

Chapter 8

Flow of a Biomolecule

We leave our analysis of fluid flows for the moment, and consider the flow in a second order dynamical system. The system we study is a coupled oscillator system that models behavior of a linear biomolecule [Du Toit 2009a]. At the end of the analysis, we shall see that the LCS approach reveals the mechanism that governs transport in the flow, and that the underlying transport structure is the homoclinic tangle. Before arriving at this point, however, we will need first to provide a suitable reduction of the full system, which has hundreds of degrees of freedom, to a reduced single degree of freedom system that adequately captures the relevant transport dynamics.

Averaging over fast variables is a widely used method to obtain coarse equations of motion in mechanical systems with many degrees of freedom [Sanders 2007]. For example, averaging methods have been successfully used to find accurate coarse models in celestial mechanics and in oscillating electrical circuits [Verhulst 2000]. The system of nonlinearly perturbed coupled oscillators under study in this chapter, however, exhibits resonances on all scales and consequently does not admit analysis using standard averaging techniques. Furthermore, the full system, as will be demonstrated using a simple bio-mechanical example, has interesting dynamics that includes spontaneous and coherent changes in global conformation, and a reduction of the system using straightforward truncation methods fails to capture the crucial influence of the fine-scale dynamics that induces this conformation change.

The class of nonlinear systems of coupled oscillators that we study are close to a coupled chain of linear harmonic oscillators. Such near-integrable systems have been

studied in [Forest 1992] where transition to equipartition of energy, and dynamical properties related to integrable instability theory of partial differential equations were investigated [Ercolani 1990]. Here we discuss the representation of dynamics of the full oscillator system as a time-dependent single degree of freedom oscillator that represents well certain aspects of the full dynamics such as coherent switching between equilibria.

We begin by presenting an approximation to the full coupled oscillator system that allows for the derivation of a single coarse equation that retains essential contributions from the higher order components. Moreover, the resulting single degree of freedom system faithfully captures the statistics of the interesting conformation change behavior observed in the full system. The approximation involves, in essence, replacing higher order components in the perturbed problem with corresponding analytic trajectories for the nearby linear system.

For the biomolecular example system, we also investigate robust actuation of conformation change and demonstrate that low-powered traveling wave perturbations provide an efficient means for achieving near optimal conformation change. Finally, we will apply the LCS method to visualize transport structures in the dynamics describing the coarse variables and consequently gain insight into the transport mechanisms that allow for global conformation change to occur.

8.1 Coarse variables and models

Consider the following system of ordinary differential equations:

$$\ddot{\theta}(t) + D \cdot \theta(t) = \epsilon F(\theta(t), t) \tag{8.1}$$

with initial conditions

$$\theta(0) = a \qquad \dot{\theta}(0) = b,$$

where a , b , and $\theta(t)$ are vectors in \mathbb{R}^N , $F : \mathbb{R}^N \times \mathbb{R} \rightarrow \mathbb{R}^N$ is a nonlinear time-dependent mapping, $\epsilon > 0$ is a small parameter that controls the size of the nonlinear

perturbation, and $D : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is a linear mapping that has the tri-diagonal matrix representation

$$D = \begin{bmatrix} 2 & -1 & 0 & & & 0 & -1 \\ -1 & 2 & -1 & 0 & & & 0 \\ 0 & -1 & 2 & -1 & 0 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & 0 & -1 & 2 & -1 & 0 \\ 0 & & & & 0 & -1 & 2 & -1 \\ -1 & 0 & & & & 0 & -1 & 2 \end{bmatrix} .$$

The smoothness conditions we require are that $F(\theta, t)$ is continuous in t and Lipschitz-continuous in θ . Equation (8.1) and the initial conditions can be written in component form using subscripts to denote the indices as

$$\begin{aligned} \ddot{\theta}_k &= \theta_{k+1} - 2\theta_k + \theta_{k-1} + \epsilon F_k(\theta, t) \\ & \qquad \qquad \qquad k = 1, \dots, N, \\ \theta_k(0) &= a_k \qquad \qquad \dot{\theta}_k(0) = b_k \end{aligned}$$

where $\theta_0 = \theta_N$ defines the periodic boundary condition.

