Stochastic modelling: Replacing fast degrees of freedom by noise

Wolfram Just†, Holger Kantz‡, Christian Rödenbeck♣, Mario Helm‡

† School of Mathematical Sciences, Queen Mary & Westfield College, Mile End Road, London E1 4NS, UK

‡ Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Str. 38, D–01187 Dresden, Germany

♣ Max Planck Institute for Biogeochemistry, Carl Zeiss Promenade 10, D–07745 Jena, Germany

E-mail: `W.Just@qmw.ac.uk`, `kantz@mpipks-dresden.mpg.de`, `roeden@mpibgc-jena.mpg.de`

Abstract. In systems with time scale separation, where the fast degrees of freedom exhibit chaotic motion, the latter are replaced by suitable stochastic processes. A projection technique is employed to derive equations of motion for the phase space density of the slow variables by eliminating the fast ones. The resulting equations can be approximated in a controlled way by Fokker–Planck equations or equivalently by stochastic differential equations for the slow degrees of freedom. We discuss some model situations and explore the accuracy of the approximations by numerical simulations.

Submitted to: *J. Phys. A: Math. Gen.*

PACS number: 05.45.-a, 05.10.Gg

1. Introduction

Many phenomena in nature are characterised by the interaction of dynamics on very different time scales. Within the context of statistical mechanics such phenomena are well known for decades and result finally in effective reduced descriptions like e.g. Boltzmann–, master–, or Fokker–Planck equations [1]. Here one usually applies a thermodynamic limit for the fast degrees of freedom and assumes to some degree local thermodynamic equilibrium which gives rise to fluctuation dissipation relations. However these thermodynamic properties of the "heat bath" are in principle dispensable, since one just relies on the decay of correlations of fast degrees of freedom. Thus, one should be able to apply similar approaches in few degrees of freedom systems having chaotic dynamics.

One particular example where these considerations play an important role is the ocean–atmosphere system generating our weather and our climate. Here on the one hand a considerable separation of time scales appears. The ocean is characterised by time scales of hundreds of years, whereas the atmospheric structures such as cyclones and anticyclones have lifetimes of a couple of days. On the other hand one cannot model the fast degrees of freedoms as a thermodynamic heat bath, although its effective dynamical degrees of freedom may be large. Nevertheless it might be highly useful to reduce the number of degrees of freedom and in particular to get rid of the fast degrees, if one is interested in the long time behaviour. The most obvious approximation would be to ignore the fast degrees of freedom completely, or at least to replace them by a function of the slow ones. Such approaches known as "adiabatic elimination" [2] or centre manifold reductions work well if the fast degrees of freedom are governed by essentially relaxatory dynamics. If the fast degrees are periodic, the "averaging theorem" [3] supplies a framework for their elimination. In both cases, effective deterministic equations of motion containing exclusively the slow degrees of freedom are found. However, for chaotic degrees of freedom, these procedures have to be improved in order not to lose relevant dynamical scenarios.

In the framework of climate research, the notion of stochastic climate models has been coined by Hasselmann [4] and recently recalled by mathematicians [5]. The idea is to replace the fast degrees of freedom by suitable stochastic processes, i.e. to replace the full deterministic system by a stochastic differential equation with only slow degrees of freedom. It is, however, until now unclear how to derive systematically a stochastic model from a deterministic one and in which sense the stochastic equation might be an approximation of the deterministic system. In this paper we will make use of a projection technique which yields answers to both questions.

Related work, however with a different goal and different techniques, was done for specific systems by C. Beck. He showed that central limit theorems can be applied to chaotic deterministic systems [6], hence showing that chaotic systems can exhibit stochastic signatures in certain limits (cf. also the approaches in [7, 8]). Moreover, he was able to prove that the Perron–Frobenius equation governing the time evolution of the invariant measure of a particular class of dynamical systems is, in a certain limit, equivalent to a Fokker–Planck equation [9]. The setting studied in [6, 9] is in fact a special case of time scale separation and hence a specific example of what we want to show here in more generality.

To specify our formal setup we restrict ourselves to the case of only two time scales, called "slow" and "fast". For simplicity and pedagogical reasons we assume that we can decompose the set of all phase space variables into two groups of dimensionality d_x and d_y, where $\boldsymbol{x}$ are the slow variables and $\boldsymbol{y}$ the fast ones. If we assume that the right hand sides of the following differential equations, $\boldsymbol{f}$ and $\boldsymbol{g}$, are of the order of unity, the time scale separation can be mediated by a small parameter ε through

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) \tag{1}$$

$$\dot{\boldsymbol{y}} = \frac{1}{\varepsilon}\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) \quad , \tag{2}$$

where $0 < \varepsilon \ll 1$. The ultimate goal of this paper will be to derive a stochastic differential equation for $\boldsymbol{x}$ alone

$$\dot{\tilde{\boldsymbol{x}}} = \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \boldsymbol{\xi}) \quad , \tag{3}$$

where $\boldsymbol{\xi}$ is a vector valued noise process with suitable correlations.

The issue of how to derive eq.(3) from eqs.(1) and (2) is directly linked to the question whether and in which sense the stochastic description approximates the deterministic system. It is evident that the introduction of any kind of noise process can never improve the accuracy of a particular solution, i.e. the error $|\boldsymbol{x}(t) - \tilde{\boldsymbol{x}}(t)|$ can never be made smaller on average by introducing uncertainty in the evolution equation of $\tilde{\boldsymbol{x}}$. Even if the error is smaller for one particular realization of the noise process, it will inevitably be larger for other realizations. As a consequence, the gain by introducing noise can only be in the sense of more realistic modelling, or more realistic averages. In Sections 4 and 5 we will study examples where the fast degrees of freedom create dynamical phenomena which disappear when eliminating the fast degrees and which reappear when introducing noise instead. So in the end, our criterion will be the accuracy by which the reduced equation reproduces the invariant measure of the full system restricted to the slow degrees of freedom, and on the accuracy of ergodic averages on this measure. Hence, a relevant and exact intermediate result will be an equation for the time evolution of the phase space density of the slow variables, where the fast degrees of freedom have been self–consistently eliminated.

