Langevin equation for slow degrees of freedom of Hamiltonian systems

RS MacKay
Mathematics Institute and Centre for Complexity Science,
University of Warwick, Coventry CV4 7AL, UK

August 26, 2008

Abstract

A way is sketched to derive a Langevin equation for the slow degrees of freedom of a Hamiltonian system whose fast ones are mixing Anosov. It uses the Anosov-Kasuga adiabatic invariant, martingale theory, Ruelle's formula for weakly non-autonomous SRB measures, and large deviation theory.

1 Introduction

Model reduction is a central theme in science. In particular, it is common to propose to replace "inessential" details of parts of dynamics by noise. This paper addresses the question of to what extent such reduction may be justified if one starts from a deterministic Hamiltonian systems, the agreed foundation for all classical mechanics.

Suppose a Hamiltonian system consists of some slow degrees of freedom coupled to a large number of fast chaotic degrees of freedom. For example, consider the conformation change degrees of freedom of a biomolecule coupled to its vibrations and movement of water molecules. It is standard to model the slow degrees of freedom by a Langevin equation, that is a stochastic differential equation where the effects of the fast degrees of freedom have been replaced by an effective Hamiltonian system, damping and noise.

Despite being in use now for 100 years [La], it seems to me there is not yet a satisfactory derivation of a Langevin equation for the slow degrees of freedom. See the reviews [GKS, Ki] for the state of affairs. An example system which has recently resurfaced is the "piston problem" [Li].

One precursor is [FKM] for the case of Hamiltonians which are quadratic in the fast degrees of freedom and subject to a continuum approximation for the distribution of frequencies of their normal modes. This approach was presented nicely by [Z], and treated more throughly in [So], but the crucial assumption (for the analysis) of harmonic fast degrees of freedom seems unrealistic and overly restrictive to me.

A second precursor is the line of investigation of the sequence of papers [O, Wi, BR, J], which depends on assuming ergodicity of the fast dynamics restricted to an energy level for frozen slow degrees of freedom, but I don't think these papers can be considered to make a mathematically complete derivation of a Langevin equation. As a side-remark, ergodicity may be considered unrealistic and overly restrictive, but perhaps true ergodicity is not really required.

This paper seeks to put more mathematical flesh on the ideas of the second approach, under the stronger assumption that the fast dynamics is mixing Anosov on relevant energy levels for frozen slow degrees of freedom. Although this is a yet more restrictive assumption than ergodicity, [GC] advocated that it may be reasonable to assume that the dynamics of a large Hamiltonian system act as if mixing Anosov on relevant energy levels (the "chaotic hypothesis") and that results derived under this assumption may apply more widely. It is worth noting that a two-degree of freedom physically relevant mixing Anosov system has been constructed [HM] and I believe analogues can be made in higher degrees of freedom.

Under the mixing Anosov assumption, some nice mathematical results can be applied, namely a formula of Ruelle for the first-order effect of weak non-autonomy on the natural measure for a chaotic system [Ru], and an almost sure invariance principle to approximate the integral of a zero-mean vector of observables for a chaotic system by a multidimensional Brownian walk [MN]. With some work I believe that these ingredients can be put together to derive a Langevin equation for the slow degrees of freedom. The paper sketches the main lines of the proposal.

2 Assumptions

1. Suppose (M, ω) is a symplectic manifold of large dimension $2m$ and $H : M \to \mathbb{R}$ is a smooth function (the Hamiltonian). Together they define a Hamiltonian vector field X_H on M, by $\omega(X_H, \xi) = dH(\xi)$ for all tangent vectors ξ. Equivalently, the symplectic form ω defines a Poisson bracket $\{F, G\} = \omega(X_F, X_G)$ on smooth functions $F, G : M \to \mathbb{R}$, and then $dF(X_H) = X_H(F) = \{F, H\}$ for all $F : M \to \mathbb{R}$. Let ϕ_t be the flow of X_H.

2. Suppose N is a symplectic manifold of moderate dimension $2n$, representing the slow degrees of freedom. It will be enough to do local analysis in N, so without loss of generality it can be considered to be a piece of $\mathbb{R}^{2n}$ with local coordinate functions Z_j. Suppose $\pi : M \to N$ is a Poisson map, i.e. for all smooth functions $F, G : N \to \mathbb{R}$, $\{F \circ \pi, G \circ \pi\}_M = \{F, G\}_N \circ \pi$.