Such a system arises naturally from a spatial discretization of a nonlinearly perturbed wave equation with periodic boundary conditions:

$$u_{tt}(x, t) = u_{xx}(x, t) + \epsilon G(x, u(x, t), t),$$

where $u : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is the amplitude of the wave in time and space (subscripts denote partial differentiation), and the real-valued function G represents the perturbation. In this context, the D matrix in equation (8.1) is simply a centered finite differencing operator that approximates the second partial derivative of u with respect to the spatial variable x , and F is obtained by evaluating G at uniformly discrete spatial positions x_k so that $F_k(\theta, t) := G(x_k, \theta, t)$, $\theta_k(t) := u(x_k, t)$, and $x_0 = x_N$.

Equivalently, equation (8.1) can be viewed as the dynamical system describing a

linear chain of N oscillators in which each oscillator is subject to a weak nonlinear potential and coupled to nearest neighbors through a harmonic potential; a specific example of such an oscillator chain related to the mechanics of biomolecules will be provided later. E. Fermi, J. Pasta, S. Ulam, and M. Tsingou used a similar chain of coupled oscillators as an example system in their pioneering numerical study of nonlinear dynamics that has since become famously known as the Fermi-Pasta-Ulam-Tsingou¹ (FPUT) problem [Fermi 1955]. The initial purpose of their study was to develop a theory of thermalization in systems with nonlinear dynamics; however, their investigation yielded unexpected results – energy initially placed in one mode did not become equally partitioned among all the modes after some time. Rather, they observed recurrences (as predicted by Poincaré) where the energy initially redistributed among some of the modes but then returned to the initial condition in which all the energy is again found in a single mode. Analyses of the FPUT problem fill a large body of literature including, for instance, the discovery of soliton and chaotic breather solutions, and have yielded insight into the interplay between chaos and integrability in nonlinear systems (see reference [Campbell 2005] for a survey of results related to the FPUT problem on the fiftieth anniversary of the introduction of the problem). The approach to the problem addressed in this chapter differs from traditional FPUT problem analyses in that, motivated by the biomolecular applications to DNA, we are interested in issues such as reduction to coarse variables and activation of global large-scale conformation change through the application of small local controls that are not typically associated with the FPUT problem.

We study the system of ordinary differential equations described by equation (8.1) when N is large, and hence the system has many degrees of freedom. Rather than determine precisely the dynamics of each degree of freedom, we are interested in describing the dynamics of only a single coarse variable. The first question to be addressed here is the following: how do we extract from the large N degree of freedom

¹Historically, this problem has been referred to as the Fermi-Pasta-Ulam (FPU) problem; however, recently the contribution of Mary Tsingou to the implementation of the numerical routines has been more widely appreciated thus motivating the addition of her name. The recent article in [Dauxois 2008] provides a discussion of the relevant history.

system an evolution equation for a single coarse variable that describes a property of interest? For the purposes of the current study, our goal is to determine a single evolution equation for the dynamics of the *average* amplitude while faithfully retaining salient dynamical features of the full system. As will be shown later, the dynamics of the coarse variable will need to include the influence of the fine scales in order to reproduce the coarse evolution correctly. We begin by first gaining insight from the unperturbed case.

8.1.1 The unperturbed case

For $\epsilon = 0$, the system in equation (8.1) becomes:

$$\ddot{\theta}(t) + D \cdot \theta(t) = 0, \quad (8.2)$$

with initial conditions

$$\theta(0) = a \quad \dot{\theta}(0) = b.$$

This is a simple linear system whose solution is easily obtained analytically. The solution is provided here in detail as it includes many building blocks required for the less tractable case when the nonlinear perturbation is included.

We begin by introducing a change of coordinates that diagonalizes the coupling matrix D . Let P be an $N \times N$ matrix whose columns contain the complete set of orthonormal eigenvectors of the real symmetric matrix D :

$$P_{kw} := \sqrt{\frac{2}{N}} \begin{bmatrix} \frac{1}{\sqrt{2}} & \cos \frac{2\pi kw}{N} & \frac{(-1)^k}{\sqrt{2}} & \sin \frac{2\pi kw}{N} & \begin{matrix} [N \dots 1] = \eta \\ \vdots \\ 1 \end{matrix} \\ w=0 & w=[1 \dots \frac{N}{2}-1] & w=\frac{N}{2} & w=[\frac{N}{2}+1 \dots N-1] & \end{bmatrix}.$$

(Here we have taken N to be even for simplicity, although the case for odd N merely has the middle column corresponding to $w = \frac{N}{2}$ removed and the column numbering

altered accordingly.)

