Smoothed Dynamics of Highly Oscillatory Hamiltonian Systems*

Sebastian Reich
Konrad-Zuse Zentrum Berlin
Heilbronner Str. 10
D-10711 Berlin

June 12, 1995

Abstract

We consider the numerical treatment of Hamiltonian systems that contain a potential which grows large when the system deviates from the equilibrium value of the potential. Such systems arise, e.g., in molecular dynamics simulations and the spatial discretization of Hamiltonian partial differential equations. Since the presence of highly oscillatory terms in the solutions forces any explicit integrator to use very small step-size, the numerical integration of such systems provides a challenging task. It has been suggested before to replace the strong potential by a holonomic constraint that forces the solutions to stay at the equilibrium value of the potential. This approach has, e.g., been successfully applied to the bond stretching in molecular dynamics simulations. In other cases, such as the bond-angle bending, this methods fails due to the introduced rigidity. Here we give a careful analysis of the analytical problem by means of a smoothing operator. This will lead us to the notion of the smoothed dynamics of a highly oscillatory Hamiltonian system. Based on our analysis, we suggest a new constrained formulation that maintains the flexibility of the system while at the same time suppressing the high-frequency components in the solutions and thus allowing for larger time steps. The new constrained formulation is Hamiltonian and can be discretized by the well-known SHAKE method.

*This work was supported in part by DOE/NSF Grant DE-FG02-91-ER25099/DMS-9304268, by NIH Grant P41R05969, and by NSF/ARPA Grant ASC-9318159

1 Introduction

We are concerned in this paper with the numerical solution of singularly perturbed Hamiltonian systems of the form

$$\begin{aligned} \frac{d}{dt}q &= M^{-1}p \\ \frac{d}{dt}p &= -\nabla V(q) - \frac{1}{\epsilon^2}G(q)^T K g(q) \end{aligned} \tag{1}$$

with Hamiltonian function

$$H(q,p) = \frac{p^T M^{-1} p}{2} + V(q) + \frac{1}{\epsilon^2}\frac{g(q)^T K g(q)}{2} \tag{2}$$

where ϵ a small parameter, $q, p \in R^n$, and $G(q) = g_q(q)$. Here M is the positive definite mass matrix of the system, $V : R^n \to R$ the potential energy function, and g is the collection of functions $g_i : R^n \to R$, $i = 1, \ldots, m$, with corresponding (scaled) force constant $K_{i,i}$, i.e.

$$\frac{g(q)^T K g(q)}{2} = \frac{1}{2}\sum_i K_{i,i}\, g_i(q)^2$$

and K the m-dimensional diagonal matrix with entries $K_{i,i}$.

We assume throughout the paper that the matrix

$$G(q)M^{-1}G(q)^T \tag{3}$$

is invertible and both $||GM^{-1}G^T||$ as well as $||[GM^{-1}G^T]^{-1}||$ take values close to one along trajectories of (1).

Note that the parameter ϵ has no immediate physical meaning and is not uniquely determined by the physical problem. It stands for the fact that the potential $g(q)^T K g(q)/(2\epsilon^2)$ grows large away from its equilibrium value $g(q) = 0$ compared to $V(q)$ and it allows one to treat the mathematical consequences of this fact in a relatively elegant way.

Throughout the paper we will use the following convention: Assume that the Hamiltonian (2) has been scaled such that

$$||V_{qq}(q)|| \leq 1 \tag{4}$$

and that the mass matrix M satisfies $M \approx I$ where I is the identity matrix. (In fact, $M = I$ can always be achieved by the canonical coordinate transformation $(q,p) \to (M^{1/2}q, M^{-1/2}p)$.) Then ϵ is chosen such that

$$||K^{-1}|| \leq 1 \tag{5}$$

where we assume that the thus defined ϵ satisfies $\epsilon \ll 1$.

Let us demonstrate this for the quadratic Hamiltonian

$$H(q,p) \;=\; \frac{p^T p}{2} + \frac{q^T W q}{2} + \frac{q^T G^T U G q}{2}$$

where W is a positive definite $n \times n$ matrix, G is a $n \times m$ matrix, $m < n$, and U is a diagonal $m \times m$ matrix. Here we would apply the canonical coordinate transformation $(q,p) \to (||W||^{-1/4} q, ||W||^{1/4} p)$ and then premultiply the resulting Hamiltonian by $||W||^{-1/2}$ which yields

$$H(q,p) \;=\; \frac{p^T p}{2} + \frac{q^T W q}{2\,||W||} + \frac{q^T G^T U G^T q}{2\,||W||}\,.$$

Then we define ϵ by

$$\epsilon \;:=\; \sqrt{||W||\,||U^{-1}||}\,.$$

Thus ϵ is proportional to the ratio of the largest eigenvalue of W to the smallest eigenvalue of U. Provided that $\epsilon \ll 1$, we would then define the function g in (2) by $g(q) := Gq$, the matrix K by $K := ||U^{-1}||\,U$, and the potential $V(q)$ by $(q^T W q)/(2\,||W||)$.

Throughout this paper we will also assume that, along solutions of (1), we have

$$\frac{p^T M^{-1} p}{2} \;\approx\; \frac{n}{2}\,\delta \tag{6}$$

with

$$\epsilon^2 \;\le\; \delta \;<\; 1 \tag{7}$$

for the chosen scaling of the Hamiltonian (2). Here n is the number of degrees of freedom of (1) and δ corresponds to $k_B T$ in statistical mechanics; T temperature and k_B the Boltzmann constant. (6) is typically satisfied for Hamiltonian systems of type (1) with n large; i.e. $n \gg 1$.

Hamiltonian systems of type (1) arise typically in the context of molecular dynamics simulations [14] (which provides the main motivation of this paper) and in the spatial discretization of Hamiltonian (hyperbolic) PDEs [13] like, for example, the Sine-Gordon equation by spectral or related methods. In the context of molecular dynamics, the potential $g(q)^T K g(q)/(2\epsilon^2)$ stands for covalent bond stretching and bond-angle bending; i.e., $g_i(q) = r - r_0$, $\epsilon^2 \approx 0.01$, and $\delta \approx 0.1$ in case of bond stretching and $g_i(q) = \phi - \phi_0$, $\epsilon^2 \approx 0.1$, and $\delta \approx 0.1$ in case of bond-angle bending. In the context of hyperbolic PDEs, the same expression is related to the high frequency modes in the Fourier spectrum of the solutions.

Differential equations of the form (1) fall into the class of singularly perturbed systems of type [6]

$$\frac{d}{dt}z \;=\; \frac{1}{\epsilon}f(z,\epsilon) \tag{8}$$

Solutions of (8) satisfy, in general,

$$||z(t)|| \;=\; O(1)$$

and

$$||\frac{d}{dt}z(t)|| \;=\; O(\epsilon^{-1})\,,$$

i.e., they are bounded but vary rapidly in t. Thus the step-size of a numerical integrator has, in general, to be of order $O(\epsilon)$. This implies a significant amount of computational work for the numerical integration over time intervals of order $O(1)$. For example, the lengths of a molecular dynamics simulation with an explicit method like Verlet [28] is for that reason restricted to a few tens of picoseconds up to a few nanoseconds, depending on the size of the problem [14]. This means that the time scale of the process that can be simulated is limited. To simulate processes over longer periods of time, new integration methods are essential.

Most of the theory has been developed for singularly perturbed problems that satisfy

1. $\operatorname{rank} f_z(z,0) = \text{const.}$

for all z. This implies that the set $\mathcal{M}_0$ defined by

$$\mathcal{M}_0 \;:=\; \{z : f(z,0) = 0\}$$

is a smooth manifold. The more stringent requirement is however that

2. $\mathcal{M}_0$ is an exponentially stable manifold of the differential equation

$$\frac{d}{dt}z \;=\; f(z,0)\,.$$

Under the Assumptions 2, one can show that there exists a family $\mathcal{M}_\epsilon$ of smooth manifolds with $\mathcal{M}_{\epsilon=0} = \mathcal{M}_0$ such that $\mathcal{M}_\epsilon$ is an exponentially stable invariant manifold of (8) [6]. Furthermore, the solutions on $\mathcal{M}_\epsilon$ reflect the long-time behavior of the general solutions of (8) with initial values in a σ-neighborhood of $\mathcal{M}_\epsilon$ up to terms of order $O(\sigma\,\epsilon)$. Since the solutions on $\mathcal{M}_\epsilon$ satisfy now $dz/dt = O(1)$, time-steps of order $O(1)$ can be used in a

numerical integrator provided that the equations are discretized by a proper (implicit) method [15].

