, 1995).
6. A. Sanders and F. Verhulst, *Averaging Methods in Nonlinear Dynamical Systems* (Springer, New York, 1985).
7. K. Hasselmann, Stochastic climate models: Part I. Theory, *Tellus* **28**:473 (1976).
8. M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, New York, 1998).
9. R. Z. Khasminskij, A limit theorem for the solutions of differential equations with random right-hand sides, *Theory Probab. Appl.* **11**:390 (1966).
10. M. Suzuki, Statistical mechanics of non-equilibrium systems. II, *Progr. Theoret. Phys.* **55**:383 (1976).
11. C. Beck, From the Perron–Frobenius equation to the Fokker–Planck equation, *J. Stat. Phys.* **79**:875 (1995).
12. W. Just, H. Kantz, C. Rödenbeck, and M. Helm, Stochastic modelling: Replacing fast degrees of freedom by noise, *J. Phys. A* **34**:3199 (2001).

13. R. Zwanzig, Ensemble method in the theory of irreversibility, *J. Chem. Phys.* **33**:1338 (1960).
14. H. Mori, H. Fujisaka, and H. Shigematsu, A new expansion of the master equation, *Progr. Theoret. Phys.* **51**:109 (1974).
15. K. Gelfert, Asymptotic behaviour of multi time-scale systems via Young measure limits, preprint (2002).
16. H. Risken, *The Fokker–Planck Equation: Methods of Solution and Applications* (Springer, Berlin, 1989).
17. P. Grigolini, L. A. Lugiato, R. Mannella, P. V. E. McClintock, M. Merri, and M. Pernigo, Fokker–Planck description of stochastic processes with colored noise, *Phys. Rev. A* **38**:1966 (1988).
18. O. Hald and R. Kupferman, Existence of orthogonal dynamics, preprint (2002).