Next, we define new coordinates by the linear transformation,

$$\hat{\theta} := P'\theta \quad (8.3)$$

where P' denotes the transpose of P . Notice that

$$\hat{\theta}_0 := \frac{1}{\sqrt{N}} \sum_{k=1}^N \theta_k \quad (8.4)$$

is (except for a constant factor of \sqrt{N}) the average amplitude. In these coordinates, the symmetric linear operator D is diagonal, yielding N uncoupled second order ODEs. In component form, they are written as

$$\ddot{\hat{\theta}}_0 = 0 \quad (8.5a)$$

$$\ddot{\hat{\theta}}_w + \alpha_w^2 \hat{\theta}_w = 0 \quad w = 1, \dots, N-1, \quad (8.5b)$$

where $\alpha_w^2 := 2(1 - \cos \frac{2\pi w}{N})$, and the initial conditions become

$$\hat{\theta}(0) = P'a =: \hat{a} \quad \dot{\hat{\theta}}(0) = P'b =: \hat{b}.$$

Here we can immediately conclude that the evolution of the average amplitude in the unperturbed case is given by the single scalar equation

$$\hat{\theta}_0(t) = \hat{a}_0 + \hat{b}_0 t.$$

Straightforward solution of the higher order components yields the analytic evolution equations

$$\hat{\theta}_w(t) = \hat{a}_w \cos \alpha_w t + \frac{\hat{b}_w}{\alpha_w} \sin \alpha_w t \quad w = 1, \dots, N-1. \quad (8.6)$$

We readily observe that the system has N integrals of motion, $\{I_w\}_{w=0}^{N-1}$, defined by

$$\begin{aligned}\dot{\hat{\theta}}_0(t) &= (\hat{b}_0) =: \sqrt{2I_0} \\ \hat{\theta}_w(t)^2 + \left(\frac{\dot{\hat{\theta}}_w(t)}{\alpha_w}\right)^2 &= (\hat{a}_w)^2 + \left(\frac{\hat{b}_w}{\alpha_w}\right)^2 =: 2I_w \quad w = 1, \dots, N-1.\end{aligned}$$

The first integral is simply a statement of the conservation of total (or average) momentum. The higher order modes evolve on circles of constant radii $\sqrt{2I_w}$ with an angular frequency of α_w . With these insights, we may write the system in completely integrable Hamiltonian form using action angle coordinates (I, ϕ) , where I and ϕ are both vectors of length N whose components I_w and ϕ_w are denoted with subscripts:

$$H^0(I_0, \dots, I_{N-1}) = \sum_{w=1}^{N-1} \alpha_w I_w \quad (8.7)$$

so that the Hamiltonian vector field becomes

$$\begin{aligned}\dot{I}_w &= -\frac{\partial H^0}{\partial \phi_w}(I_0, \dots, I_{N-1}) = 0 \\ \dot{\phi}_w &= \frac{\partial H^0}{\partial I_w}(I_0, \dots, I_{N-1}) = \alpha_w \quad w = 0, \dots, N-1.\end{aligned}$$

Before proceeding to an analysis of the perturbed case, we first observe here some of the important properties of the unperturbed system.

Remark. *The Hamiltonian H^0 is degenerate.*

The Hamiltonian H^0 is linear in the components of I ; hence

$$\det \left(\frac{\partial^2 H^0}{\partial I_i \partial I_j}(I_0, \dots, I_{N-1}) \right) = 0 \quad i, j = 0, \dots, N-1, \quad (8.8)$$

and the frequency map $I \rightarrow \alpha(I)$ is not a diffeomorphism. This degeneracy of the Hamiltonian implies that the system does not satisfy the assumptions of classical KAM theorems and hence straightforward KAM theory cannot be applied [Wiggins 2003].

Remark. *The frequency vector α is resonant.*

Since the first element of the frequency vector α is 0, any $\kappa \in \mathbb{Z}^N \setminus \{0\}$ of the form

$\kappa = [z \ 0 \ \dots \ 0]'$ for any $z \in \mathbb{Z} \setminus \{0\}$ yields $\kappa \cdot \alpha = 0$. The irrational structure of the eigenvalues in the higher modes leads to the conclusion that the resonant frequency vector α is of multiplicity 1.