However, Assumption 2 is not satisfied for singularly perturbed Hamiltonian systems. In particular, solutions of (1) oscillate highly about the manifold $\mathcal{M}_0$. Thus, as we will show in Section 2, the manifold $\mathcal{M}_0$ does not even satisfy the weaker assumption of normal-hyperbolicity [6],[8]. This leaves us with the task of finding a different approach to the long-time integration of (1). In this paper, we attempt to do so by introducing the notion of the *smoothed dynamics* of highly oscillatory Hamiltonian systems. By this we mean the following:

Because of (4), the shortest period in the motion of (1) due to the potential $V(q)$ is of order $O(1)$. In contrast to this, the potential $g(q)^T K g(q)/(2\epsilon^2)$ contributes high-frequency terms with period of order $O(\epsilon)$. To separate these high frequency components from the slowly varying parts, we introduce the smoothing operator

$$\langle w \rangle_\alpha (t) \quad := \quad \frac{1}{\alpha} \int_{-\infty}^{+\infty} \rho(\frac{t-t'}{\alpha})\, w(t')\, dt' \tag{9}$$

with $0 < \alpha \ll 1$ and $w : R \to R$. Here $\rho : R \to R$ is an appropriate weight function such that for any (bounded) continuous function w there is a smooth (C^∞) function $\bar{w}$ with

$$\langle w \rangle_\alpha (t) - \bar{w}(t) \quad = \quad O(\alpha^s) \tag{10}$$

and for any smooth (C^∞) function w we have

$$\langle w \rangle_\alpha (t) - w(t) \quad = \quad O(\alpha^s) \tag{11}$$

where s is a fixed integer with $s \gg 1$. One could, for example, chose for ρ the Meyer scaling function [5]. Note that, in the frequency domain, the smoothing operator (9) corresponds to a low pass filter with cut-off frequency $\omega_c = O(1/\alpha)$.

From now on we will always identify $\langle w \rangle_\alpha$ with a smooth function $\bar{w}$ such that (10) holds and assume that $\langle w \rangle_\alpha$ is C^∞.

The idea is now to replace the rapidly varying solutions $q(t)$ of (1) by $\langle q \rangle_\alpha (t)$ with

$$\alpha \quad = \quad \sqrt{\epsilon}$$

and then to seek an approximation to the smooth $\langle q \rangle_{\sqrt{\epsilon}}$ rather then to the rapidly varying $q(t)$. We call the functions $\langle q \rangle_{\sqrt{\epsilon}}(t)$, corresponding to solutions $q(t)$ of (1), the *smoothed dynamics* of (1).

The aim of this paper is to derive a (constrained) Hamiltonian system with Hamiltonian $H_\epsilon(Q,P)$ and constraint manifold $\mathcal{M}_\epsilon$ such that the corresponding (smooth) solutions $(Q(t),P(t))$ satisfy

$$\langle q \rangle_{\sqrt{\epsilon}}(t) - Q(t) \;=\; O(\delta\,\epsilon^2)$$

and

$$\langle p \rangle_{\sqrt{\epsilon}}(t) - P(t) \;=\; O(\delta\,\epsilon^2)$$

over bounded intervals of time.

The approximation of (1) by a constrained Hamiltonian systems has been considered before (see, for example, [21],[27]). In a naive approach, one would introduce the new variable

$$\lambda \;:=\; \frac{1}{\epsilon^2} K g(q)$$

and rewrite (1) as

$$\begin{aligned} \frac{d}{dt}q &= M^{-1}p \\ \frac{d}{dt}p &= -\nabla V(q) - G(q)^T\lambda \\ \epsilon^2 K^{-1}\lambda &= g(q) \end{aligned} \tag{12}$$

In the limit $\epsilon \to 0$, we obtain the constrained system

$$\begin{aligned} \frac{d}{dt}Q &= M^{-1}P \\ \frac{d}{dt}P &= -\nabla V(Q) - G(Q)^T\Lambda \\ 0 &= g(Q) \end{aligned} \tag{13}$$

which is Hamiltonian [13] on the constrained manifold

$$\mathcal{M}_0 \;=\; \{(Q,P) \in R^{2n} : g(Q) = 0,\, G(Q)M^{-1}P = 0\,\} \tag{14}$$

provided that the matrix (3) is invertible as assumed earlier.

This approximation has been used, for example, in MD simulations to remove the bond stretching modes [22], [26].

The solutions on $\mathcal{M}_0$ are now smooth. However, the approximation (8) introduces, in general, an error of order $O(\delta)$ over bounded time intervals (see Section 6). While, for example, this error turns out to be not significant

for the covalent bond stretching the same formulation (13) yields qualitatively wrong results when applied to the bond-angle bending or the harmonic dihedral bending [26].

The constrained formulation derived in this paper approximates the smoothed dynamics of (1) up to terms of order $O(\delta\,\epsilon^2)$ over bounded intervals of time and provides therefore a qualitative improvement over (13). Furthermore, let $\langle \mathcal{A}(q) \rangle$ denote the time average of an observable $\mathcal{A}(q)$ along trajectories $q(t)$ of (1). Then, assuming ergodicity of (1), we show that our approximation H_ϵ also satisfies

$$\langle \mathcal{A}(q) \rangle \;\; = \;\; \langle \mathcal{A}(Q) \rangle + O(\delta\,\epsilon^2)$$

where $\langle \mathcal{A}(Q) \rangle$ is the time average of the observable $\mathcal{A}$ along the corresponding smooth trajectories $Q(t)$ of H_ϵ. Finally, in Section 8, we discuss numerical aspects of our new method and demonstrate its properties by means of a simple numerical example.

Another approach to the long-time integration of highly oscillatory Hamiltonian system has been taken by Simo and his collaborators [25]. They advocate the direct discretization of (1) by an implicit energy-momentum method and the usage of a large step-size. However, there do not exist rigorous stability and order of convergence results for these methods when applied to general systems of type (1) with a step-size $\Delta t \gg \epsilon$.

2 Mathematical Background

In the first part of this section we show how to reformulate (1) as a singularly perturbed problem (8). To do so, we introduce local coordinates (q_1, q_2) by

$$\begin{aligned} q_1 &= g(q) \\ q_2 &= b(q) \end{aligned}$$

where $b(q)$ is a vector valued function such that $B(q)M^{-1}G(q)^T = 0$, $B(q) = b_q(q)$, and the composed matrix $[G(q)^T\, B(q)^T]$ is invertible and well conditioned. The existence of such a coordinate system follows, at least locally, from the Frobenius Theorem [2]. The corresponding conjugate momenta are given by

$$[G(q)^T\, B(q)^T] \left[\begin{array}{c} p_1 \\ p_2 \end{array} \right] \;\; = \;\; p$$

which results in the Hamiltonian

$$H(q,p) \;\; = \;\; \frac{p_1^T G M^{-1} G^T p_1}{2} + \frac{p_2^T B M^{-1} B^T p_2}{2} + V + \frac{1}{\epsilon^2}\frac{q_1^T K q_1}{2}\,. \quad (15)$$

The equations of motion are now given by

$$\begin{aligned} \frac{d}{dt}q_1 &= GM^{-1}G^T\,p_1 \\ \frac{d}{dt}p_1 &= -\nabla_{q_1}V - \frac{1}{\epsilon^2}Kq_1 - \nabla_{q_1}\frac{p_1^TGM^{-1}G^Tp_1 + p_2^TBM^{-1}B^Tp_2}{2} \end{aligned} \tag{16}$$

and

$$\begin{aligned} \frac{d}{dt}q_2 &= BM^{-1}B^T\,p_2 \\ \frac{d}{dt}p_2 &= -\nabla_{q_2}V - \nabla_{q_2}\frac{p_1^TGM^{-1}G^Tp_1 + p_2^TBM^{-1}B^Tp_2}{2} \end{aligned} \tag{17}$$

where, for notational convenience, we suppressed the arguments in the mappings $V(q_1,q_2)$, $G(q_1,q_2)$, and $B(q_1,q_2)$.

Upon rescaling p_1 in (16) by ϵ, the equations (16) become

$$\begin{aligned} \frac{d}{dt}q_1 &= \frac{1}{\epsilon}GM^{-1}G^T\,p_1 \\ \frac{d}{dt}p_1 &= -\epsilon\,\nabla_{q_1}V - \frac{1}{\epsilon}Kq_1 - \nabla_{q_1}\frac{p_1^TGM^{-1}G^Tp_1 + \epsilon^2\,p_2^TBM^{-1}B^Tp_2}{2\,\epsilon} \end{aligned} \tag{18}$$

which are now of the form (8). The corresponding manifold $\mathcal{M}_0$ is given by $q_1 = p_1 = 0$ or, in the original variables by (14). Linearization of (18) about the manifold $\mathcal{M}_0$ yields a linear system with eigenvalues on the imaginary axis. Thus the manifold $\mathcal{M}_0$ is not normally hyperbolic and the persistence of $\mathcal{M}_o$ for $\epsilon > 0$ cannot be concluded [8].