Section 2 is devoted to the formal elimination scheme to derive Fokker–Planck–like evolution equations for the phase space density of the slow degrees of freedom. Among others, one likes to convert the formal solution into Langevin equations for the slow variables in order to be able to perform numerical solutions of the initial value problem, which is the typical setting in most numerical modelling tasks. The resulting Langevin equation assumes a particularly simple form for skew systems, which are discussed in Section 3. Sections 4 and 5 are devoted to two different numerical examples. The first one, presented in Section 4, will allow us to make a quantitative comparison between the numerical results and our analytic predictions. One particular issue will be how the accuracy of the approximations made at the end of Section 2 will depend on the magnitude of the time scale separation. The second example which is discussed in Section 5 exploits some aspect of the full nonlinear coupling between the different types of degree of freedom. Bifurcations in the fast subsystem may considerably influence the nature of the effective noise process. Some technical details are summarised in three appendices to keep the paper self contained.

2. Temporal evolution in terms of Fokker–Planck equations

In order to perform the elimination of the fast degrees of freedom we switch from the trajectory–wise description of eqs.(1), (2) to an equivalent description in terms of the phase space density $\rho_t(\boldsymbol{x}, \boldsymbol{y})$ and the corresponding Liouville–like equation

$$\frac{\partial \rho_t}{\partial t} = -\mathcal{L}\rho_t \quad . \tag{4}$$

Here

$$\mathcal{L} := \frac{1}{\varepsilon}\mathcal{L}^{(0)} + \mathcal{L}^{(1)} := \frac{1}{\varepsilon}\frac{\partial}{\partial \boldsymbol{y}}\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) + \frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) \tag{5}$$

denotes the generator of the dynamics. We assume for simplicity that all densities are smooth and can be treated like ordinary functions‡. Elimination of fast variables is now most conveniently done by considering the distribution of slow variables

$$\bar{\rho}_t(\boldsymbol{x}) = \int d\boldsymbol{y}\, \rho_t(\boldsymbol{x}, \boldsymbol{y}) \quad . \tag{6}$$

It is the goal to derive a closed equation of motion for the distribution (6) from the full equation of motion (4). The spirit of such a procedure consists in approximating the full distribution by a reduced density, where the fast variables are in some sense adiabatically eliminated

$$\rho_t^{red}(\boldsymbol{x}, \boldsymbol{y}) = \rho^{cond}(\boldsymbol{y}|\boldsymbol{x})\bar{\rho}_t(\boldsymbol{x}) \quad . \tag{7}$$

Regardless of the special form of ρ^{cond} it is possible to write down an exact equation of motion for $\bar{\rho}_t$ by using well established projection operator techniques [10]. However, the usefulness and the evaluation of such equations in terms of a perturbation expansion depends crucially on the particular choice for ρ_{cond}.

As discussed also in a previous study [11], two different choices seem to be appealing. From the mathematical point of view one would guess that as long as $\boldsymbol{y}$ is fast enough it adjusts itself almost instantaneously according to the invariant distribution of the fast equation (2) with $\boldsymbol{x}$ being considered as a fixed parameter. The corresponding density $\rho_{\boldsymbol{x}}^{add}(\boldsymbol{y})$ has the advantage that averages with respect to $\boldsymbol{y}$ can be computed as time

‡ One may add a small diffusive like contribution to eq.(5) to ensure smoothness and consider the limit of vanishing diffusion at the end. In addition, a generalisation in terms of corresponding measures seems to be possible, but would increase the amount of notation considerably.

averages provided certain ergodic properties are met. If $\boldsymbol{\eta}_{\boldsymbol{x}}[t, \boldsymbol{y}]$ denotes the solution of eq.(2) with initial condition $\boldsymbol{\eta}_{\boldsymbol{x}}[0, \boldsymbol{y}] = \boldsymbol{y}$, then

$$\langle h \rangle_{\boldsymbol{y}} := \int d\boldsymbol{y}\, h(\boldsymbol{x}, \boldsymbol{y}) \rho_{\boldsymbol{x}}^{add}(\boldsymbol{y}) = \lim_{T\to\infty} \int_0^T dt\, h(\boldsymbol{x}, \boldsymbol{\eta}_{\boldsymbol{x}}[t, \boldsymbol{y}]) \tag{8}$$

holds for typical initial conditions and observables $h(\boldsymbol{x}, \boldsymbol{y})$.

On the other hand one may take the stationary conditional distribution

$$\rho^{cond}(\boldsymbol{y}|\boldsymbol{x}) := \frac{\rho_*(\boldsymbol{x}, \boldsymbol{y})}{\bar{\rho}_*(\boldsymbol{x})} \quad , \tag{9}$$

where ρ_* and $\bar{\rho}_*$ denote the corresponding stationary densities. As we shall show, this choice leads to formally consistent results. Fortunately, the just mentioned two cases coincide in the limit of small ε provided some regularity assumptions are met. Taking eqs.(5) and (9) into account the condition of stationarity results in

$$0 = \mathcal{L}\rho_* = \frac{1}{\varepsilon}\bar{\rho}_* \mathcal{L}^{(0)} \rho^{cond} + \mathcal{L}^{(1)} \rho^{cond} \bar{\rho}_* \quad . \tag{10}$$

Assuming that all densities stay uniformly smooth in the limit $\varepsilon \to 0$, the first term has to vanish in the leading order of the small parameter ε. Therefore, ρ^{cond} tends towards the invariant density of $\mathcal{L}^{(0)}$ and finally coincides with $\rho_{\boldsymbol{x}}^{add}$.

If we now apply the standard projection operator technique to eq.(4) we obtain finally a closed equation of motion for the density (6) (namely eq.(A.8), cf. Appendix A for details). In lowest order perturbation theory this equation reduces to the Fokker–Planck equation (cf. Appendix B)

$$\frac{\partial \bar{\rho}_t}{\partial t} = -\frac{\partial}{\partial \boldsymbol{x}} \langle \boldsymbol{f} \rangle_{\boldsymbol{y}} \bar{\rho}_t(\boldsymbol{x}) + \frac{\partial}{\partial \boldsymbol{x}} \left(D(\boldsymbol{x}) \bar{\rho}_*(\boldsymbol{x}) \frac{\partial}{\partial \boldsymbol{x}} \frac{\bar{\rho}_t(\boldsymbol{x})}{\bar{\rho}_*(\boldsymbol{x})} \right) \quad . \tag{11}$$

Here the d_x–dimensional diffusion matrix is determined by the fluctuation of the vector field f,

$$\delta \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) - \langle \boldsymbol{f} \rangle_{\boldsymbol{y}} \tag{12}$$

in terms of the autocorrelation function as

$$D(\boldsymbol{x}) = \int_0^\infty dt \langle \delta \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\eta}_{\boldsymbol{x}}[t, \boldsymbol{y}]) \delta \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) \rangle_{\boldsymbol{y}} \quad . \tag{13}$$

The average $\langle \ldots \rangle_{\boldsymbol{y}}$ expresses according to the definition (8) an average over initial conditions $\boldsymbol{y}$. As long as the dynamics of eq.(2) is (exponentially) mixing the diffusion matrix is well defined since the correlations decay sufficiently fast. Furthermore, the diffusion matrix is of order ε, since this expansion parameter governs the time scale on which the correlation decays according to eq.(2).