Furthermore, since z is any element in $\mathbb{Z} \setminus \{0\}$, we can conclude the following:

Remark. *The frequency vector α has resonances at all orders.*

A standard approach to achieve reduction of order in a dynamical system is to perform averaging over the fast angular coordinates. However, for the system of interest, the resonances, or more specifically the zero eigenvalue corresponding to $\alpha_0 = 0$ in the equation for the average amplitude, precludes such a treatment. Standard statements of averaging theorems require all the components of the frequency vector α to be strictly greater than 0 (see the statement of the Averaging Theorem in [Arnold 1978], for example, or the discussion in Chapter 8 of [Sanders 2007] regarding passage through resonance and the absence of theory to treat the fully resonant case). Hence, the oscillator system represents a special case to which routine averaging methods cannot be applied.

Moreover, in our attempt to obtain a single closed equation for the evolution of the average variable, straightforward averaging or truncation approaches are ineffective for an even more subtle—yet crucial—reason. Averaging over the higher order components (laying the resonance issues aside) yields a single degree of freedom, and hence an integrable system. As such, the reduced equation fails to capture important details in the dynamics arising from the inherent non-integrability, and in particular the intricate influence of the higher order components on the average mode. A thorough review of averaging methods is not within the scope or purpose of this thesis. However, an important point to be made is that a central feature of the proposed method for reduction is that it incorporates in an approximate yet effective way the influence of the higher order modes, and consequently more accurately captures non-trivial dynamics associated with the full non-integrable system.

8.1.2 Perturbed case

We use the same definitions and process used in the unperturbed case for the perturbed case. We begin by making the same linear change of coordinates using the matrix of eigenvectors P so that the transformed equations of motion become:

$$\ddot{\hat{\theta}}_0 = \epsilon P_{k0} F_k(P\hat{\theta}) \quad (8.9a)$$

$$\ddot{\hat{\theta}}_w + \alpha_w^2 \hat{\theta}_w = \epsilon P_{kw} F_k(P\hat{\theta}), \quad w = 1, \dots, N-1, \quad (8.9b)$$

where we sum over k from 1 to N on the right hand side.

As expected, the coordinate transformation diagonalizes the nearest neighbor coupling term, but the nonlinear forcing term remains globally coupled. At this stage, no approximations have been made and equations (8.9a) and (8.9b) recover the full solution exactly. For nonzero ϵ , the system is no longer integrable and the solutions that were previously observed to evolve on circles of fixed radii are perturbed.

At this point, we must consider how to break the coupling between the zeroth order equation for the average amplitude and the higher order modes in order to obtain a closed equation for the average amplitude. As previously mentioned, the system is not amenable to averaging methods because of the zero eigenvalue in the frequency vector. Rather, we introduce an approximation by replacing equations (8.9b) with the analytical solutions of the unperturbed linear system as defined in equation (8.6). This approach effectively removes the need to integrate the higher order modes by replacing their evolution with the explicit analytic solution of the nearby integrable system. In so doing, we obtain the improved approximation, $\bar{\theta}_0$, of the exact solution $\hat{\theta}_0$, whose dynamics is prescribed by

$$\ddot{\bar{\theta}}_0 = \frac{\epsilon}{\sqrt{N}} \sum_{k=1}^N F_k(P\bar{\theta}), \quad (8.10)$$

where in the right hand side P is a constant matrix, and $\bar{\theta}$ is a vector whose first component is the scalar dependent variable $\bar{\theta}_0$ and whose remaining components are

elementary functions of time and the initial conditions of the full system provided analytically by equation (8.6).

Remark. *The solution trajectories of equation (8.10) are within $\mathcal{O}(\epsilon)$ of the solution trajectories of the original full system described in equation (8.9a) for times $\mathcal{O}(1)$.*

This result is shown by applying a standard error analysis technique: substituting a formal expansion of the solution, using the Lipschitz continuity of F , and then applying the Gronwall lemma as is done in the proof of Theorem 9.1 in [Verhulst 2000], for example.