Now we want to derive a few important properties of the smoothing operator (9). We assume that $\rho : R \to R$ is a smooth function that goes to zero, as $|t| \to \infty$, faster than any inverse power of t, $\rho(0) = 1$, and

$$\int_{-\infty}^{\infty}\rho(t)dt = 1$$

The proper construction of a ρ, such that in addition (10) and (11) hold, falls into the subject of filter design and wavelet analysis [5].

The following four propositions will be used for the derivation of the smoothed dynamics of (1).

Proposition 1. *Let $w(t)$ be a differentiable function, then*

$$\langle\frac{d}{dt}w\rangle_\alpha = \frac{d}{dt}\langle w\rangle_\alpha$$

Proof. We have

$$\begin{aligned}\frac{d}{dt}\langle w\rangle_\alpha(t) &= \frac{d}{dt}\frac{1}{\alpha}\int_{-\infty}^{+\infty}\rho(\frac{t-t'}{\alpha})\,w(t')\,dt' \\ &= \frac{1}{\alpha}\int_{-\infty}^{+\infty}-\frac{d}{dt'}[\rho(\frac{t-t'}{\alpha})]\,w(t')\,dt'\end{aligned}$$

Now, since $\rho(\pm\infty)=0$, integration by parts yields

$$\begin{aligned}\frac{1}{\alpha}\int_{-\infty}^{+\infty}-\frac{d}{dt'}\rho(\frac{t-t'}{\alpha})\,w(t')\,dt' &= \frac{1}{\alpha}\int_{-\infty}^{+\infty}\rho(\frac{t-t'}{\alpha})\,\frac{d}{dt'}w(t')\,dt' \\ &= \langle\frac{d}{dt}w\rangle_\alpha(t)\end{aligned}$$

□

Proposition 2. *Let $w(t)\in R^n$ be an arbitrary function such that the matrix correlation function of $\Delta w := w-\langle w\rangle_\alpha$ satisfies*

$$\langle\Delta w\Delta w^T\rangle_\alpha = O(\alpha^l)$$

Let us also assume that the higher order momenta of $\Delta w(t)$ w.r.t. the smoothing operator (9) are small compared to $O(\alpha^l)$, then

$$\langle f(w)\rangle_\alpha = f(\langle w\rangle_\alpha)+O(\alpha^l)$$

where $f: R\to \boldsymbol{R}$ is a smooth function and l a positive real integer.

Proof. Taylor expansion of f about $\langle w\rangle_\alpha$ yields

$$f(w) = f(\langle w\rangle_\alpha)+f'(\langle w\rangle_\alpha)\Delta w+\frac{1}{2}\Delta w^T f''(\langle w\rangle_\alpha)\Delta w+\ldots$$

Applying our smoothing operator and noting that $\langle\Delta w\rangle_\alpha \approx 0$ and $\langle f(\langle w\rangle_\alpha)\rangle_\alpha \approx f(\langle w\rangle_\alpha)$ etc., we obtain

$$\langle f(w)\rangle_\alpha = f(\langle w\rangle_\alpha)+O(\langle\Delta w\Delta w^T\rangle_\alpha)$$

and the desired result follows. □

Proposition 3. *Let $w(t)\in R^n$ be an arbitrary function. As before, we write $w(t)$ as*

$$w(t) = \langle w\rangle_\alpha(t)+\Delta w(t)\,.$$

Then, up to terms of order $O(\alpha^s)$,

$$\langle ww^T\rangle_\alpha = \langle w\rangle_\alpha\langle w^T\rangle_\alpha+\langle\Delta w\Delta w^T\rangle_\alpha$$

Proof. Since, up to terms of order $O(\alpha^s)$, $\langle\langle w\rangle_\alpha \Delta w^T\rangle_\alpha = \langle w\rangle_\alpha \langle \Delta w^T\rangle_\alpha = 0$ and $\langle\langle w\rangle_\alpha \langle w\rangle_\alpha^T\rangle_\alpha = \langle w\rangle_\alpha \langle w\rangle_\alpha^T$. □

Proposition 4. *Let $w(t) \in R^n$ be an integrable function. We write $w(t)$ again as $w(t) = \langle w\rangle_\alpha + \Delta w(t)$. If $\Delta w(t) = O(\alpha^l)$ and $\Delta\bar{w}(\tau) := \Delta w(\alpha^2\tau)$ is smooth, then*

$$\int \Delta w(t)\, dt \;=\; O(\alpha^{l+2})\,.$$

Proof. Note that, by assumption, $\Delta\bar{w}(\tau)$ is smooth in τ and $\int \Delta\bar{w}(\tau)\, d\tau = O(\alpha^l)$. Thus

$$\begin{aligned}\int \Delta w(t)\, dt &= \alpha^2 \int \Delta w(\alpha^2\tau)\, d\tau \\ &= \alpha^2\, O(\alpha^l)\,.\end{aligned}$$

□

Let us finally review a few results from statistical mechanics. A Hamiltonian system with Hamiltonian H is called ergodic [16] if the time average

$$\langle \mathcal{A}\rangle \;:=\; \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \mathcal{A}(q(t),p(t))dt \tag{19}$$

of an observable $\mathcal{A}(q,p)$ is equal to the microcanonical (constant energy $E = H(q,p)$) ensemble average [16]

$$\langle \mathcal{A}\rangle_{\rm ens} \;:=\; \int\int \rho_{\rm ens}(q,p)\mathcal{A}(q,p)\, dqdp \tag{20}$$

with the microcanonical density function

$$\rho_{\rm ens}(q,p) \;:=\; \frac{\delta(H(q,p)-E)}{\int\int \delta(H(q,p)-E)\, dqdp}$$

where $\delta(x)$ denotes Dirac's delta function.

Remark. (i) If the Hamiltonian H possesses first integrals, ergodicity is always understood as ergodicity on the level sets of these first integrals.

(ii) Because of (6), the microcanonical ensemble average of (1) is almost identical to the macrocanonical (constant temperature) ensemble average with density function

$$\rho_{\rm ens}(q,p) \;:=\; \frac{\exp\left(-H(q,p)/\delta\right)}{\int\int \exp\left(-H(q,p)/\delta\right) dqdp}\,.$$

This is always true for systems with n very large; n the number of degrees of freedom [1]. Note that identifying microcanonical and macrocanonical ensemble averages is common practice in molecular dynamics [1].

Under the assumption that a given Hamiltonian system is ergodic, equipartition of energy [16] implies then that

$$\langle p^i \frac{\partial H}{\partial p^i} \rangle \quad = \quad \delta \tag{21}$$

and

$$\langle q^i \frac{\partial H}{\partial q^i} \rangle \quad = \quad \delta \tag{22}$$

where q^i and p^i, $i = 1, \ldots, n$, denote the ith component of the vector q, p respectively. Here H is the Hamiltonian of the system and, as introduced in (6), δ corresponds to $k_B T$ in statistical mechanics; T the temperature and k_B the Boltzmann constant.

Now, if the Hamiltonian happens to be a quadratic function of its coordinates; so, through a canonical transformation, it can be brought into the form

$$H \quad = \quad \sum_i \frac{A_i}{2} (\tilde{p}_i)^2 + \sum_i \frac{B_i}{2} (\tilde{q}_i)^2$$

and each solution component $(\tilde{q}_i(t), \tilde{p}_i(t))$ is Gaussian with $\langle \tilde{q}_i \rangle = \langle \tilde{p}_i \rangle = 0$, $\langle (\tilde{q}_i)^2 \rangle = \delta / B_i$, and $\langle (\tilde{p}_i)^2 \rangle = \delta / A_i$. Of course, linear systems cannot be ergodic. But, upon assuming that the linear system is the local approximation of a nonlinear system, the assumption of equipartitioning of energy among the harmonic degrees of freedom is justified.

3 The Quasi-Stationarity Assumption

In this and the following section we conduct an analysis of the Hamiltonian formulation (16) and (17) in terms of our smoothing operator (9). Let us denote the non-smooth solution component of $q(t)$ by $\Delta q(t)$; i.e.

$$\Delta q \quad := \quad q - \langle q \rangle_{\sqrt{\epsilon}} .$$

Note that $\Delta q(\epsilon \tau)$ is smooth in the fast time τ. We also write $Q(t)$ instead of $\langle q \rangle_{\sqrt{\epsilon}}$ or approximations thereof. (The same notations are used later on for the momentum $p(t)$.)