Thus, in lowest order perturbation expansion the deterministic equations of motion reduce to a stochastic system with the Fokker–Planck equation being given by eq.(11). The drift consists of the $\boldsymbol{y}$–averaged vector field, whereas the diffusion is given by the autocorrelations of the fluctuations of the vector field. Both quantities are easily calculated by integrating the fast system (2) alone.

The resulting Fokker–Planck–equation is a self–consistent equation for the slow variables and can be converted into a Langevin equation. The latter is often easier to handle by numeric integration since time dependent solutions of Fokker–Planck equations are usually cumbersome to compute. Moreover, dynamical aspects such as stability of an individual solution with respect to perturbations of the initial condition are simpler to analyse on the basis of stochastic differential equations. If $D^{1/2}(\boldsymbol{x})$ denotes the canonical square root of the diffusion matrix, i.e. $D^{1/2}$ is the symmetric positive semi–definte matrix whose square yields the diffusion matrix (which by construction is positive semi–definite), then the corresponding Langevin equation using Stratonovich calculus reads [12]

$$\dot{\boldsymbol{x}}(t) = \langle \boldsymbol{f} \rangle_{\boldsymbol{y}} + D(\boldsymbol{x}) \frac{\partial}{\partial \boldsymbol{x}} \ln \bar{\rho}_*(\boldsymbol{x}) + D^{1/2}(\boldsymbol{x}) \left(\frac{\partial}{\partial \boldsymbol{x}} D^{1/2}(\boldsymbol{x}) \right) + D^{1/2}(\boldsymbol{x}) \boldsymbol{\xi}(t) \quad . \tag{14}$$

Here $\boldsymbol{\xi}$ denotes a Gaussian white noise normalised according to

$$\langle \xi_\nu(t) \xi_\mu(t') \rangle = 2\delta_{\nu\mu} \delta(t - t') \quad . \tag{15}$$

In the lowest nontrivial order in ε, eq.(14) reduces to a Langevin equation which contains only $\boldsymbol{x}$ and the noise process but not the invariant density $\bar{\rho}_*$. Hence, it can be treated numerically or even analytically quite straightforwardly.

3. Skew systems and averaging

The simplest class of systems consists of those where there is no feedback from the slow to the fast degrees of freedom, i.e. where $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{g}(\boldsymbol{y})$. Then the diffusion matrix (13) inherits its phase space dependence solely from the coupling of the fast degrees of freedom to the slow one, $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})$, but not from the internal fast dynamics $\boldsymbol{\eta}$. Hence an additive/multiplicative coupling of the fast degrees of freedom in $\boldsymbol{f}$ directly transfers to the coupling of the noise in the Langevin equation (14). Such a feature is clearly not present in non skew models discussed in the next section.

There are two prominent examples of elimination of fast degrees of freedom which behave regularly and yield a deterministic reduced dynamical system (cf. e.g. [2]). Both of these cases are included in our analysis too. The first case, often refered to as "slaving" corresponds to a fast dynamical system (2) relaxing to a fixed point. Here a full discussion which goes to arbitray order in perturbation theory can be supplied (cf. Appendix C). The second case, often referred to as averaging, corresponds to a periodc fast dynamics with $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{g}(\boldsymbol{y})$, so that the period does not depend on $\boldsymbol{x}$. The approach adopted here is not well taylored to this case since the autocorrelation entering the diffusion matrix (13) is periodic and the Markov approximation needs a more thorough justification. Nevertheless, the diffusion matrix still has a meaning in the distributional sense§. Above all, the diffusion matrix coincides with the zero frequecy contribution in the spectrum of the periodic quantity (12). Since the fluctuation does not contain a constant contribution, the diffusion matrix vanishes. Therefore no stochastic force appears in the Langevin equation (14). The deterministic drift is given by the

§ From the technical point of view one may include a factor $\exp(-\alpha t)$ in the intergal eq.(13) and consider the limit $\alpha \to 0+$.

time average of the slow dynamics since the density $\bar{\rho}_*$ is in general nonsingular. Thus, the result coincides to leading order with the usual averaging procedure.

4. Stochastic vs. deterministic resonance

Let us now come back to the case where the fast dynamics is chaotic, so that we expect a finite diffusion in the reduced dynamics. As an example, we study the motion of a damped particle in a double well potential driven by a small periodic forcing. One component of a fast chaotic Lorenz model is added to the slow equation of motion.

$$\begin{aligned}
\dot{x} &= x - x^3 + a\cos(\omega t) + \kappa y_1 \\
\dot{y}_1 &= \frac{\sigma}{\varepsilon}(y_2 - y_1) \\
\dot{y}_2 &= \frac{1}{\varepsilon}(r y_1 - y_1 y_3 - y_2) \\
\dot{y}_3 &= \frac{1}{\varepsilon}(y_1 y_2 - b y_3)\ .
\end{aligned} \tag{16}$$

We choose the canonical parameters, σ=10, r=28, b=8/3, and $\omega = 0.001$, $a = 0.1$. Due to the lack of feedback from the slow to the fast variables and due to the additive nature of the coupling of the fast to the slow variables, the diffusion constant D of eq.(13) is independent of x. If we denote by $\langle y_1(t')y_1(t'+t)\rangle_{t'}$ the auto–correlation function of the Lorenz–y_1–coordinate for $\varepsilon = 1$, then the diffusion constant D in the Fokker–Planck equation corresponding to eq.(16) is given by

$$D = \varepsilon\kappa^2 \int_0^\infty \langle y_1(t')y_1(t'+t)\rangle_{t'}\, dt \quad . \tag{17}$$

The effective description in terms of the Langevin equation (14) in order $\sqrt{\varepsilon}$ for the x–dynamics leads to the well known classical scenario of stochastic resonance (for an excellent review on stochastic resonance see [14]). With $\kappa = 0$ and $a = 0$, the particle x settles down in one of the two wells of the potential, i.e. $x(t) \to \pm 1$. For $0 < a \ll 1$, the periodic forcing is still so weak that it cannot induce any transitions from one well to the other. For $\kappa > 0$ and hence $D > 0$ we expect a finite probability for such a jump from one well to the other. It is approximately given by the Kramers rate $1/T_K = \exp(-\Delta/D)$ where $\Delta = 1/4$ is the height of the barrier between the two wells, and D measures the noise amplitude. It is known that there occurs a maximal synchronization between the jumps and the periodic forcing, if the period of the forcing and the Kramers time fulfill the condition $2\pi/\omega = 2T_K$.