3. Suppose for each $Z \in N$, the "fibre" $\pi^{-1}(Z)$ is a symplectic submanifold of M, i.e. the restriction of ω to its tangent space is everywhere non-degenerate. It follows that the restriction H_Z of H to $\pi^{-1}(Z)$ defines a Hamiltonian vector field X_{H_Z} on $\pi^{-1}(Z)$ (the constrained vector field). Furthermore, X_{H_Z} preserves the $2(m-n)$-dimensional volume form
$$\Omega_Z = \omega^{\wedge(m-n)}/(m-n)!$$
on $\pi^{-1}(Z)$, the level sets $(H, \pi)^{-1}(E, Z)$, and the $(2m-2n-1)$-dimensional volume form $\mu_{Z,E}$ on $(H, \pi)^{-1}(E, Z)$ defined by $\mu \wedge dH = \Omega_Z$. Denote its flow by ψ_t.

4. Suppose
$$W_Z(E) := \int_{\{H \le E\}} \Omega_Z$$
is finite for relevant Z, E and differentiable (actually, $W_Z(E)$ finite is not really necessary as long as its derivative with respect to Z and E is definable). Its E-derivative can be written as
$$W_Z'(E) = \int_{H^{-1}(E)} \mu_{Z,E}.$$
Then $\mu_{Z,E}$ induces an invariant probability measure
$$\lambda_{Z,E}(A) = \frac{1}{W_Z'(E)} \int_A \mu_{Z,E}$$
for subsets A on the level set for X_{H_Z}, which Boltzmann called an "ergode" and Gibbs a "microcanonical emsemble".

5. Suppose
$$V_j = \{Z_j \circ \pi, H\}$$
are slow compared to X_{H_Z}, in a sense to be made more precise in the next assumption.

6. Suppose X_{H_Z} is mixing Anosov on the level sets for relevant E, Z. In particular, suppose the autocorrelation of the deviation $v(s)$ of $V \circ \psi_s$ from its mean decays in a short time $\varepsilon \ll 1$ on the "slow" timescale, that for significant change in Z under the mean of V.

7. Suppose the size of v (in units for the slow timescale) scales like $\sqrt{\delta/\varepsilon}$ for some δ.
8. Suppose the temperature $T > 0$ is significantly higher than a value determined by δ and the symplectic form on N, to be made explicit in Section 4.3.
9. Finally, suppose the heat capacity per fast degree of freedom is positive and bounded.

3 Aim

The aim is to show that for ε small the distribution of paths $\pi \circ \phi_t(Y)$ for random Y with respect to λ_{Z_0,E_0} and for t in any order-one time interval, is close to that for the solutions of a system of stochastic ordinary differential equations

$$dZ_i = (J_{ij} - \beta D_{ij})\nabla_j F \; dt + \sigma_{ik} \; dW_k, \quad Z(0) = Z_0, \tag{1}$$

with $\nabla_j F$ a shorthand for $\frac{\partial F}{\partial Z_j}$, J representing the Poisson bracket on N (i.e. $\{F, G\} = \nabla_i F J_{ij} \nabla_j G$), β the inverse temperature for the initial energy, $F : N \to \mathbb{R}$ the free energy function at the given temperature,

$$D_{ij} = \int_{-\infty}^{0} \lambda(v_i(0)v_j(s)) \; ds \tag{2}$$

W a multidimensional Wiener process, σ any matrix function satisfying the Einstein-Sutherland relation $\sigma\sigma^T = D + D^T$, and the Klimontovich interpretation of the stochastic differential equation.

Definitions of β, F and the Klimontovich interpretation will be recalled at the appropriate points.

Equation (1) with Klimontovich interpretation implies Klein-Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = -\mathrm{div}\,(\rho(J - \beta D)\nabla F - S\nabla\rho) \tag{3}$$

for the evolution of probability densities ρ with respect to volume $\wedge dZ_j$ on N, where $S = (D + D^T)/2$, the symmetric part of D.

4 Strategy

The strategy of proof has six stages. First $\lambda(V)$ is evaluated in terms of the microcanonical free energy. Secondly, the fluctuations $v(t)$ of V from its mean are approximated by a multidimensional white noise with covariance $D+D^T$. Thirdly, a correction to the ergode is derived when Z is moving, which in general yields damping. Fourthly, the effect of autonomous rather than externally imposed Z-motion is argued to make no difference, to the order of approximation considered. Fifthly, the ergode is approximated by the monode (canonical ensemble) for large number of degrees of freedom. Lastly, the Klimontovich interpretation is shown to be necessary.