The approach proposed here for obtaining a closed equation for the average amplitude includes the influence of higher order modes by incorporating explicit time-dependence in the perturbing term, and hence leads to a one and a half degree of freedom system. The coordinate transformation and subsequent approximation yield a single *non-autonomous* ordinary differential equation that includes approximate dynamics for the higher order scales, and whose solution approximates the dynamics of the average angle of the system. In effect, the information contained in the higher order modes persists in the lower order description via the initial conditions.

The method just described is now applied to a simple coupled-oscillator model for biomolecules where retention of the influence of the higher modes in the dynamics is essential for accurately recovering nontrivial dynamics of conformation change.

8.2 Conformation change in biomolecules

Biomolecules undergo rapid and global conformation change as a crucial part of their function. Many statistical mechanical models have been proposed in which these conformation changes are the result of increased thermal fluctuations [Yakushevich 2004, Peyrard 2004, Weber 2006, Dauxois 1993, Causo 2000, Kafri 2002], or an external agent that provides an overwhelming force [Bhattacharjee 2000, Cocco 1999]. Presently, we are interested in investigating whether this conformation change phenomenon can be induced simply by utilizing the natural dynamics inherent to the system. A simple

model will be presented in which the intrinsic design of the dynamics ensures the robustness of conformational states to random perturbations, yet global conformation change can be robustly induced by the action of a low-energy local control. At the heart of this dynamical behavior is the exchange of energy between smaller and larger scales. The approximation technique presented in the first section allows passage to a low-dimensional model that effectively captures this behavior.

8.2.1 The model

We consider a class of biopolymers that can be modeled as a long circular chain of identical pendula attached to a rigid backbone. Each pendulum has one rotational degree of freedom in the plane orthogonal to the backbone. The motion of the pendula is governed by two interactions: each pendulum interacts with its nearest neighbors through a harmonic potential that models torsional coupling through the backbone; and secondly, each pendulum moves in a Morse potential that models the weaker hydrogen bonding interaction between pendulum pairs on a complementary chain. We immobilize one of the strands and consider only the motion of the pendula on the opposing strand as depicted in Figure 8.1.

Using the pendulum mass, m , for the mass scale; the pendulum length, h , for the length scale; and the nearest neighbor coupling strength, S , for the energy scale; the non-dimensional Lagrangian that describes the motion of N coupled pendula is given by

$$L(\theta, \dot{\theta}) = \sum_{k=1}^N \left[\frac{1}{2} \dot{\theta}_k^2 - \frac{1}{2} (\theta_k - \theta_{k-1})^2 - \epsilon \left(e^{-a[1 - \cos \theta_k - x_0]} - 1 \right)^2 \right],$$

where a , x_0 , and ϵ are the Morse potential decay coefficient, the Morse potential equilibrium distance, and the Morse potential amplitude respectively. The angular displacement of the k^{th} pendulum is denoted θ_k . The argument of the Morse potential is the distance between complementary base pairs projected onto the vertical. The time, t , has been scaled by the induced non-dimensional time, $\tau = \sqrt{S/(mh^2)} t$.

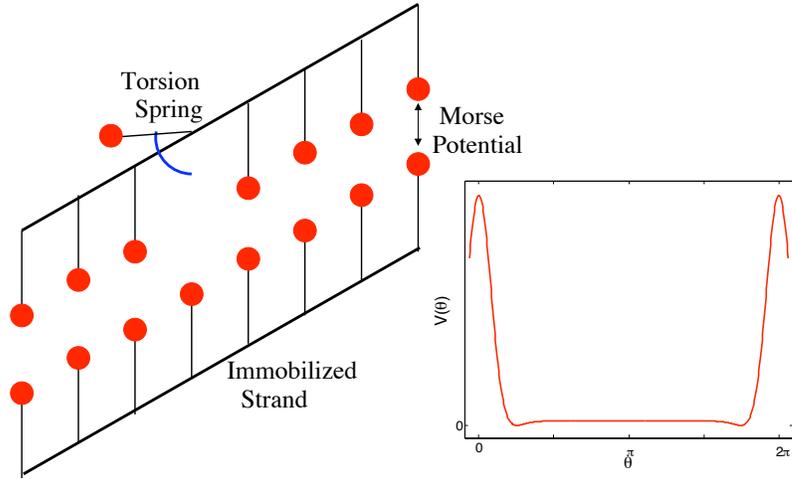


Figure 8.1: The biomolecule is modelled as a chain of pendula that rotate about the axis of a fixed backbone. The pendula interact with nearest neighbors along the backbone through harmonic torsional coupling, and with pendula on the opposing strand through a Morse potential.