If $\varepsilon \ll 1$ the theory developed in the last section tells us that the Lorenz–y_1–coordinate should in fact act like noise. In Figure 1 we show numerical results of the full system eq.(16) in comparison to white noise driven motion in the double well, i.e. a numerical simulation of the corresponding Langevin equation‖. We compute the

‖ In a naive numerical simulation, white noise is realized as a piecewise constant function with step widths δt, where δt is the time step of the Euler–integrator. Consequently, the auto–correlation function of this numerical noise is $c(\tau) = \sigma^2$ if $\tau \in [-\delta t/2, \delta t/2]$ and 0 elsewhere, where σ^2 is the variance of correlation between the coordinate $x(t)$ and the periodic driving

$$c = \sqrt{\langle x(t)\cos\omega t\rangle_t^2 + \langle x(t)\sin\omega t\rangle_t^2} \tag{18}$$

and the phase difference between driving force and the well-hopping,

$$\phi = \arctan\frac{\langle x(t)\sin\omega t\rangle_t^2}{\langle x(t)\cos\omega t\rangle_t^2}\ . \tag{19}$$

In the numerical simulation, we fix ε and tune κ to produce the curves shown in Figure 1.

The comparison of the results shows that for $\epsilon < 0.1$, there is within the statistical accuracy a perfect agreement between the noise driven system and the Lorenz driven system. For larger ϵ, i.e. less pronounced time scale separation, deviations of the curves from the noise limit are evident. The onset of synchronous well–hopping is shifted towards larger equivalent noise amplitudes, and the correlation between periodic forcing and hopping is slightly more pronounced. The first comes from the clear non–Gaussian character of the short time accumulation of y_1. The tails are finite as can be seen from the inset in Figure 2.

This example shows impressively the benefits of stochastic modelling. The ad hoc complete elimination of the fast degrees of freedom is here equivalent with the zeroth order approximation in ε. However, it destroys the relevant dynamical feature of hopping between the wells, whereas our properly derived noise process which is first order in $\sqrt{\varepsilon}$ produces quantitatively correct hopping rates and synchronisation effects. Thus the invariant measure in the extended phase space, i.e. including the driving force, is correctly reproduced (cf. also [7, 11]).

Only for $\varepsilon \geq 0.1$, one needs to go beyond the first order in $\sqrt{\varepsilon}$ here. In order to see of what nature the higher order corrections might be, we created so called surrogate data of the Lorenz y_1–coordinate. A sequence of random numbers was created which has an identical autocorrelation function and marginal probability distribution as the original sequence of the Lorenz system but no nontrivial higher order correlations [15]. It reflects thus a non–Gaussian, linearly correlated noise process. Replacing the fast Lorenz-system by these surrogates and considering $\varepsilon = 1$, we find indeed a shift of the onset of the resonance to higher noise levels. But we do not reproduce the increased value of correlation between hopping and periodic driving of the double well (cf. Figure 2). We therefore conclude that the nontrivial distribution (cf. insert of Figure 2) is only one source of the differences. It can be taken into account by a non–Gaussian noise process. We should recall that the linear correlations of the fast variable can be fully absorbed in the diffusion constant, so that alternatively to the sophisticated surrogates one can use a simple scrambling of the fast deterministic variable (conserving their nontrivial distribution but destroying all of their temporal correlations) and feed this data instead of ξ into the Langevin equation. In this case, correlations and distributions are considered separately. The remaining discrepancy has to be caused by higher order

the random numbers. Hence, the diffusion coefficient of this process is $1/2\delta t\sigma^2$. Of course, in a proper integration scheme, this $\sqrt{\delta t}$-factor is correctly respected.

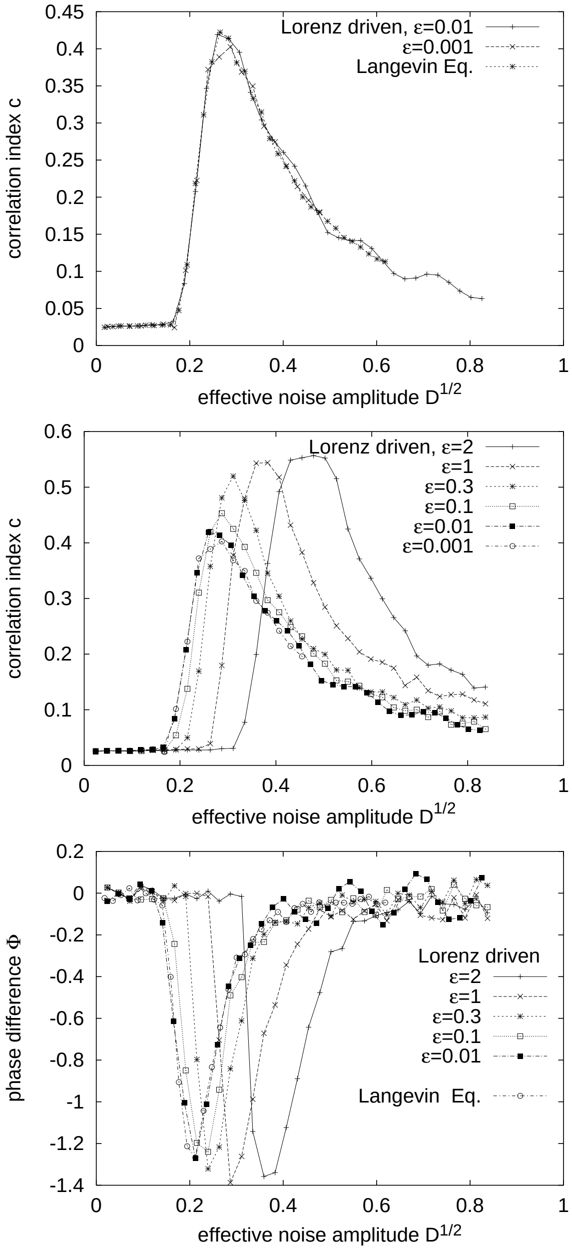

Figure 1. Upper panel: Correlation c between $x(t)$ and the periodic driving for sufficiently strong time scale separation compared to the numerical integration of the Langevin eq.(14). The "effective noise amplitude" is, for the deterministic simulation, obtained according to eq.(17).