4.1 Zeroth order mean velocity

Anosov and Kasuga (e.g. [LM]) proved that $W_Z(E)$ is an adiabatic invariant for X_{H_Z} ergodic on energy levels, with respect to slow external change of Z, i.e. for most trajectories the energy $E(t)$ changes in such a way to keep $W_{Z(t)}(E(t)) \approx w_0$. Define the "microcanonical free energy" $f : N \to \mathbb{R}$ for given w_0 by

$$f(Z) = W_Z^{-1}(w_0). \tag{4}$$

The following calculation shows that

$$\lambda(V) = J\nabla f. \tag{5}$$

Firstly, $W_Z(f(Z)) = w_0$, so $\nabla W + W'\nabla f = 0$, i.e.

$$\nabla f = -\frac{1}{W'}\nabla W. \tag{6}$$

Thus $(J\nabla f)_j = \{Z_j, f\} = -\frac{1}{W'}\{Z_j, W\}$. Next, the flow χ_u of $X_{Z_j\circ\pi}$ preserves Ω (because it is Hamiltonian) and the fibration π (because π is Poisson). Thus the change of $W_Z(E) = \int_{\{H\leq E\}} \Omega$ from $Z(0)$ to $Z(u)$ along the flow χ_u is the Ω-measure of the band in $\pi^{-1}(Z(u))$ between $H^{-1}(E)$ and $\chi_u((H,Z)^{-1}(E, Z(0))$. The rate of change of H along the flow χ_u is $\{H, Z_j\circ\pi\}$ and we can write $\Omega = \mu\wedge dH$ in a fibre, so

$$\{W, Z_j\} = -\int_{H^{-1}(E)} \{H, Z_j\circ\pi\}\mu.$$

Finally, $\lambda(V_j) = \frac{1}{W'}\int_{H^{-1}(E)}\{Z_j\circ\pi, H\}\mu$.

Remark: A similar calculation with respect to the canonical ensemble $e^{\beta F}e^{-\beta H}\Omega$ on $\pi^{-1}(Z)$ (Boltzmann's "monode") gives mean velocity $J\nabla F$, where the canonical free energy F is defined by

$$e^{-\beta F(Z)} = \int_{\pi^{-1}(Z)} e^{-\beta H}\Omega_Z. \tag{7}$$

Nevertheless, the canonical ensemble is not ergodic (energy is conserved), so it is not clear how to proceed further (unless we just ignore the mixing requirement for the next results).

Anosov's averaging theorem (e.g. [LM]) proves that most trajectories follow the mean dynamics closely on short enough timescales. Thus from (5) the zeroth order dynamics of the slow degrees of freedom is Hamiltonian with Hamiltonian function the microcanonical free energy f. Note in particular that the zeroth order dynamics preserves f, which is a restatement of the adiabatic invariance of $W_Z(E)$.

I am interested in capturing effects that go beyond zeroth order, however, in particular on longer timescales where the invariance of $W_Z(E)$ breaks down.

4.2 Fluctuations

The fluctuations $v(t)$ of $V\circ\psi_t$ from its mean for fixed Z produce an effect approximately equivalent to white noise with covariance matrix $R = \int_{-\infty}^{\infty} C(t)\, dt$, where $C_{ij}(t) = \lambda(v_i(t)v_j(0))$, provided the integral converges. This is well known but derivations vary in level of sophistication.

The simplest version is to let $z(t) = \int_0^t v(s)\, ds$ (I denote it by z rather than Z because this expression does not include the mean velocity of Z nor the fact that the distribution of v changes as Z moves) and prove that

$$\lambda(z_i(t)z_j(t))/t \to R_{ij}$$

as $t\to +\infty$ (Green-Kubo formula).

Here is a proof. From the definitions of z and C, $\lambda(z_i(t)z_j(t)) = \int_{-t}^t (t-|u|)C_{ij}(u)\, du$. So

$$\lambda(z_i(t)z_j(t))/t = \int_{-t}^t (1-\frac{|u|}{t})C_{ij}(u)\, du.$$

Tackle the positive and negative ranges of u separately. Convergence of the integral for R implies that given $\varepsilon > 0$ there is a t_0 such that $|\int_u^t C(v)\, dv| \leq \varepsilon$ for all $t\geq u\geq t_0$. Then for $t\geq t_0$,

$$\int_0^\infty C(u)\, du - \int_0^t (1-\frac{u}{t})C(u)\, du = \int_t^\infty C(u)\, du + \frac{1}{t}\int_0^{t_0} uC(u)\, du + \frac{1}{t}\int_0^t\int_{\max(u,t_0)}^t C(v)\, dv\, du.$$