It follows from the equations (17) that $dq_2/dt = O(1)$ and $d^2 q_2 / dt^2 = O(1)$. Thus $d^2 \Delta q_2 / dt^2 = O(1)$ and Proposition 4 implies that $\Delta q_2 = O(\epsilon^2)$

and q_2 is smooth up to terms of order $O(\epsilon^2)$. This and Proposition 2 yield then that

$$\langle f(q_2)\rangle_{\sqrt{\epsilon}} \;=\; f(Q_2) + O(\epsilon^4)$$

for any smooth function f. Since we are only interested in an order $O(\delta\epsilon^2)$ approximation of the smoothed dynamics, we identify from now on q_2 with $Q_2 = \langle q_2\rangle_{\sqrt{\epsilon}}$.

The analysis is more complicated for the variable q_1. Here the above simple argument yields only $\Delta q_1 = O(\epsilon)$. Thus we have to analyse the fast subsystem (16) in more detail. To make such an analysis tractable, we make the following crucial assumptions:

Let us assume for a moment that we keep (q_2, p_2) in (16) constant. We call (16) *ergodic in the fast variable* (q_1, p_1) if for any observable $\mathcal{A}(q_1, p_1)$

$$\langle\mathcal{A}\rangle_{\sqrt{\epsilon}}(t) \;=\; \langle\mathcal{A}\rangle^{(1)}_{\rm ens}(q_2, p_2) + O(\delta\,\epsilon^2) \tag{23}$$

where $\langle\mathcal{A}\rangle^{(1)}_{\rm ens}$ denotes the microcanonical ensemble average over the variable (q_1, p_1); i.e.

$$\langle\mathcal{A}\rangle^{(1)}_{\rm ens}(q_2, p_2) \;:=\; \int\!\!\int \mathcal{A}(q_1, p_1)\rho^{(1)}_{\rm ens}(q_1, p_1, q_2, p_2)\, dq_1 dp_1$$

with the reduced density function

$$\rho^{(1)}_{\rm ens}(q_1, p_1, q_2, p_2) \;:=\; \frac{\delta(H(q_1, p_1, q_2, p_2) - E)}{\int\int \delta(H(q_1, p_1, q_2, p_2) - E)\, dq_1 dp_1}\,.$$

Note that the solutions $(q_1(t), p_1(t))$ vary on a time-scale of order $O(\epsilon)$ and $\langle\mathcal{A}\rangle_{\sqrt{\epsilon}}$ can be considered as the time-average over a time interval of order $O(\sqrt{\epsilon})$. Thus it seems reasonable to assume that $\langle\mathcal{A}\rangle_{\sqrt{\epsilon}} \approx \langle\mathcal{A}\rangle$ which, together with the standard ergodicity assumption, yields (23).

Now solutions $(q_2(t), p_2(t))$ are not constant but vary on a time-scale of order $O(1)$ and we assume that, along solutions $(q(t), p(t))$ of (1), we have

$$\langle\mathcal{A}\rangle_{\sqrt{\epsilon}}(t) \;=\; \langle\mathcal{A}\rangle^{(1)}_{\rm ens}(q_2(t), p_2(t)) + O(\delta\,\epsilon^2)$$

or, in other words, the solution component $(q_1(t), p_1(t))$ is always in thermodynamic equilibrium on time-scales of order $O(\sqrt{\epsilon})$. We call this the *quasi-stationarity* of (16).

Quasi-Stationarity Assumption. *From now on we assume that the fast system (16) is quasi-stationary along solutions $(q(t), p(t))$ of (1).*

4 Smoothed Dynamics – The Fast Subsystem

Let us now start with the analysis of (1) by applying the smoothing operator (9) to the fast system (16). In this section we will derive an order $O(\delta\,\epsilon^2)$ estimate of $Q_1 = \langle q_1 \rangle_{\sqrt{\epsilon}}$ which will provide us with the proper constraint function for our approximation of the smoothed dynamics of (1). In the following section we will then derive the corresponding Hamiltonian $H_\epsilon(Q,P)$.

Applying the smoothing operator (9) and Proposition 1 to (16), we obtain

$$\begin{aligned} \frac{d}{dt}Q_1 &= \langle GM^{-1}G^T p_1 \rangle_{\sqrt{\epsilon}} \\ \frac{d}{dt}P_1 &= -\langle \nabla_{q_1} V \rangle_{\sqrt{\epsilon}} - \frac{1}{\epsilon^2} K Q_1 - \\ &\quad - \langle \nabla_{q_1} \frac{p_1^T G M^{-1} G^T p_1 + p_2^T B M^{-1} B^T p_2}{2} \rangle_{\sqrt{\epsilon}} \end{aligned} \tag{24}$$

Since $(Q_1(t), P_1(t))$ is smooth, we have $dP_1/dt = O(1)$ and Q_1 has to satisfy

$$Q_1 = O(\epsilon^2)\,. \tag{25}$$

which is equivalent to

$$g(Q) = 0 \tag{26}$$

up to terms of order $O(\epsilon^2)$. Note that (26) is the constraint used in (13).

(25) and equipartitioning of energy in the variable q_1 (which is a consequence of our Quasi-Stationarity Assumption) allows us now to derive an estimate for $\langle \Delta q_1 \Delta q_1^T \rangle_{\sqrt{\epsilon}}$. In particular, (6) and equipartitioning of energy imply that

$$\frac{1}{\epsilon^2}\langle g(q)^T K g(q) \rangle_{\sqrt{\epsilon}} = \frac{1}{\epsilon^2}\langle q_1^T K q_1 \rangle_{\sqrt{\epsilon}} = O(m\,\delta)\,.$$

Thus

$$\langle q_1\, q_1^T \rangle_{\sqrt{\epsilon}} = O(\delta\,\epsilon^2)\,.$$

Knowing that $Q_1\, Q_1^T = O(\epsilon^4)$, we conclude from Proposition 3 that

$$\langle \Delta q_1\, \Delta q_1^T \rangle_{\sqrt{\epsilon}} = O(\delta\,\epsilon^2)\,.$$

This result and the similar result for the slow variable q_2, as derived in the previous section, allow us now to conclude from Proposition 2 that

$$\langle f(q_1, q_2) \rangle_{\sqrt{\epsilon}} = f(Q_1, Q_2) + O(\delta\,\epsilon^2)$$

for any smooth function f. Furthermore, it also follows that

$$\begin{aligned}\langle G(q)M^{-1}G(q)^T p_1\rangle_{\sqrt{\epsilon}} &= \langle G(Q)M^{-1}G(Q)^T(P_1+\Delta p_1)\rangle_{\sqrt{\epsilon}} + O(\delta\,\epsilon^2)\\ &= G(Q)M^{-1}G(Q)^T P_1 + O(\delta\,\epsilon^2)\,.\end{aligned}$$

Hence, (24) can now be rewritten as

$$\begin{aligned}\frac{d}{dt}Q_1 &= G(Q)M^{-1}G(Q)^T P_1\\ \frac{d}{dt}P_1 &= -\nabla_{Q_1}V(Q) - \frac{1}{\epsilon^2}KQ_1 -\\ &\qquad -\langle\nabla_{q_1}\frac{p_1^T GM^{-1}G^T p_1 + p_2^T BM^{-1}B^T p_2}{2}\rangle_{\sqrt{\epsilon}}\end{aligned}$$

Since $Q_1 = O(\epsilon^2)$, we also have $dQ_1/dt = O(\epsilon^2)$ which, with the above equations, implies that $P_1 = O(\epsilon^2)$ and $dP_1/dt = O(\epsilon^2)$. Thus

$$\begin{aligned}O(\epsilon^4) &= -K\,Q_1 - \epsilon^2\,\nabla_{Q_1}V(Q) -\\ &\qquad -\epsilon^2\langle\nabla_{q_1}\frac{p_1^T GM^{-1}G^T p_1 + p_2^T BM^{-1}B^T p_2}{2}\rangle_{\sqrt{\epsilon}}\,.\end{aligned}$$

Now

$$\begin{aligned}p_1^T GM^{-1}G^T p_1 + p_2^T BM^{-1}B^T p_2 &= p^T M^{-1}p\\ &= O(n\,\delta)\end{aligned}$$

and, therefore,

$$\langle\nabla_{q_1}\frac{p_1^T GM^{-1}G^T p_1 + p_2^T BM^{-1}B^T p_2}{2}\rangle_{\sqrt{\epsilon}} = O(\delta) \qquad (27)$$

which finally yields the estimate

$$Q_1 + \epsilon^2\,K^{-1}\,\nabla_{Q_1}V(Q) = O(\delta\,\epsilon^2)\,.$$

In the original Cartesian coordinate Q, this can be written as

$$0 = g(Q) + \epsilon^2\,K^{-1}\,[G(Q)M^{-1}G(Q)^T]^{-1}G(Q)M^{-1}\nabla V(Q) \qquad (28)$$

up to terms of order $O(\delta\,\epsilon^2)$.