Central panel: The convergence of the behaviour of c as a function of the time scale separation ε.

Lower panel: The phase shift ϕ between periodic driver and response of $x(t)$, comparing the Lorenz driven system to the Langevin dynamics. In [14] approximate analytic descriptions of both c and ϕ are given.

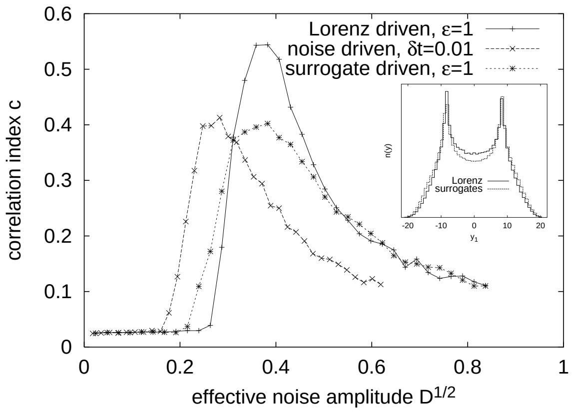

Figure 2. Curve as in Figure 1 for white noise, for the Lorenz driven motion with ε=1, and for surrogate driven motion, also with ε=1. The inset shows the distribution of the Lorenz data and the surrogates.

correlations. A proper treatment of this feature would require much more sophisticated stochastic models. Fortunately such correlations are only relevant if the time scale separation is poor.

We finally have to add however a reservation concerning topological constraints. Such constraints may appear if the slow subsystems displays relaxatory dynamics, e.g. if it is one–dimensional. Then more sophisticated considerations have to be used. If one considers the model (16) then as explained above the corresponding stochastic description yields a simple one–dimensional stochastic differential equation which of course coincides with the famous Kramers problem if the periodic driving is neglected. In such cases hopping between the potential wells is triggered by the Gaussian nature of the noise which allows for arbitrarily large amplitudes. These large amplitudes are indeed required to induce transitions between the wells. That property already holds for the full system (16). Thus our stochastic description (11) and (14) which we have obtained in lowest order perturbation expansion is correct only if the amplitudes of the fast system are large enough. If they are too small then all higher orders of the master equation have to be taken into account (cf. eq.(A.8)) rendering the analysis much more complicated. In fact, this particular feature yields the main difference between the treatment of statistical mechanics and the analysis of low dimensional dynamical systems. In statistical mechanics large amplitudes can be always achieved because on takes the thermodynamic limit of the heat bath. Here we cannot guarantee for these amplitudes in advance, and one has to check such a condition in each case separately. However, if no topological constraint appears then the approximation by Gaussian noise does not pose a problem as shown in this section.

5. Non skew systems and intrinsic noise bifurcation

The general situation of a nonlinear coupling between fast and slow degrees of freedom is too difficult to be treated in generality. We discuss here a particular effect, namely fast degrees of freedom whose chaoticity depends on the values of the slow variables. Whenever the fast variables are chaotic, the diffusion constant $D(x)$ is non–vanishing and the fast variables act as noise. When they behave periodically, the periodicity of the auto–correlation function finally integrates up to zero, and no noise is visible. We combine these two features by considering a fast equation of motion whose parameters are modulated by the slow variables in such a way that a change from regular to chaotic motion can be triggered. For that purpose we consider two Lorenz systems. The slow one with the variables x_1, x_2, x_3 and standard parameter values is driven by the fast variable y_1 additively. The fast system with y_1, y_2, y_3 is driven by the slow one through the parameter $r_y(x_1) = 210 + 60\arctan(2x_1)/\pi$. In order to simplify the treatment we have chosen a sigmoid function $r_y(x_1)$ for the feedback which acts as a kind of switch.

$$\begin{aligned}
\dot{x}_1 &= 10(x_2 - x_1) + \kappa y_1 \\
\dot{x}_2 &= 28x_1 - x_1x_3 - x_2 \\
\dot{x}_3 &= x_1x_2 - \frac{8}{3}x_3 \\
\dot{y}_1 &= \frac{10}{\varepsilon}(y_2 - y_1) \\
\dot{y}_2 &= \frac{1}{\varepsilon}(r_y(x_1)y_1 - y_1y_3 - y_2) \\
\dot{y}_3 &= \frac{1}{\varepsilon}(y_1y_2 - \frac{8}{3}y_3)\ .
\end{aligned} \tag{20}$$

A projection of the attractor onto the x_1–x_2–plane is shown in Figure 3, whereas Figure 4 displays the corresponding time trace of y_1 and r_y. The time scale relation was set to $\varepsilon = 2 \cdot 10^{-3}$ and the coupling strength to $\kappa = 10$.

The asymmetry of the attractor shape is related to the fact that the mean value of y_1 is non–zero for both the chaotic and periodic motion. Moreover, if $r_y \approx 240$, the trajectory of the fast system settles after some chaotic transient on either of two coexisting pairwise symmetric limit cycles. Despite the fact that chaotic and limit cycle oscillations have almost identical amplitude and the same time scale, their effect on the slow variables is evidently different. In the case of chaotic oscillations, the fast variable effectively acts as noise. However for the limit cycle behaviour, i.e. if $r_y \approx 240$, then due to the averaging theorem the stochastic character of the fast variable disappears in good approximation. Eq.(20) demonstrates that through the coupling of the slow dynamics to the fast degrees of freedom the nature of the effective noise changes considerably if bifurcations within the fast system appear.