The first and third terms are each at most ε in absolute value. The second is at most ε as soon as $t\geq \frac{1}{\varepsilon}\int_0^{t_0} uC(u)\, du$. Hence $\int_0^t (1-\frac{u}{t})C(u)\, du \to \int_0^\infty C(u)\, du$ as $t\to\infty$. Similarly for $u<0$ and hence the result. Note that in contrast to a statement in [Ga] it is not necessary to assume $\frac{1}{T}\int_{-T}^{T}|t|C(t)\,dt \to 0$ as $T \to \infty$: it follows automatically from $\int_{-\infty}^{\infty} C(t)\,dt < \infty$.

As a corollary this shows the covariance matrix to be non-negative. The covariance matrix R can be written as $D + D^T$, with D as defined in (2). Thus it also follows that the symmetric part S of D is non-negative.

More sophisticated results use the theory of martingales: stochastic processes on vector spaces such that the expectation of the future value given the present one is just the present value. The best is to prove a "vector-valued almost sure invariance principle", in the sense that the paths $z(T)$ in $\mathbb{R}^{2n}$ are distributed within $O(T^\delta)$ of the distribution for the corresponding Brownian paths, for some $\delta < 1/2$. This follows from a general result of [MN] provided the fast dynamics is sufficiently rapidly mixing (exponential suffices).

To increase the accuracy of the approximation by white noise, it could be a good idea to iteratively improve the choice of fibration π to make C decay as fast as possible, in particular to remove major changes in sign.

4.3 Correction to ergode

If $Z(t)$ is moved slowly along some path then the natural probability measure on the moving fast system lags slightly behind the instantaneous ergode $\lambda_{Z,E}$ for the given value of w_0. If we assume that $W_Z(E)$ is conserved exactly, the first order difference can be computed by a formula of [Ru] (extrapolating a bit beyond his hypotheses, but see [EHL] for a statement of what should suffice).

Ruelle's formula assumes a direction of time, in the sense that the probability distribution for the fast system is assumed to be absolutely continuous along unstable manifolds, whereas one could have asked for it along stable manifolds. Many people regard this as justified by a hypothesis of a low entropy initial condition for the universe [Le], but hypotheses on initial conditions seem to me inadequate to explain the direction of causality. I suspect a true explanation lies in quantum gravity: probably a consistent theory of quantum gravity will exhibit two phases, differing in the direction of interaction of radiation and matter. Our patch of space-time is in one of these phases and we choose to orient time accordingly.

For a slowly varying vector field X_t on a manifold M, which at each time t is mixing Anosov, Ruelle's formula specifies the first order change to the expectation $\langle O(t)\rangle$ of any smooth observable $O : M \to \mathbb{R}$ at time t from that for the frozen system X_t:

$$\delta\langle O(t)\rangle = \int_{-\infty}^{t} \langle d(O \circ \psi_{ts})(X_s - X_t)\rangle \, ds, \tag{8}$$

where ψ_{ts} is the flow of X_t from time s to time t. The term X_t can be dropped since $\langle d(O \circ \psi_{ts})X_t\rangle$ is the expectation of the rate of change of a function with respect to an invariant measure, so zero (it was included to make clear that the result is first order in the change in the vector field). In our case, the state space $\pi^{-1}(Z)$ also changes with time so evaluating X_s requires choosing some diffeomorphism between the state spaces at times s and t, but the result does not depend on the choice.

Let us calculate the change in the mean of V due to slow motion of Z. For X_t we use $X_{H_{Z(t)}}$ and for the ensemble average we use $\lambda_{Z(t),E}$. Any motion of Z can be specified as the result of a (possibly time-dependent) Hamiltonian flow on N, with some Hamiltonian G, so $\dot{Z} = J\nabla G$. The function G lifts to $G \circ \pi$ on M and so induces a fibre-preserving flow χ on M, which we can use to identify points of different fibres. In particular for X_s in Ruelle's formula we can use $\chi_{ts}^* X_{H_{Z(s)}}$, which can be written as $X_{(H\circ\chi_{st})_{Z(t)}}$. Then

$$d(V_j \circ \psi_{ts})X_s = \{V_j \circ \psi_{ts}, H \circ \chi_{st}\}_{Z(t)}, \tag{9}$$

where $\{,\}_Z$ is the Poisson bracket on $\pi^{-1}(Z)$, defined via the restriction of the symplectic form to the fibre. Thus Ruelle's formula gives a time-integral of an energy-level average of a Poisson bracket on a fibre.