This equation will provide the holonomic constraints for our approximation of the smoothed dynamics of (1). In contrast to the *hard* constraint (26) which introduces an error of order $O(\epsilon^2)$, we call (28) *soft* constraints.

In the following section we will derive the Hamiltonian that describes the smoothed dynamics on

$$\mathcal{M}_\epsilon := \{(Q,P)\in R^{2n} : \tilde{g}(Q) = 0,\, \tilde{G}(Q)M^{-1}P = 0\,\}$$

with $\tilde{G}(q) := \tilde{g}_Q(Q)$ and the soft constraint function

$$\tilde{g}(Q) := g(Q) + \epsilon^2\,K^{-1}\,[G(Q)M^{-1}G(Q)^T]^{-1}G(Q)M^{-1}\nabla V(Q)\,.$$

5 Smoothed Dynamics – The Slow Subsystem

In this section we derive the smoothed equations of the slow subsystem (17).

First we want to obtain an estimate for Δp_2. Because of (27) and $\langle \nabla_{q_2} V(q_1, q_2) \rangle_{\sqrt{\epsilon}} = \nabla_{Q_2} V(Q_1, Q_2) + O(\delta\, \epsilon^2)$, $d\Delta p_2/dt$ satisfies the estimate

$$\frac{d}{dt}\Delta p_2 \;=\; O(\delta)$$

which, according to Proposition 4, implies that

$$\Delta p_2 \;=\; O(\delta\, \epsilon)\,.$$

Thus

$$\langle f(p_2) \rangle_{\sqrt{\epsilon}} \;=\; f(P_2) + O(\delta^2\, \epsilon^2)$$

for any smooth function f and we can identify from now on p_2 with P_2 (as done before for the variable q_2). With this in mind, the smoothing of (17) results, up to terms of order $O(\delta\, \epsilon^2)$, in

$$\begin{aligned} \frac{d}{dt}Q_2 &= B(Q) M^{-1} B(Q)^T\, P_2 \\ \frac{d}{dt}P_2 &= -\nabla_{Q_2} V(Q) + \nabla_{q_2} \frac{P_2^T B(Q) M^{-1} B(Q)}{2} - \\ &\qquad - \langle \nabla_{q_2} \frac{p_1^T G(q) M^{-1} G(q)^T p_1}{2} \rangle_{\sqrt{\epsilon}} \end{aligned} \tag{29}$$

This leaves us with the task of finding the smoothed expression for

$$\langle \nabla_{q_2} \frac{p_1^T G(q) M^{-1} G(q)^T p_1}{2} \rangle_{\sqrt{\epsilon}}\,.$$

To do so we have to make use of the Quasi-Stationarity Assumption which implies equipartitioning of energy in the variable p_1. In our case here, equipartitioning of energy means that

$$\begin{aligned} \langle p_1^T G(q) M^{-1} G(q)^T p_1 \rangle_{\sqrt{\epsilon}} &= \langle p_1^T \frac{\partial H}{\partial p_1}(q,p) \rangle_{\sqrt{\epsilon}} \\ &= m\, \delta \end{aligned}$$

with m the dimension of p_1. Now let $W(q)$ be an orthogonal matrix such that $W(q)^T G(q) M^{-1} G(q)^T W(q)$ is a diagonal matrix $D(q)$ with entries $d_{i,i}(q)$ and define $\tilde{p}_1 := W(q)^T p_1$. Then

$$\begin{aligned} \langle \tilde{p}_1^T D(q) \tilde{p}_1 \rangle_{\sqrt{\epsilon}} &= \langle p_1^T G(q) M^{-1} G(q)^T p_1 \rangle_{\sqrt{\epsilon}} \\ &= m\, \delta \end{aligned}$$

and equipartitioning of energy between the degrees of freedom yields

$$\langle (\tilde{p}_1^i)^2 \, d_{i,i}(q) \rangle_{\sqrt{\epsilon}} \;=\; \delta$$

$i = 1, \ldots, m$. Now, up to terms of order $O(\delta\,\epsilon^2)$,

$$\langle (\tilde{p}_1^i)^2 \, d_{i,i}(q) \rangle_{\sqrt{\epsilon}} \;=\; \langle (\tilde{p}_1^i)^2 \rangle_{\sqrt{\epsilon}} \, d_{i,i}(Q)$$

and

$$\langle (\tilde{p}_1^i)^2 \rangle_{\sqrt{\epsilon}} \;=\; \frac{\delta}{d_{i,i}(Q)} + O(\delta\,\epsilon^2) \, .$$

With, up to terms of order $O(\delta\,\epsilon^2)$,

$$\begin{aligned} \langle \nabla_{q_2} \frac{p_1^T G(q) M^{-1} G(q)^T p_1}{2} \rangle_{\sqrt{\epsilon}} \;&=\; \langle \nabla_{Q_2} \frac{p_1^T G(Q) M^{-1} G(Q)^T p_1}{2} \rangle_{\sqrt{\epsilon}} \\ &=\; \sum_i \frac{\langle (\tilde{p}_1^i)^2 \rangle_{\sqrt{\epsilon}}}{2} \nabla_{Q_2} d_{i,i}(Q) \, , \end{aligned}$$

we finally obtain

$$\langle \nabla_{q_2} \frac{p_1^T G(q) M^{-1} G(q)^T p_1}{2} \rangle_{\sqrt{\epsilon}} \;=\; \frac{\delta}{2} \nabla_{Q_2} \ln \left[\det \, D(Q) \right] + O(\delta\,\epsilon^2)$$

which, in terms of the original matrix $G(Q) M^{-1} G(Q)^T$, leads to the potential

$$V_F(Q) = \frac{\delta}{2} \ln \, \det \left[G(Q) M^{-1} G(Q)^T \right] \tag{30}$$

where we used that

$$\det \, D(Q) = \det \left[W(Q) D(Q) W(Q)^T \right] .$$

Summarizing the results of this section, we find that smoothing of (17) yields the system

$$\begin{aligned} \frac{d}{dt} Q_2 \;&=\; B(Q) M^{-1} B(Q)^T \, P_2 \\ \frac{d}{dt} P_2 \;&=\; -\nabla_{Q_2} V(Q) + \nabla_{Q_2} V_F(Q) + \nabla_{Q_2} \frac{P_2^T B(Q) M^{-1} B(Q)^T P_2}{2} \end{aligned} \tag{31}$$

Remark. The potential (30) has been introduced before by Fixman [7] in the context of statistical mechanics. He showed that (30) has to be included into the constrained formulation (13) to make sure that, in the limit $\epsilon \to 0$, the unconstrained system (1) and the constrained system (13) possess the same reduced density function $\rho_{\mathrm{ens}}(q_2, p_2)$. Similar results can be found in [27] and [17].

6 Constraint Formulations

The coordinates (q_1, p_1, q_2, p_2) were only introduced for theoretical purposes. This leaves us with the task of reformulating (31) in terms of the Cartesian coordinates (q, p). In fact, this turns out to be straightforward and we obtain

Theorem 1. *The constrained Hamiltonian equations*

$$\begin{aligned} \frac{d}{dt}Q &= M^{-1}P \\ \frac{d}{dt}P &= -\nabla V(Q) - \nabla V_F(Q) - \frac{1}{\epsilon^2}\, G(Q)^T K g(Q) - \tilde{G}(Q)^T \Lambda \\ 0 &= \tilde{g}(Q) \end{aligned} \tag{32}$$

provide an order $O(\delta\,\epsilon^2)$ approximation to the smoothed dynamics of (1) over bounded periods of time. The corresponding Hamiltonian is

$$H_\epsilon = \frac{P^T M^{-1} P}{2} + V(Q) + V_F(Q) + \frac{g(Q)^T K g(Q)}{2\,\epsilon^2} + \tilde{g}(Q)^T \Lambda$$

with the (soft) constraint manifold

$$\mathcal{M}_\epsilon = \{(Q,P) \in R^{2n} : \tilde{g}(Q) = 0,\, \tilde{G}(Q) M^{-1} P = 0\}\,.$$

Proof. In terms of the variable (Q_1, P_1, Q_2, P_2), (32) is equivalent to

$$\begin{aligned} \frac{d}{dt}Q_2 &= BM^{-1}B^T P_2 \\ \frac{d}{dt}P_2 &= -\nabla_{Q_2} V - \nabla_{Q_2} V_F - [BM^{-1}B^T]^{-1} BM^{-1}\tilde{G}^T\Lambda - \\ &\qquad - \nabla_{Q_2} \frac{P_1^T G M^{-1} G^T P_1 + P_2^T B M^{-1} B^T P_2}{2} \end{aligned} \tag{33}$$

and

$$\begin{aligned} Q_1 &= -\epsilon^2\, K^{-1}\, \nabla_{Q_1} V(Q) \\ P_1 &= [G(Q) M^{-1} G(Q)^T]^{-1} \frac{d}{dt} Q_1 \end{aligned}$$

where we used in (33) that

$$\nabla_{Q_2} f = [BM^{-1}B^T]^{-1} BM^{-1} \nabla_Q f$$

for any differentiable function $f : R^n \to R$.