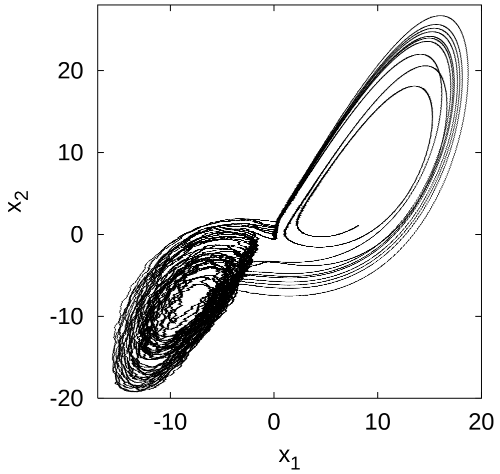

Figure 3. The projection of the attractor of Eq.(20) into the x_1-x_2-plane. On the left part of the attractor, $r_y(t) \approx 180$, so that the fast motion is chaotic, whereas on the right part, $r_y(t) \approx 240$, so that it is periodic (limit cycle).

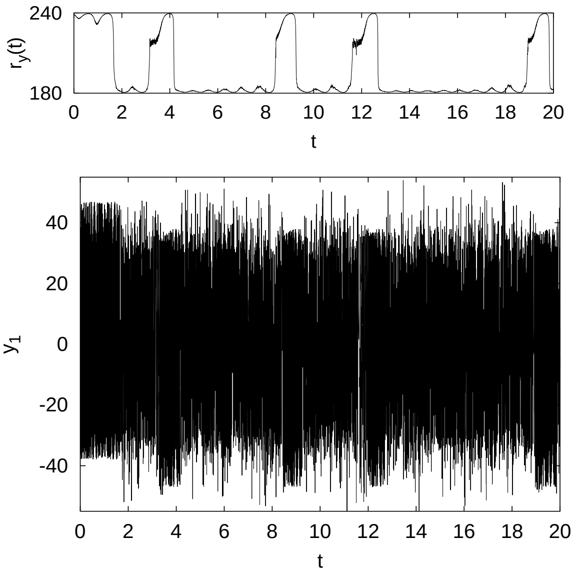

Figure 4. The parameter $r_y(x_1)$, as a function of time (upper panel) and the fast variable y_1 driving the slow Lorenz (lower panel) during the same time interval.

6. Conclusions

If a system can be decomposed into a slow and a fast subsystem, the time evolution of the phase space density of the slow variables can be described by an effective equation of motion. In lowest order of time scale separation such an equation reduces to a Fokker–Planck equation where the drift and diffusion coefficients are given explicitly in terms of averages over the fast dynamics. This result extends the usual averaging theorem for periodic modulation to chaotic degrees of freedom. Our features are in agreement with generalised central limit theorems for chaotic systems [6] which state that chaotic degrees of freedom can accumulate to a Gaussian noise process.

If the time scale separation becomes less pronounced we still can write down a closed non–Markovian equation of motion for the density of the slow variables. Expanding in terms of the time scale separation we observe that the exact stationary density is preserved in each order, so that even an analysis beyond the simple Fokker–Planck approach seems to be promising. Of course the simple relation to stochastic differential equations gets lost, since these higher order contributions are no longer of Fokker–Planck type. Thus, we expect that modelling of fast degrees of freedom by Gaussian noise is limited to the case of well pronounced time scale separation. This conjecture is supported by our numerical analysis. We have found that a modelling by an adapted non–Gaussian noise which takes the distribution of the chaotic process into account is not sufficient. Temporal higher order correlations become important if one wants to cope with all features of the full dynamical system. In [9] a particular higher order effect was found, which in our setting, however, seems less relevant, since we are dealing with time continuous systems. However, translated into our language, it means: In a skew system, the fast dynamics can stay arbitrarily long time spans arbitrarily close to an unstable periodic orbit. Such regular episodes can potentially create strong non–stochastic effects on the slow dynamics, and since they can be arbitrarily long, they might appear (with corresponding low probability) for every finite ϵ.

Last but not least we emphasise that contrary to applications in statistical mechanics the slow variables may have dramatic effects on the fast dynamics, since no thermodynamic limit for a heat bath is involved. We have studied this effect in a simple model system showing that the influence of the fast degrees of freedom may change considerably due to the dynamical coupling of the slow variables. It would be interesting to find other physical models which display this novel type of instability.

Appendix A. Mori–Fujisaka–Shigematsu expansion

Within statistical mechanics there exists a well developed formally exact expansion of the master equation [13]. We are adopting here a variant of this approach to derive an equation of motion for the distribution of slow variables (6). In order to keep the presentation self contained we recall these steps in some detail for the special system given by eqs.(1) and (2).

Using the notation $\mathrm{Tr}_{\boldsymbol{y}}\{\ldots\} = \int d\boldsymbol{y}\ldots$ for the phase space integral with respect to fast variables we introduce the projection operator

$$\mathcal{P}\ldots := \rho^{cond}\mathrm{Tr}_{\boldsymbol{y}}\{\ldots\} \tag{A.1}$$

which generates the reduced density (7). Applying this projection to the full equation of motion (4) one obtains the Nakajima–Zwanzig equation [10]

$$\frac{\partial \rho_t^{red}}{\partial t} = -\mathcal{P}\mathcal{L}\rho_t^{red} + \int_0^t ds\, \mathcal{P}\mathcal{L}\mathcal{Q}e^{-\mathcal{Q}\mathcal{L}s}\mathcal{Q}\mathcal{L}\rho_{t-s}^{red} \tag{A.2}$$

where $\mathcal{Q} := \mathcal{I} - \mathcal{P}$ denotes the orthogonal projection. For the initial condition the constraint $\mathcal{Q}\rho^{rel}_{t=0} = 0$ has been adopted. Employing the definitions (5), (9), and (A.1) and using the stationarity of ρ_* we are left with¶

$$\frac{\partial \bar{\rho}_t}{\partial t} = -\frac{\partial}{\partial \boldsymbol{x}} \mathrm{Tr}_{\boldsymbol{y}}\{\boldsymbol{f}\rho^{cond}\}\bar{\rho}_t(\boldsymbol{x}) + \frac{\partial}{\partial \boldsymbol{x}}\int_0^t ds\, \mathrm{Tr}_{\boldsymbol{y}}\{\Delta\boldsymbol{f} e^{-\mathcal{Q}\mathcal{L}s}\Delta\boldsymbol{f}\rho_*\}\frac{\partial}{\partial \boldsymbol{x}}\frac{\bar{\rho}_{t-s}(\boldsymbol{x})}{\bar{\rho}_*(\boldsymbol{x})} \quad , \tag{A.3}$$

where

$$\Delta\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y}) := \boldsymbol{f}(\boldsymbol{x},\boldsymbol{y}) - \mathrm{Tr}_{\boldsymbol{y}}\{\boldsymbol{f}\rho^{cond}\} \tag{A.4}$$

denotes the fluctuation with respect to the conditional stationary density (9). Eq.(A.3) constitutes almost the Fokker–Planck type equation we are dwelling on. But at the moment the integral kernel is still an operator and we are going to evaluate its action on the distribution $\bar{\rho}_t$ explicitly.