Lemma: For symplectic manifold K with volume form Ω, Hamiltonian H, energy level E, normalised energy level volume λ_E and any smooth functions $F, G : K \to \mathbb{R}$ for which the required integrals converge,

$$\int \{F, G\}\ d\lambda_E = \frac{1}{W'(E)} \frac{\partial}{\partial E} \left(W'(E) \int \{F, H\} G\ d\lambda_E \right). \tag{10}$$

Proof: For any smooth functions $F, U : K \to \mathbb{R}$ for which the integral converges,

$$\int \{F, U\}\ d\Omega = \int dF(X_U)\ d\Omega = 0,$$

because it is the integral of the rate of change of F along orbits of X_U with respect to an invariant measure. Apply this to a product $U = GA$ and use Leibniz' rule and antisymmetry for Poisson brackets to deduce that

$$\int A\{F, G\}\ d\Omega = \int \{A, F\} G\ d\Omega. \tag{11}$$

Now take A to be (a sequence of smooth approximations to) $\delta(E - H)$:

$$\int \delta(E - H)\{F, G\}\ d\Omega = \int \{\delta(E - H), F\} G\ d\Omega = -\int \delta'(E - H)\{H, F\} G\ d\Omega, \tag{12}$$

since $\{., F\}$ is a derivation. The right hand side can be written as

$$\frac{\partial}{\partial E} \int \delta(E - H)\{F, H\} G\ d\Omega.$$

All that remains is to write $\delta(E - H)\ d\Omega = W'(E)\ d\lambda$ on both sides. □

Applying the lemma to (9) produces

$$\delta\langle V_j \rangle = \frac{1}{W'(E)} \frac{\partial}{\partial E} \left(W'(E) \int_{-\infty}^{t} ds \langle \{H, H \circ \chi_{st}\}_Z V_j \circ \psi_{ts} \rangle \right). \tag{13}$$

Now

$$\frac{\partial}{\partial s} H \circ \chi_{st} = -\{H, G \circ \pi\}_M,$$

so for times s out to some decorrelation time, which we supposed to be $\varepsilon \ll 1$ (Assumption 6), we can write to leading order

$$\{H, H \circ \chi_{st}\} = (t - s)\{H, \{H, G \circ \pi\}_M\}_Z.$$

Specialising to $G = Z_k$ gives $\{H, G \circ \pi\}_M = -V_k$. So the integral in (13) becomes

$$\int_{-\infty}^{t} ds\ (t - s)\langle \{H, V_k\}_Z(s) V_j(t) \rangle.$$

To justify this approximation properly requires some hypothesis on the rate of decay of the correlation function of v (probably $\int_{-\infty}^{0} |tC(t)|\ dt < \infty$ suffices). Now $\{H, V_k\}_Z = -\frac{dV_k}{ds}$ along the flow of X_{H_Z}, so integration by parts (with again some assumption about sufficiently rapid convergence of the autocorrelation integral) transforms the integral to

$$-\int_{-\infty}^{t} ds \langle (V_k(s) - \langle V_k \rangle) V_j(t) \rangle = -D_{jk},$$

with D given by (2). Taking G to be an arbitrary linear combination of Z_k yields

$$\delta\langle V(t) \rangle = -\frac{(W'D)'}{W'} J^{-1} \dot{Z}. \tag{14}$$

This result agrees with [BR], except that the term "f_1" which they chose to neglect does not arise here. Perhaps I lost it through making a constant $\dot{Z}$ approximation, or perhaps it really gives nothing to leading order.

The derivation requires $\dot{Z}$ to be small, since Ruelle's formula is first order. This is achieved for small enough ε, because the relevant notion of smallness of $\dot{Z}$ is on the fast timescale.

For $k = m-n$ large, one can expect W' to vary much more rapidly than D with energy E. In particular, under the assumption of Section 4.5 to come, $W'(E) \sim e^{ks(E/k)}$. Then (14) can be approximated by $-\beta DJ^{-1}\dot{Z}$ with

$$\beta = (\log W')',$$

the inverse temperature. Thus we obtain

$$\dot{Z} = J\nabla f - \beta DJ^{-1}\dot{Z} \tag{15}$$

for the mean motion of Z.