Lagrange's equations of motion yield

$$\ddot{\theta}_k - \theta_{k+1} + 2\theta_k - \theta_{k-1} = 2\epsilon a \left(e^{-a[1-\cos\theta_k-x_0]} - 1 \right) e^{-a[1-\cos\theta_k-x_0]} \sin\theta_k$$

for $k = 1, \dots, N$. These equations can also be written in the form introduced in equation (8.1): $\ddot{\theta} + D \cdot \theta = \epsilon F(\theta)$ where D is the constant tridiagonal $N \times N$ matrix that describes the nearest neighbor coupling, and F is the nonlinear vector-valued forcing term due to the Morse potential.

Numerical integration of the full system of equations was performed for a chain of $N = 200$ pendula using a fourth order symplectic integrator with excellent energy preservation properties [McLachlan 1992]. The parameter values were chosen to best represent typical values for biomolecules. The nondimensional scales are determined using parameter values $m = 300$ AMU, $h = 1$ nm, and $S = 42$ eV, that collectively induce a model time unit of 0.272 ps. In nondimensional units, the Morse potential parameters are $a = 7$, $x_0 = 0.3$, and $\epsilon = 1/1400$. For these parameter values, the nonlinear Morse potential term represents a small perturbation to the linear nearest

neighbor coupling interaction.

8.2.2 Properties of the model

The shape of the Morse potential induces two stable equilibria corresponding to global energy minima achieved when all the pendula have identical angular displacements (thus nullifying the nearest neighbor coupling) and are positioned at the Morse potential equilibrium distance, x_0 , from their complementary pendula. For small energies, typical motions involve uncoordinated oscillations of the pendula near these stable equilibria, although Mezić has previously observed that a local perturbation can cause the pendula to undergo a coherent global change of conformation from one energy basin to the other [Mezic 2006]. By definition, we say that a global conformational change has occurred when the average angle of the pendula passes through π radians. This motion is referred to in the rest of this chapter as “flipping”. A convenient way to represent this flipping behavior is to project the trajectory of the system onto the average variables $\Theta := \frac{1}{N} \sum_{k=1}^N \theta_k$ and $\dot{\Theta}$ as shown in Figure 8.2(b). In this projection we see that the pendula at first oscillate about an energy minimum in one conformational state, and then undergo conformational change to the other energy basin where they continue to oscillate.

8.2.3 The reduced order model

As noted above, the average angle variables, Θ and $\dot{\Theta}$, provide a good coarse description of the flipping process and begs the derivation of a single closed equation for the average variable. There is, however, no separation of time scales in this system so that simple truncation to a low order model does not retain sufficient dynamics to incorporate spontaneous flipping events. Indeed, any method that yields an autonomous single degree of freedom system for the coarse variable cannot possibly capture the flipping event. Furthermore, as previously noted, routine averaging methods are not applicable since intrinsic resonances induce coupling on all scales.

The approach presented in the first section is now applied to this pendulum chain