Employing the formally adjoint operators

$$\mathcal{P}^\dagger \ldots = \mathrm{Tr}_{\boldsymbol{y}}\{\ldots\rho^{cond}\}, \quad \mathcal{L}^\dagger = -\frac{1}{\varepsilon}\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y})\frac{\partial}{\partial \boldsymbol{y}} - \frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y}) \tag{A.5}$$

we obtain for arbitrary functions $G(\boldsymbol{x},\boldsymbol{y})$ and $H(\boldsymbol{x},\boldsymbol{y})$

$$\begin{aligned} &\mathrm{Tr}_{\boldsymbol{y}}\{G\cdot e^{-\mathcal{Q}\mathcal{L}s}H\} - \mathrm{Tr}_{\boldsymbol{y}}\{e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger s}G\cdot H\} \\ &= \int_0^s ds_1 \frac{d}{ds_1}\mathrm{Tr}_{\boldsymbol{y}}\{e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger(s-s_1)}G\cdot e^{-\mathcal{Q}\mathcal{L}s_1}H\} \\ &= -\frac{\partial}{\partial \boldsymbol{x}}\int_0^s ds_1\, \mathrm{Tr}_{\boldsymbol{y}}\{\boldsymbol{f}\mathcal{Q}^\dagger e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger(s-s_1)}G\cdot e^{-\mathcal{Q}\mathcal{L}s_1}H\} \quad , \end{aligned} \tag{A.6}$$

where in the last step the definition (5) has been employed. Iterating eq.(A.6) we are left with

$$\begin{aligned} \mathrm{Tr}_{\boldsymbol{y}}\{G\cdot e^{-\mathcal{Q}\mathcal{L}s}H\} &= \mathrm{Tr}_{\boldsymbol{y}}\{e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger s}G\cdot H\} \\ &- \frac{\partial}{\partial \boldsymbol{x}}\int_0^s ds_1\, \mathrm{Tr}_{\boldsymbol{y}}\{e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger s_1}\boldsymbol{f}\mathcal{Q}^\dagger e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger(s-s_1)}G\cdot H\} \\ &+ \left(-\frac{\partial}{\partial \boldsymbol{x}}\right)^2\int_0^s ds_1\int_0^{s_1} ds_2 \\ &\quad \mathrm{Tr}_{\boldsymbol{y}}\{e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger s_2}\boldsymbol{f}\mathcal{Q}^\dagger e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger(s_1-s_2)}\boldsymbol{f}\mathcal{Q}^\dagger e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger(s-s_1)}G\cdot H\} \\ &+ \ \ldots \quad . \end{aligned} \tag{A.7}$$

Using the expansion (A.7) in eq.(A.3) we obtain the final result

$$\frac{\partial \bar{\rho}_t}{\partial t} = -\frac{\partial}{\partial \boldsymbol{x}}\mathrm{Tr}_{\boldsymbol{y}}\{\boldsymbol{f}\rho_*\}\frac{\bar{\rho}_t(\boldsymbol{x})}{\bar{\rho}_*(\boldsymbol{x})} + \sum_{n=0}^{\infty}(-1)^n\left(\frac{\partial}{\partial \boldsymbol{x}}\right)^{n+1}\int_0^t ds\, \Gamma(s,x)\frac{\partial}{\partial \boldsymbol{x}}\frac{\bar{\rho}_{t-s}(\boldsymbol{x})}{\bar{\rho}_*(\boldsymbol{x})} \quad , \tag{A.8}$$

where

$$\begin{aligned} \Gamma_0(s,\boldsymbol{x}) \quad &:= \mathrm{Tr}_{\boldsymbol{y}}\{e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger s}\Delta\boldsymbol{f}\cdot\Delta\boldsymbol{f}\rho^{cond}\} \\ \Gamma_n(s,\boldsymbol{x}) \quad &:= \int_0^s ds_1\ldots\int_0^{s_{n-1}} ds_n \\ &\quad \mathrm{Tr}_{\boldsymbol{y}}\{e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger s_n}\boldsymbol{f}\mathcal{Q}^\dagger e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger(s_n-s_{n-1})}\ldots\boldsymbol{f}\mathcal{Q}^\dagger e^{-\mathcal{L}^\dagger\mathcal{Q}^\dagger(s-s_1)}\Delta f\cdot\Delta\boldsymbol{f}\rho^{cond}\} \end{aligned} \tag{A.9}$$

¶ Vector notation has to be understood in the sense that products between $\partial/\partial\boldsymbol{x}$ and $\boldsymbol{f}$ are scalar products. All other products are wedge products.

denote the tensorial integral kernels. It is worth to mention that the exact stationary distribution $\bar{\rho}_*$ fulfils eq.(A.8) term by term since stationarity ensures that the average of the slow vector field has neither sources nor sinks

$$\frac{\partial}{\partial \boldsymbol{x}} \mathrm{Tr}_{\boldsymbol{y}}\{\boldsymbol{f}\rho_*\} = -\mathrm{Tr}_{\boldsymbol{y}}\{\mathcal{L}\rho_*\} = 0 \quad . \tag{A.10}$$

Appendix B. Perturbation expansion

Employing the small parameter in eq.(5) the expressions (A.9) may be evaluated by a perturbation expansion. Using

$$\mathcal{L}^{(0)\dagger}\mathcal{P}^\dagger = 0, \quad \mathcal{L}^\dagger\mathcal{Q}^\dagger = \frac{1}{\varepsilon}\mathcal{L}^{(0)\dagger} + \mathcal{L}^{(1)\dagger}\mathcal{Q}^\dagger \tag{B.1}$$

and neglecting the small contribution $\mathcal{L}^{(1)\dagger}\mathcal{Q}^\dagger$ in the operator exponential eq.(A.9) yields to leading nonvanishing order

$$\Gamma_0(s, \boldsymbol{x}) = \mathrm{Tr}_{\boldsymbol{y}}\{e^{-\mathcal{L}^{(0)\dagger}s/\varepsilon}\delta\boldsymbol{f}\cdot\delta\boldsymbol{f}\rho^{add}_{\boldsymbol{x}}\} + \mathcal{O}(\varepsilon) \quad . \tag{B.2}$$