Now invoke Assumption 8 of Section 2, which I'll express explicitly as $\beta\|DJ^{-1}\| \ll 1$. Then (15) can be rewritten as

$$\dot{Z} = (I + \beta DJ^{-1})^{-1}J\nabla f \approx (J - \beta D)\nabla f. \tag{16}$$

So the result of the correction is to modify the matrix representing the Poisson bracket. D can have an antisymmetric component, which just changes J to a different antisymmetric matrix (an effect called "geometric magnetism" by [BR], though it is not clear to me whether the result automatically satisfies the Jacobi identity). Assuming $\beta > 0$, the symmetric part of D, being non-negative, produces damping because under (16), $df(Z)/dt = -\beta\nabla_i f D_{ij}\nabla_j f \leq 0$ to first order.

Ruelle's formula could also be used to determine the first order change to the covariance of the fluctuations from the mean velocity, but since we are considering the fluctuations to already be first order small it does not make sense to determine this second order effect on its own.

In principle, the effects of deviations from conservation of $W_Z(E)$ should also be analysed; indeed, in the view of [O] they are responsible for the dissipation.

4.4 Effect of autonomous slow motion

In reality $Z(t)$ does not move along a predetermined path but is driven by $\dot{Z} = V$, which depends on the choice of initial condition from λ_{Z_0,E_0}. The difficulties induced by back-reaction of the slow motion on the fast variables have been addressed by Kifer in some contexts [Ki]. It seems reasonable to add the fluctuations to the mean determined as above, but one ought to verify that correlations do not cause a further change of the same order. This produces nearly the claimed result (1), but with ∇f instead of ∇F and the interpretation of " = " unspecified.

4.5 Ergode to monode

For large number of fast degrees of freedom, the microcanonical free energy can be approximated by the canonical free energy (up to a possible constant), assuming a thermodynamic limit exists for the heat capacity. Here is a derivation.

From (7), the canonical free energy for given β_0 satisfies $e^{-\beta_0 F(Z)} = \int e^{-\beta_0 E} W'_Z(E)\, dE$, so

$$-\beta_0 e^{-\beta_0 F}\nabla F = \int e^{-\beta_0 E}\nabla W'\, dE.$$

Integrating this by parts yields $\int \beta_0 e^{-\beta_0 E}\nabla W\, dE$, which by (6) can be written as $-\beta_0 \int e^{-\beta_0 E} W'\nabla f\, dE$, where the E-dependence of f is via $w_0 = W_Z(E)$. Thus

$$\nabla F = \int e^{-\beta_0 E} W'\nabla f\, dE \;/ \int e^{-\beta_0 E} W'\, dE.$$

For large $k = m - n$, assume the heat capacity per degree of freedom $c(\epsilon) = \frac{1}{kT'(k\epsilon)}$ as a function of the energy ϵ per degree of freedom is positive and bounded uniformly in k (say for simplicity that the limit as $k \to \infty$ exists). It follows by integration that $\beta(E) = 1/T(E)$ is a function of ϵ nearly independent of k, and by another integration the same for $\frac{1}{k}\log W'(E)$; write the latter as $s(\epsilon)$, the entropy per degree of freedom. Then the function $e^{-\beta_0 E}W'(E)$ of E is approximately $e^{-k(\beta_0\epsilon - s(\epsilon))}$, which is sharply peaked around the ϵ_0 such that $s'(\epsilon_0) = \beta_0$ (because $s''(\epsilon) = -\frac{\beta^2}{c(\epsilon)} < 0$), i.e. $\beta(E_0) = \beta_0$. Thus ∇F for β_0 is approximated by ∇f for this E_0.

4.6 Klimontovich interpretation

The Klimontovich interpretation [Kl] of the stochastic differential equation is as the limit as τ decreases to 0 of the discrete-time process

$$Z_i((n+1)\tau) - Z_i(n\tau) = \tau(J_{ij} - \beta D_{ij})\nabla_j F + \sigma_{ij}(Z((n+1)\tau))w_j(n) \tag{17}$$

with $w_j(n)$ independent random steps of zero mean and variance τ. The distinguishing feature is that σ is evaluated at the end of a step rather than at the beginning (as for Ito) or averaged between the ends (as for Stratonovich). More formally, the Klimontovich interpretation is defined as the Ito equation with added drift $\nabla_j S_{ji}$ where S is the symmetric part of D.

The differences between interpretations of the stochastic differential equation are probably beyond first order, but there is a clear preference for Klimontovich interpretation, because it is the only one for which the measure $e^{-\beta F}\Omega_N$ (the marginal on N of the canonical ensemble) is stationary if S depends on Z. This can be verified by substitution in (3). It was also understood by Hänggi [DH].