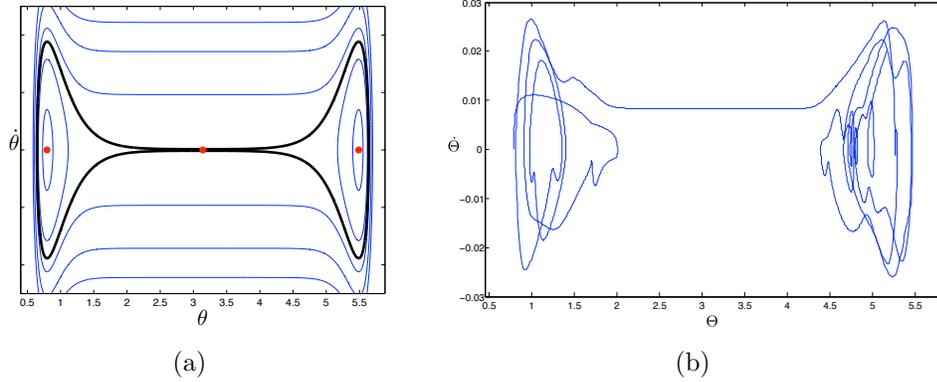


Figure 8.2: Figure (a) shows sample phase space trajectories of a single pendulum in the Morse potential when no coupling is present. The homoclinic trajectory is emphasized. The locations of the equilibrium points are marked with red dots. Trajectories inside the homoclinic trajectory associated with the equilibrium point at $(\pi, 0)$ are always oscillating, while those outside the homoclinic trajectory are always flipping. Figure (b) shows a single trajectory of the fully-coupled model projected onto the average variable phase space and indicates a flipping event from one conformational state to the other. The trajectory resembles the phase portrait for a single pendulum moving in the Morse potential; however, the harmonic nearest neighbor coupling provides resonant kicks that allow the trajectory to escape from one energy basin and then become trapped by the other.

to obtain a low order model that retains the essential influence of the higher order scales on the global flipping behavior. With no Morse potential ($\epsilon = 0$), the remaining linear system has an explicit solution in which the average velocity $\hat{\theta}_0$ is constant, and the remaining coordinate pairs $(\hat{\theta}_w, \hat{\theta}_w)$ (after scaling $\hat{\theta}_w$ by its corresponding frequency, α_w) evolve on circles of fixed radii. For small nonzero ϵ , this integrable solution is perturbed as shown in Figure 8.3. Certainly, deviations from the linear solution are evident in the lower modes, whereas the trajectories of the higher order components remain close to the integrable circular solutions of the unperturbed case. Numerical experiments reveal that as epsilon is increased, the higher order components in a typical flipping trajectory remain close to the unperturbed solution until epsilon has increased by a factor of 10. As epsilon increases further, the higher order components still exhibit oscillatory behavior, but the stronger nonlinear coupling causes large deviations from the trajectories along fixed radii and the circular

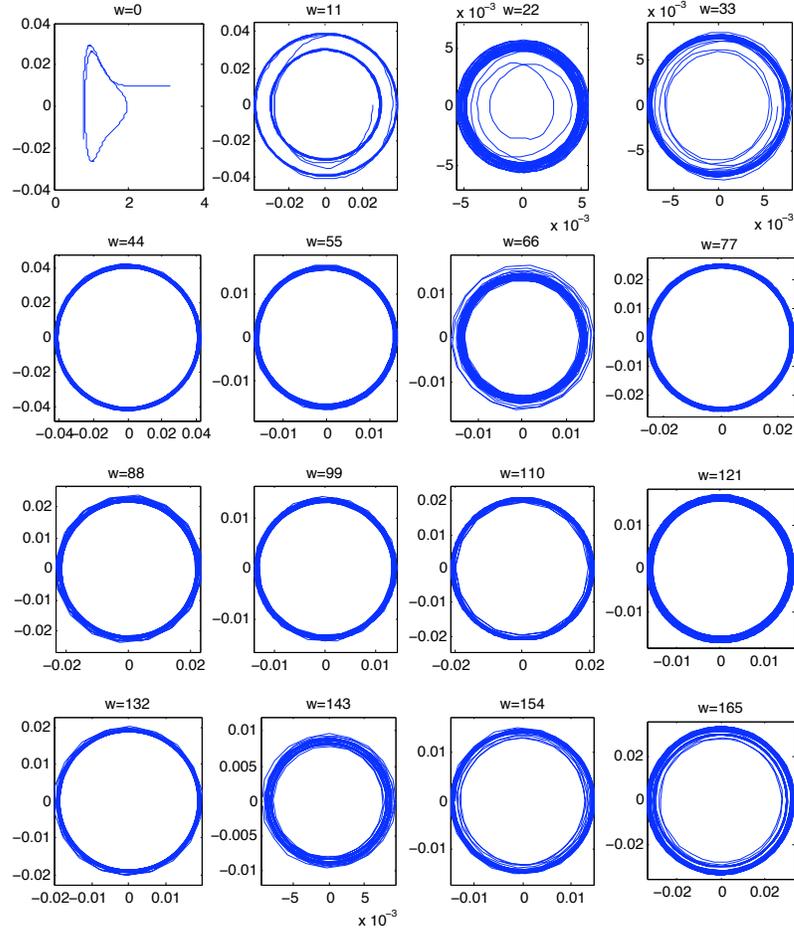


Figure 8.3: The evolution of a few of the $\hat{\theta}_w$ and $\dot{\hat{\theta}}_w$ variables is shown here. The average mode ($w = 0$) indicates that flipping occurs as in Figure 8.2(b) while the remaining modes simply oscillate with small deviations from the unperturbed ($\epsilon = 0$) integrable solution.

solutions disintegrate.