Here we have taken into account that the density ρ^{cond} is to leading order identical to $\rho^{add}_{\boldsymbol{x}}$. Since the correlation given by eq.(B.2) decays on a time scale of order ε we may replace the density $\bar{\rho}_{t-s}$ in the corresponding integral of eq.(A.8) by $\bar{\rho}_t$ and may extend the remaining integral to infinity (Markov approximation). For the same reason the kernels appearing in the higher derivatives are of higher order

$$\Gamma_n(s, \boldsymbol{x}) = \mathcal{O}(\varepsilon^n) \tag{B.3}$$

and will be neglected. Thus we arrive at the Fokker–Planck equation (11).

Appendix C. Slow manifold

Eq.(A.8) constitutes a formally exact result of the full evolution equation (4) where no approximation was involved, apart from a constraint on the initial condition. Our previous considerations have shown that within a straightforward perturbation expansion the Fokker–Planck equation (11) can be derived. Such a result follows in a different manner too, if the projection operator (A.1) were based on $\rho^{add}_{\boldsymbol{x}}$ instead of ρ_{cond}. However, such a choice somehow anticipates the perturbation expansion in the definition of the projection operator. In fact, higher order terms differ and might be even formally inconsistent. Such features are already well known from statistical mechanics where there has been a long lasting debate how to include the interaction with a heat bath appropriately in the projection operator.

We will illustrate these features by considering a special case of eqs.(1) and (2) namely when there exists an invariant attracting manifold determined by $\boldsymbol{y} = \boldsymbol{\eta}_*(\boldsymbol{x})$. First of all, the condition of invariance requires

$$\frac{1}{\varepsilon}\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{\eta}_*(\boldsymbol{x})) = D\boldsymbol{\eta}_*(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\eta}_*(\boldsymbol{x})) \quad , \tag{C.1}$$

whereas the motion on the manifold is given by

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\eta}_*(\boldsymbol{x})) \quad . \tag{C.2}$$

If $\bar{\rho}_*(\boldsymbol{x})$ denotes the invariant distribution of eq.(C.2) then the full stationary density is given by

$$\rho_*(\boldsymbol{x}, \boldsymbol{y}) = \delta(\boldsymbol{y} - \boldsymbol{\eta}_*(\boldsymbol{x}))\bar{\rho}_*(\boldsymbol{x}) \quad . \tag{C.3}$$

The latter expression is confirmed easily by evaluating the condition of invariance $\mathcal{L}\rho_* = 0$ taking eqs.(C.1) and (C.2) into account. Hence, the conditional stationary density (9) reads

$$\rho^{cond}(\boldsymbol{y}|\boldsymbol{x}) = \delta(\boldsymbol{y} - \boldsymbol{\eta}_*(\boldsymbol{x})) \quad . \tag{C.4}$$

Now the full drift coefficient of the Master equation (A.8) is given by

$$\mathrm{Tr}_{\boldsymbol{y}}\{\boldsymbol{f}\rho^{cond}\} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\eta}_*(\boldsymbol{x})) \tag{C.5}$$

in complete accordance with eq.(C.2). By virtue of eqs.(C.4) and (A.4)

$$\Delta\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})\rho^{cond}(\boldsymbol{y}|\boldsymbol{x}) \equiv 0 \tag{C.6}$$

holds, so that all diffusion kernels (A.9) vanish identically. Hence, the full Master equation(A.8) reduces to the deterministic system (C.2) on the slow manifold. In this sense the drift given by the weight ρ^{cond} is a renormalised drift to infinite order of perturbation theory. Such a property depends crucially on the appropriate choice for the projection operator.

References

[1] E. Fick and G. Sauermann, *The quantum statistics of dynamic processes*, (Springer, Berlin, 1990); R. Kubo, M. Toda, and N. Hashitsume, *Statistical physics: Nonequilibrium statistical mechanics*, (Springer, Berlin, 1991)

[2] N. G van Kampen, *Elimination of fast variables*, Phys. Rep. **124**, 69–160 (1985).

[3] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, (Springer, New–York, 1993)

[4] K. Hasselmann, *Stochastic climate models: Part I. Theory.*, Tellus **28**, 473–485 (1976).

[5] L. Arnold, *Hasselmann's program revisited: The analysis of stochasticity in deterministic climate models*, in: P. Imkeller, J.–S. von Storch, eds., *Stochastic Climate Models*, (Birkhäuser, Boston, 2001)

[6] C. Beck, G. Roepsdorff, and C. Schroer, *Driven chaotic motion in single– and double well potentials*, Physica **D 72**, 211–229 (1994).

[7] M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, 2nd edition, (Springer, New York, 1998)

[8] Y. Kifer, *Averaging in dynamic–systems and large deviations*, Invent. Math. **110**, 337–370 (1992).

[9] C. Beck, *From the Perron–Frobenius Equation to the Fokker–Planck Equation*, J. Stat. Phys. **79** 875–894 (1995).

[10] R. Zwanzig, *Ensemble method in the theory of irreversibility*, J. Chem. Phys. **33**, 1338–1341 (1960).

[11] C. Rödenbeck, C. Beck, and H. Kantz, *Dynamical systems with time scale separations: Averaging, stochastic modelling, and central limit theorems*, in: P. Imkeller, J.–S. von Storch, eds., *Stochastic Climate Models*, (Birkhäuser, Boston, 2001)

[12] H. Risken, *The Fokker–Planck equation*, (Springer, Berlin, 1989)

[13] H. Mori, H. Fujisaka, and H. Shigematsu, *A new expansion of the master equation*, Prog. Theor. Phys. **51**, 109–122 (1974).

[14] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, *Stochastic Resonance*, Rev. Mod. Phys. **70**, 223–287 (1998).

[15] T. Schreiber and A. Schmitz, *Improved surrogate data for nonlinearity tests*, Phys. Rev. Lett. **77**, 635–638 (1996).