5 Case of standard mechanical system

If N is a cotangent bundle T^*L, the Hamiltonian has the form $H(Q,P,z) = \frac{1}{2}P^T M^{-1}P + h(Q,z)$ for slow degrees of freedom $(Q,P) \in T^*L$ and positive definite mass matrix $M(Q)$, and the symplectic form is $\omega = \sum_j dQ_j \wedge dP_j + \tilde{\omega}$ with $\tilde{\omega}$ a symplectic form in z (possibly depending on Q and P), then the formulae simplify. In particular $\dot{Q} = M^{-1}P$ exactly, F takes the form $F(Q,P) = \frac{1}{2}P^T M^{-1}P + G(Q)$ with G defined by $e^{-\beta G(Q)} = \int e^{-\beta h(Q,z)}\Omega$, D is independent of P, and $D_{ij} = 0$ if any index refers to a component of Q. So for the rest of this section, D denotes just its PP-block. G is called the "potential of mean force". Hence

$$d(M(Q)\dot{Q}) = -(\nabla G + \beta D\dot{Q})dt + \sigma dW. \tag{18}$$

Note that this can be derived even without making the approximation used in (16).

This simpler form of Langevin equation permits a further reduction in the "overdamped" case. Specifically, if D is invertible and the maximal rate of change of M, D and G along solutions is slow compared to the slowest damping rate $\beta\|MD^{-1}\|^{-1}$ then P relaxes onto a slow manifold and is thereafter slaved to the motion of Q, resulting to a good approximation in

$$\beta D_{ij} dQ_j = -\nabla_i G(Q)\ dt + \sigma_{ij} dW_j(t), \tag{19}$$

or equivalently (with $T = 1/\beta$ and taking D symmetric for simplicity)

$$dQ = -TD^{-1}\nabla G\ dt + 2T\sigma^{-T}\ dW \tag{20}$$

(using Klimontovich interpretation again). For a proof see Theorem 10.1 of [Ne]. This is the form commonly used in biochemistry, e.g. [BEL, MB] and perhaps justifies the idea of reactions following steepest descent curves on the free energy surface. Note that it produces Smoluchowski-Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = \mathrm{div}\left(T(D^{-1}\nabla G)\rho + T^2 D^{-1}\nabla\rho\right), \tag{21}$$

with stationary measure $e^{-\beta G}dQ$ (dQ being volume on L).

6 Quantum degrees of freedom

Quantum mechanics can be viewed as Hamiltonian. For Hermitian operator h on complex Hilbert space U, take the state space to be the projectivisation $P(U)$ of U (i.e. the set of 1D complex subspaces), endowed with the Fubini-Study symplectic form [MS], and take $H(\psi) = \langle\psi|h\psi\rangle/\langle\psi|\psi\rangle$. The resulting Hamiltonian vector field gives Schrödinger evolution $i\frac{d\psi}{dt} = h\psi$.

Alternatively, take M to be the dual of the Lie algebra of Hermitian operators on U with inner product $\langle A, B\rangle = \text{Tr } AB$ and its Lie-Poisson bracket, and $H(A) = \text{Tr } hA$. This gives von Neumann evolution $idA/dt = [h, A]$.

Thus one can incorporate quantum degrees of freedom in the above framework, for example electronic degrees of freedom involved in the conformation change of rhodopsin after absorbing a photon.

There is the problem however that the quantum dynamics is not Anosov, so after all perhaps the approach of [FKM, Z] is more appropriate.

7 Kinetics out of chemical equilibrium

The slow state space N can be a covering space. For example, its base can represent the conformation of myosin and associated momenta, and the decks differ by the number of ATP molecules. This is the appropriate way to view a biomolecular system in thermal equilibrium but out of chemical equilibrium. Indeed such problems motivated the present paper [MM].

8 Conclusion and Problems

A strategy has been sketched to derive a Langevin equation for slow degrees of freedom of a Hamiltonian system, under suitable assumptions. It would be good to carry out this programme in full.

I conclude with a few related problems for the future:

- Obtain higher accuracy by slightly different well selected choice of $Z(0)$.
- Determine the rank of S (degeneracies can arise only from coboundaries [CLB]), and conditions for hypoellipticity.
- Attempt to extend the results beyond mixing Anosov fast dynamics.
- Try to use the theory of partial hyperbolicity, because it can lead to ergodicity under fairly general circumstances.
- Adapt the approach for a constant pressure ensemble, e.g. by adding a heavy piston under gravity.
- Investigate what to do if there is no gap in timescales?
- Study ways in which the result fails. This would be more interesting than the whole programme.