Furthermore, the Lagrange multiplier Λ is obtained by differentiating the constraint $\tilde{G}(Q)M^{-1}P = 0$ with respect to time and replacing the time derivatives dQ/dt and dP/dt by the corresponding expressions in (32). This yields, in local coordinates,

$$\begin{aligned} O(\epsilon^2) &= \Lambda + \nabla_{Q_1}(V + V_F) + \epsilon^{-2} K\, Q_1 + \\ &\qquad + \nabla_{Q_1} \frac{P_1^T G M^{-1} G^T P_1 + P_2^T B M^{-1} B^T P_2}{2} \\ &= \Lambda + \nabla_{Q_1} V_F + \nabla_{Q_1} \frac{P_1^T G M^{-1} G^T P_1 + P_2^T B M^{-1} B^T P_2}{2} \end{aligned}$$

and

$$\Lambda \;=\; O(\delta)$$

where we used that $GM^{-1}\tilde{G}^T = GM^{-1}G^T + O(\epsilon^2)$ and

$$0 \;=\; \tilde{g}(Q) \;=\; \epsilon^2 K^{-1} \nabla_{Q_1} V(Q) + Q_1 \,.$$

Thus

$$B(Q)M^{-1}\tilde{G}(Q)\Lambda \;=\; O(\delta\,\epsilon^2)$$

and, since

$$P_1^T G(Q) M^{-1} G(Q)^T P_1 \;=\; O(\epsilon^4)\,,$$

the equations (33) are equivalent to (31) up to terms of order $O(\delta\,\epsilon^2)$. Standard perturbation results for differential equations (see, e.g., [23]) imply that the same is true for the solutions over bounded intervals of time. □

Corollary 1. *An order $O(\epsilon^2)$ approximation of the smoothed dynamics of (1) over bounded intervals of time is given by the constrained Hamiltonian system*

$$\begin{aligned} \frac{d}{dt}Q &= M^{-1}P \\ \frac{d}{dt}P &= -\nabla V(Q) - \nabla V_F(Q) - G(Q)^T \Lambda \\ 0 &= g(Q) \end{aligned} \tag{34}$$

with Hamiltonian

$$H_\epsilon(Q,P) \;=\; \frac{P^T M^{-1} P}{2} + V(Q) + V_F(Q) + g(Q)^T \Lambda$$

and the (hard) constraint manifold

$$\mathcal{M}_\epsilon \;=\; \{(Q,P) \in R^{2n} : g(Q) = 0, G(Q)M^{-1}P = 0\}\,.$$

Furthermore, if the potential V_F is set equal to zero in (34), then the approximation is only of order $O(\delta)$.

Proof. Same as proof of Theorem 1, except that

$$Q_1 \;=\; 0$$

provides only an approximation of order $O(\epsilon^2)$ to $\langle q_1 \rangle_{\sqrt{\epsilon}}$. □

Remarks. (i) Corollary 1 implies that the smoothed dynamics of (1) cannot, in general, be approximated by the *slow solutions* of (1) as introduced by Kreiss in [9]. One can show that, up to terms of order (ϵ^2), the slow solutions of (1) are given by the constrained equations (13) which differ from (34) by the Fixman potential (30) and thus by a term of order $O(\delta)$.

(ii) A similar result to Corollary 1 has been published before, e.g., by van Kampen [27] and Pear & Weiner [17] in the context of statistical mechanics. In [21], Rubin & Unger considered in detail the case $p_1(0) = 0$ which leads to the formulation (13) and the case $p_1(0) \neq 0$ for a single constraint; i.e. $m = 1$.

An important aspect of Hamiltonian systems is the presence of symmetries which imply the conservation of the corresponding momentum maps (first integrals) [13]. Here we have the following

Theorem 2. *Let a Lie group Γ be a symmetry of (1) [13], i.e., $H(\gamma q, \gamma^{-T} p) = H(q,p)$ for all $\gamma \in \Gamma$, then Γ is also a symmetry of the constrained system (32), (34) respectively.*

Proof. We have to show that $H_\epsilon(\gamma Q, \gamma^{-T} P) = H_\epsilon(Q,P)$ for all $\gamma \in \Gamma$. Since $G(\gamma Q) = G(Q)\gamma^{-1}$, $\gamma^{-1}M^{-1}\gamma^{T} = M^{-1}$, $\nabla V(\gamma Q) = \gamma^{-T}\nabla V(Q)$, and

$$\begin{aligned} G(\gamma Q)M^{-1}G(\gamma Q)^T &= G(Q)\gamma^{-1}M^{-1}\gamma^{-T}G(Q)^T \\ &= G(Q)M^{-1}G(Q)\,, \end{aligned}$$

we indeed have $V_F(Q) = V_F(\gamma Q)$ and $\tilde{g}(Q) = \tilde{g}(\gamma Q)$. □

Example 1. To show the effect of our two approximations (34) and (32) to the slow dynamics, we looked at a one-dimensional chain of two soft and

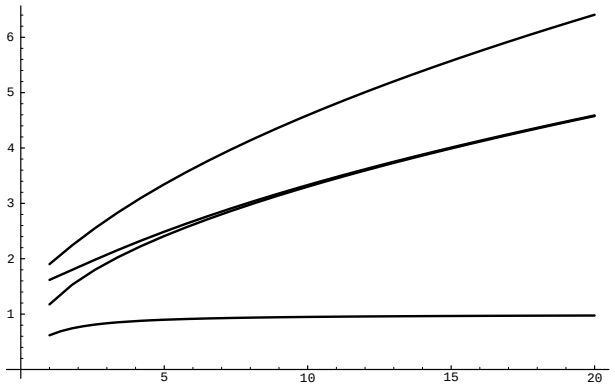

Figure 1: Natural frequencies of the unconstrained system as a function of the parameter $1/\epsilon^2$.

three hard springs with both ends of the chain held fixed. The Hamiltonian, we used, is given by

$$\begin{aligned} H(q,p) &= \frac{p^T p}{2} + \frac{1}{2\epsilon^2}([q_1 - 1]^2 + [q_3 - q_2 - 1]^2 + [5 - q_4 - 1]^2) \\ &\quad + \frac{1}{2}([q_2 - q_1 - 1]^2 + [q_4 - q_3 - 1]^2) \, . \end{aligned}$$

We computed the natural frequencies of the corresponding (linear) unconstrained system (1) (Fig. 1) and compared those to the ones obtained for the (linear) constrained system (34) with hard constraints (Fig. 2) and those with flexible constraints (32) (Fig. 3). Note that the smoothed dynamics is given by the smallest natural frequency of the unconstrained system; i.e., (9) acts here as a low-pass filter cutting off all frequencies except the lowest one. While both constrained methods correctly eliminate the three highest frequencies in the system, the low frequency component is far better approximated by the system (32) with flexible constraints. This is crucial especially for moderate values of $1/\epsilon^2$. (Note that for linear problems the Fixman potential is constant and does not need to be included into the constrained dynamics and that Theorem 1 applies with $\delta = \epsilon^2$.)

7 Smoothed Dynamics and Time Averages

In this section we want to investigate how accurate the smooth solutions of (32) approximate time averages of the highly oscillatory system (1). To

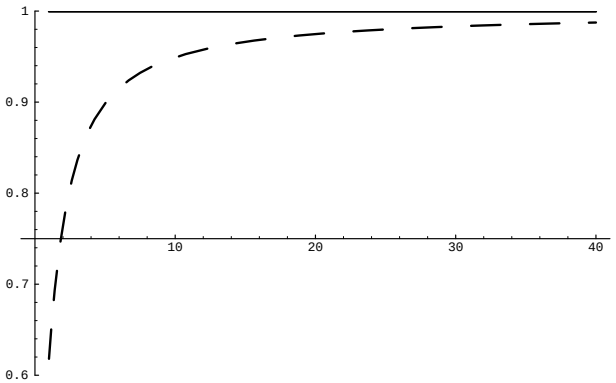

Figure 2: Natural frequency of the constrained system with hard constraints compared to the lowest frequency of the unconstrained system (dashed line) as a function of the parameter $1/\epsilon^2$.