Acknowledgements

This paper has benefitted from the responses of participants at the "Mathematics of Model Reduction" conference in Leicester, 28-30 August 2007, and a workshop on "Microscopic origins of dissipation and noise" at the Max Planck Institute for Mathematical Sciences, Leipzig, 31 Oct - 3 Nov 2007.

References

[BEL] Bemporad D, Essex JW, Luttmann C, Permeation of small molecules through a lipid bilayer: a computer simulation study, J Phys Chem B 108 (2004) 4875–84.

[BR] Berry MV, Robbins JM, Chaotic classical and half-classical adiabatic reactions: geometric magnetism and deterministic friction, Proc R Soc Lond A 442 (1993) 659–672.

[CLB] Conze J-P, Le Borgne S, Méthode de martingales et flot géodesique sur une surface de courbure constante négative, Ergod Th Dyn Sys 21 (2001) 421–441.

[DH] Dunkel J, Hänggi P, Theory of relativistic Brownian motion: the $(1+1)$-dimensional case, Phys Rev E 71 (2005) 016124.

[EHL] Eyink GL, Haine TWN, Lea DJ, Ruelle's linear response formula, ensemble adjoint schemes and Lévy flights, Nonlin 17 (2004) 1867–89.

[FKM] Ford GW, Kac M, Mazur P, Statistical mechanics of assemblies of coupled oscillators, J Math Phys 6 (1965) 504–515.

[GC] Gallavotti G, Cohen EGD, Dynamical ensembles in stationary states, J Stat Phys 80 (1995) 931–970.

[Ga] Gaspard P, Chaos, scattering and statistical mechanics (Cambridge Univ Press, 1998).

[GKS] Givon D, Kupferman R, Stuart A, Extracting macroscopic dynamics: model problems and algorithms, Nonlinearity 17 (2004) R55-R127.

[HM] Hunt TJ, MacKay RS, Anosov parameter values for the triple linkage and a physical system with a uniformly chaotic attractor, Nonlin 16 (2003) 1499–1510.

[J] Jarzynski C, Thermalization of a Brownian particle via coupling to low-dimensional chaos, Phys Rev Lett 74 (1995) 2937–40.

[Ki] Kifer Yu, Some recent advances in averaging, in: Modern Dynamical Systems and Applications, eds Brin M, Hasselblatt B, Pesin Ya (Cambridge Univ Press, 2004) 385–403.

[Kl] Klimontovich Yu L, Nonlinear Brownian motion, Physics Uspekhi 37 (1994) 737–766.

[La] Langevin P, Sur la théorie du mouvement brownien, Comptes Rendus Acad Sci Paris 146 (1908) 530–3.

[Le] Lebowitz JL, Statistical mechanics: a selective review of two central issues, Rev Mod Phys 71 (1999) S346–S357.

[Li] Lieb E, Some problems in statistical mechanics that I would like to see solved, Physica A 263 (1999) 491–9.

[LM] Lochak P, Meunier C, Multiphase averaging for classical systems (Springer, 1988).

[MM] MacKay RS, MacKay DJC, Ergodic pumping: a mechanism to drive biomolecular conformation changes, Physica D 216 (2006) 220–234.

[MB] Marrink S-J, Berendsen HJC, Simulation of water transport through a lipid membrane, J Phys Chem 98 (1994) 4155–68.

[MS] McDuff D, Salamon D, Introduction to Symplectic Topology (Oxford, 1995).

[MN] Melbourne I, Nicol M, A vector-valued almost sure invariance principle for hyperbolic dynamical systems, arXiv/math.DS/0606535

[Ne] Nelson E, Dynamical theories of Brownian motion (Princeton Univ Press, 1967).

[O] Ott E, Goodness of ergodic adiabatic invariants, Phys Rev Lett 42 (1979) 1628–31.

[Ru] Ruelle D, Differentiation of SRB states for hyperbolic flows, IHES/M/04/33

[So] Soffer, Dissipation through dispersion, CRM Proc 27 (2001) 175–184.

[Wi] Wilkinson M, Dissipation by identical oscillators, J Phys A 23 (1990) 3603–11.

[Z] Zwanzig R, Nonlinear generalized Langevin equations, J Stat Phys 9 (1973) 215–220.