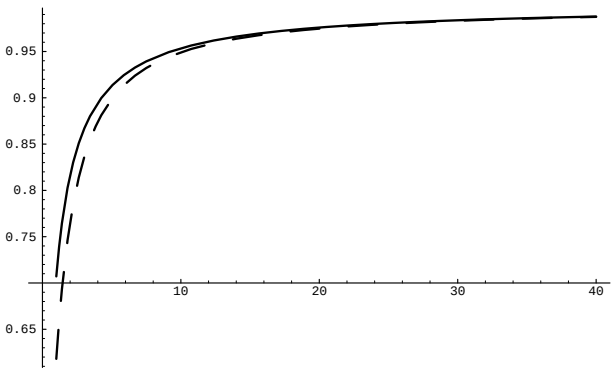

Figure 3: Natural frequency of the constrained system with flexible constraints compared to the lowest frequency in the unconstrained system (dashed line) as a function of the parameter $1/\epsilon^2$.

be more precise, let $\mathcal{A}(q)$ be an observable of the system (1) and let $\langle \mathcal{A}(q) \rangle$ denote its time-average along a solution $q(t)$ of (1). Here we assume that (1) is ergodic and, thus, $\langle \mathcal{A}(q) \rangle$ does not depend on the chosen trajectory $q(t)$. We are interested how accurate, if at all, the time-average $\langle \mathcal{A}(Q) \rangle$ of $\mathcal{A}$ along trajectories of (32) approximates $\langle \mathcal{A}(q) \rangle$.

We have first to review a few more facts from statistical mechanics. Let $H(q_1, p_1, q_2, p_2)$ be an arbitrary Hamiltonian in the variables (q_1, p_1, q_2, p_2). As in Section 3, $\rho^{(1)}_{\rm ens}(q_1, p_1, q_2, p_2)$ denotes the corresponding reduced density function and we consider the ensemble average over the variable (q_1, p_1); i.e.

$$\langle \mathcal{A} \rangle^{(1)}_{\rm ens}(q_2, p_2) \;=\; \int\int \mathcal{A}(q_1, q_2) \rho^{(1)}_{\rm ens}(q_1, p_1, q_2, p_2)\, dq_1 dp_1 \,.$$

The free energy $\mathcal{H}(q_2, p_2)$ of the remaining variables (q_2, p_2) is a function that satisfies

$$\nabla_{q_2} \mathcal{H}(q_2, p_2) \;=\; \langle \nabla_{q_2} H \rangle^{(1)}_{\rm ens}(q_2, p_2)$$

and

$$\nabla_{p_2} \mathcal{H}(q_2, p_2) \;=\; \langle \nabla_{p_2} H \rangle^{(1)}_{\rm ens}(q_2, p_2)$$

[10]. Furthermore, let $\rho^{(2)}_{\rm ens}(q_2, p_2)$ denote the density function corresponding to the (Hamiltonian) free energy $\mathcal{H}(q_2, p_2)$, then the total ensemble average

$$\langle \mathcal{A} \rangle_{\rm ens} \;:=\; \int\int\int\int \mathcal{A}(q_1, p_1, q_2, p_2) \rho_{\rm ens}(q_1, p_1, q_2, p_2)\, dq_1 dp_1 dq_2 dp_2$$

satisfies

$$\langle \mathcal{A} \rangle_{\rm ens} \;=\; \langle \langle \mathcal{A} \rangle^{(1)}_{\rm ens} \rangle^{(2)}_{\rm ens}$$

with

$$\langle \langle \mathcal{A} \rangle^{(1)}_{\rm ens} \rangle^{(2)}_{\rm ens} \;:=\; \int\int \langle \mathcal{A} \rangle^{(1)}_{\rm ens}(q_2, p_2) \rho^{(2)}_{\rm ens}(q_2, p_2)\, dq_2 dp_2$$

[10].

Let us now return to the system (16) and (17). By our Quasi-Stationarity Assumption, the smoothed equations (33) are given, up to terms of order $O(\delta\, \epsilon^2)$, by

$$\begin{aligned} \frac{d}{dt} Q_2 &= +\langle \nabla_{p_2} H \rangle^{(1)}_{\rm ens}(Q_2, P_2) \\ \frac{d}{dt} P_2 &= -\langle \nabla_{q_2} H \rangle^{(1)}_{\rm ens}(Q_2, P_2) \end{aligned}$$

The corresponding free energy $\mathcal{H}(Q_2, P_2)$ is, up to terms of order $O(\delta\,\epsilon^2)$, given by the the restriction of H_ϵ of Theorem 1 to the constraint manifold $\mathcal{M}_\epsilon$ and subsequent canonical coordinate transformation $(Q, P) \in \mathcal{M}_\epsilon \to (Q_2, P_2)$. Thus, using the same notations as above, the ensemble average of an observable $\mathcal{A}(q_1, q_2)$ is given by

$$\langle \mathcal{A} \rangle_{\rm ens} \;=\; \langle \langle \mathcal{A} \rangle^{(1)}_{\rm ens} \rangle^{(2)}_{\rm ens} + O(\delta\,\epsilon^2)\,. \tag{35}$$

Now, since

$$\langle \mathcal{A} \rangle^{(1)}_{\rm ens}(Q_2) \;=\; \mathcal{A}\,(Q_1, Q_2) + O(\delta\,\epsilon^2)$$

with Q_1 given as a function of Q_2 by (28); i.e.

$$\begin{aligned} Q_1 \;&=\; \langle q_1 \rangle^{(1)}_{\rm ens}(Q_2, P_2) \\ &=\; \epsilon^2\, K^{-1}\, \nabla_{Q_1} V(0, Q_2) + O(\delta\,\epsilon^2)\,, \end{aligned}$$

(35) simplifies to

$$\langle \mathcal{A} \rangle_{\rm ens} \;=\; \langle \mathcal{A} \rangle^{(2)}_{\rm ens} + O(\delta\,\epsilon^2) \tag{36}$$

where

$$\langle \mathcal{A} \rangle^{(2)}_{\rm ens} \;=\; \int\!\!\int \mathcal{A}(Q_1(Q_2), Q_2)\, \rho^{(2)}_{\rm ens}(Q_2, P_2)\, dQ_2 dP_2\,.$$

Formula (36) has a remarkable consequence. Upon assuming that (1) is ergodic (in the presence of symmetries, ergodic on the level sets of the corresponding first integrals); i.e.

$$\langle \mathcal{A} \rangle \;=\; \langle \mathcal{A} \rangle_{\rm ens} \tag{37}$$

we obtain

Theorem 3. *Let $\mathcal{A}(q)$ be an observable of the system (1). Assume that (1) as well as the smoothed system (32) are ergodic. Then the time-average of $\mathcal{A}$ along trajectories $q(t)$ of (1) is reproduced by the time-average of $\mathcal{A}$ along trajectories $Q(t)$ of (32) up to terms of order $O(\delta\,\epsilon^2)$.*

Proof. Ergodicity of (32) implies that

$$\langle \mathcal{A}(Q) \rangle \;=\; \langle \mathcal{A}(Q) \rangle^{(2)}_{\rm ens}\,.$$

The result follows from equations (36) and (37). □

8 Numerical Discretization

Any constrained Hamiltonian system of the form (13) can efficiently be discretized by the SHAKE extension [22]

$$\begin{aligned} Q_{k+1} &= Q_k + \Delta t\, M^{-1} P_{k+1/2} \\ P_{k+1/2} &= P_{k-1/2} - \Delta t\, [\nabla V(Q_k) + G(Q_k)^T \Lambda_k] \\ 0 &= g(Q_{k+1}) \end{aligned} \tag{38}$$

of the Verlet scheme [28] which requires now the solution of an implicit equation in the variable Λ_k. It has been shown [12] that this scheme preserves the symplectic structure [13],[24] of Hamiltonian flows, is time-reversible, and conserves first integrals related to symmetries of the system [29],[19]. Furthermore, as shown in [18], the numerical solutions can asymptotically be considered as the exact solution of a perturbed constrained Hamiltonian system.

The same scheme can also be applied to the Hamiltonian system (32) with flexible constraints. This time we obtain

$$\begin{aligned} Q_{k+1} &= Q_k + \Delta t\, M^{-1} P_{k+1/2} \\ P_{k+1/2} &= P_{k-1/2} - \Delta t\, [\nabla V(Q_k) + \nabla V_F(Q_k) + \\ &\qquad + \frac{G(Q_k)^T K g(Q_k)}{\epsilon^2} + \tilde{G}(Q_k)^T \Lambda_k] \\ 0 &= \tilde{g}(Q_{k+1}) \end{aligned} \tag{39}$$

Again the method is symplectic, time-reversible, and momentum conserving.

The method (39) is computational expensive. An effective implementation of (39) and the discretization of (32) by less expensive methods can be found in [3] and [11]. Note that one could also discretize (32) by a proper modification of the energy-momentum methods proposed in [25].

Example 2. In this example we consider a three-bead-two-bond structure where the structure is restricted to move in a finite volume by the potential

$$V_r(q) = \sum_i K_r \left(\frac{r_i}{\sigma}\right)^{12}.$$

Here r_i denotes the distance of each of the three beads to the origin, $\sigma = 3.0$, and $K_r = 50.0$. We set the mass of all three beads equal to $m = 1$ and choose $r_0 = 1$ as the equilibrium bond-length and $\phi_0 = 120^o$ as the equilibrium bond-angle. The force constant for the harmonic bond-angle bending

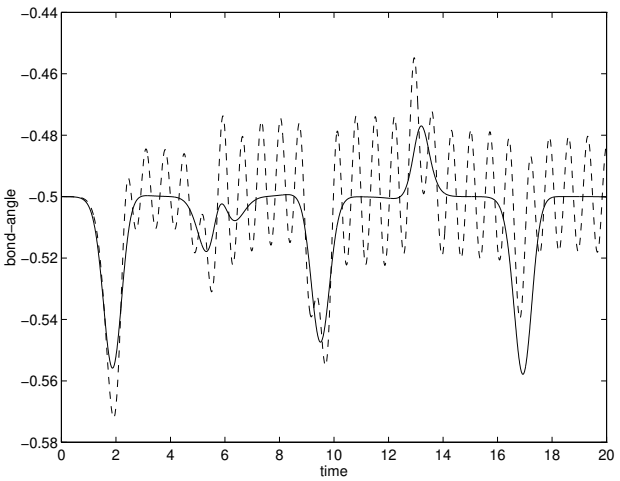

Figure 4: Trajectory of the bond-angle for the unconstrained formulation (dashed line) compared to the one for the formulation with hard constraints on the bond-lengths and soft constraints on the bond-angle (solid line).

potential was $K_a = 50\,\mathrm{deg}^{-2}$ and $K_b = 500$ for the corresponding bond stretching potentials. Note that these values correspond to force constants typically found in molecular dynamics simulations [26].

We started the structure from its equilibrium position with the initial velocities in x-direction equal to $p_x = 1.0$. The impact of the structure clashing with the potential-wall V_r can be seen in Fig. 4. We computed the trajectory of the bond-angle (plotted as $\cos(\phi)$) for the unconstrained formulation and compared this trajectory with the one obtained by constraining the bond-lengths by hard constraints and the bond-angle by a soft constraint. Note that for this particular example the Fixman potential (30) is almost constant and needed not to be included into the calculations. This implies also that equipartitioning of kinetic energy is not necessary to reproduce the smoothed dynamics up to terms of order $O(\delta\,\epsilon)$. In fact, we choose this example to demonstrate the effect of using soft constraints in the bond-angles versus hard constraints which, in our case here, would freeze the bond-angle at a value $\cos(\phi) = -0.5$.

9 Concluding Remark.

We have provided a theoretical framework for removing the fastest degrees of motion in highly oscillatory Hamiltonian systems as they arise, for example, in molecular dynamics. Our approach can be understood as a mean force field approach where the force field is obtained by an averaging process over

the fastest degrees of motion. As well known from other mean force field approaches (for example, Brownian motion), fluctuations play an important role in reproducing correct rate constants [4]. This requires an embedding of the mean force field into stochastic dynamics [4]. For our particular system (32), this will be discussed in a forthcoming publication [20].

Acknowledgements. This work was started while the author was visiting the Beckman Institute in Urbana-Champaign. We like to thank Klaus Schulten for providing a very stimulating environment and Robert D. Skeel for many fruitful discussions.

References

[1] Allen, M.P. and Tildesley, D.J., *Computer simulations of liquids*, Oxford University Press, New York, 1987.

[2] Boothby, W.M., *An introduction to differentiable manifolds and Riemannian geometry*, 2nd edition, Academic Press, 1986.

[3] Brooks, B.R., Zhou, J., and Reich, S., Elastic molecular dynamics with flexible constraints, in preparation.

[4] Brooks III, C.L., Karplus, M., and Pettitt, B.M., *Proteins: A theoretical perspective of dynamics, structure, and thermodynamics*, John Wiley & Sons, New York, 1988.

[5] Daubechies, I., *Ten lectures on wavelets*, SIAM, Philadelphia, 1992.

[6] Fenichel, N., Geometric singular perturbation theory for ordinary differential equations, *J.D.E.*, **31**, 53–98, 1979.

[7] Fixman, M., Classical statistical mechanics of constraints: A theorem and applications to polymers, *Proc. Nat. Acad. Sci.*, **71**, 3050–53, 1974.

[8] Kopell, N., Invariant manifolds and the initialization problem for some atmospheric equations, *Physica D*, **14**, 203–215, 1985.

[9] Kreiss, H.-O., Problems with different time scales for ordinary differential equations, *SIAM J. Numer. Anal.*, **16**, 980–998, 1979.

[10] Kuczera, K., Conformational free energy from thermodynamic integration simulations, Techn. Report, University of Kansas, Lawrence, 1995.

[11] Leimkuhler, B., Reich, S., and Skeel, R.D., Integration methods for molecular dynamics, to appear in *IMA Volumes in Mathematics and its Applications*, Mesirov, J. and Schulten, K., editors, Springer Verlag, 1995.

[12] Leimkuhler, B. and Skeel, R.D., Symplectic numerical integrators in constrained Hamiltonian systems, *J. Comput. Phys.*, **112**, 117–125, 1994.

[13] Marsden, J.R. and Ratiu, T., *An introduction to mechanics and symmetry*, Springer Verlag, 1994.

[14] McCammon, J.A. and Harvey, S.C., *Dynamics of proteins and nucleic acids*, Cambridge University Press, 1987.

[15] Nipp, K. and Stoffer, D., Invariant manifolds of numerical integration schemes applied to stiff systems of singular perturbation type – Part I: RK-methods, Research Report 92-14, Seminar für Angewandte Mathemathik, ETH Zürich, 1992.

[16] Pathria, P.K., *Statistical mechanics*, Pergamon Press, Oxford, 1972.

[17] Pear, M.R. and Weiner, J.H., Brownian dynamics study of a polymer chain of linked rigid bodies, *J. Chem. Phys.*, **71**, 212–224, 1979.

[18] Reich, S., Symplectic integration of constrained Hamiltonian systems by composition methods, *SIAM J. Numer. Anal.*, to appear.

[19] Reich, S., Momentum conserving symplectic integrators, *Physica D*, **76**, 375–383, 1994.

[20] Reich, S., A free energy approach to the torsion dynamics of macromolecules, in preparation, 1995.

[21] Rubin, H. and Ungar, P., Motion under a strong constraining force, *Comm. Pure Appl. Math.*, **10**, 65–87, 1957.

[22] Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C., Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes, *J. Comput. Phys.*, **23**, 327–342, 1977.

[23] Sanders, J.A. and Verhulst, F., *Averaging methods in nonlinear dynamical systems*, Springer Verlag, 1985.

[24] Sanz-Serna, J.M. and Calvo, M.P., *Numerical Hamiltonian Problems*, Chapman and Hall, London, 1994.

[25] Simo, J., Tarnow, N., and Wong, K.K., Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics, *Comp. Meth. Appl. Mech. Eng.*, **1**, 63–116, 1992.

[26] van Gunsteren, W.F. and Karplus, M., Effects of constraints on the dynamics of macromolecules, *Macromolecules*, **15**, 1528–44, 1982.

[27] van Kampen, N.G. and Lodder, J.J., Constraints, *Am. J. Phys.*, **52**, 419–424, 1984.

[28] Verlet, L., Computer experiments on classical fluids, Part I, *Phys. Rev.*, **159**, 98–103, 1967.

[29] Zhang, M.-Q. and Skeel, R.D., Symplectic integrators and the conservation of angular momentum, *J. Comp. Chem.*, to appear.