The $1\frac{1}{2}$ degree of freedom reduced system obtained using the approximation in equation (8.10) retains sufficient dynamics of the higher order modes in the time-dependent terms to capture statistics of the the flipping event remarkably well. The distributions of numerically computed flipping times for 5000 random initial conditions for the $1\frac{1}{2}$ degree of freedom reduced system and the full 200 degree of freedom system are compared in Figure 8.4. The histograms in Figures 8.4(a), 8.4(b), and 8.4(c) are computed using values of epsilon equal to 1/1400, 5/1400, and 10/1400

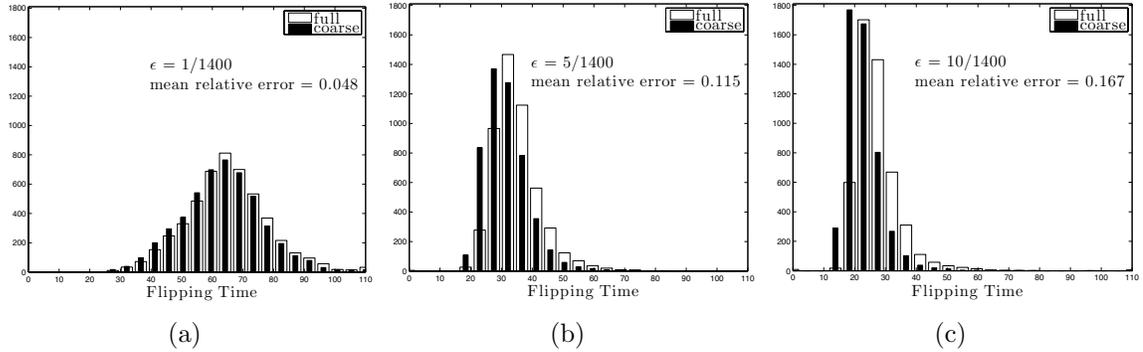


Figure 8.4: The histograms provided here compare the distribution of flipping times predicted by the $1\frac{1}{2}$ degree of freedom reduced model (shown in solid) with the distribution of flipping times predicted by the full 200 degree of freedom system (shown in outline) for different values of epsilon. The histograms are computed using 5000 random initial conditions. The plot in (a) was computed using a value of epsilon equal to $1/1400$ consistent with the biomolecule model, while (b) and (c) were computed using values of epsilon that are larger by a factor of 5 and 10 respectively. In (a) we see that the reduced model captures the distribution of flipping times remarkably well with a relative error of 4.8 percent. As expected, the relative error increases for larger values of epsilon.

respectively. When epsilon is equal to $1/1400$ (the value provided by the biomolecule model), the reduced model captures the shape of the flipping time distribution remarkably well and the mean relative error in the predicted flipping time is less than five percent. As expected, the relative error in the flipping time prediction increases as epsilon is increased. The quasi-periodic forcing introduced in the right hand side of equation (8.10) by the solution for the linear system provides the perturbation required to induce global flipping from a local perturbation and can be thought of as a time-dependent control.

The rigorous error estimate obtained in Section 8.1.2 stated that trajectories of the reduced model are within $\mathcal{O}(\epsilon)$ of the true solution for times $\mathcal{O}(1)$. The application of this error estimate to the biomolecule model merits the following two observations. First, the portion of a trajectory during which the pendulum experiences the Morse potential (which is responsible for inducing flipping motion) is very brief; the interaction occurs on a time-scale that is $\mathcal{O}(1)$ rather than the much longer $\mathcal{O}(1/\epsilon)$ time-scale, so that the estimate has validity, as evidenced numerically, for the

prediction of flipping times. Second, the error estimate was obtained for an arbitrary forcing function $F(\theta(t), t)$. In the biomolecule model, the exponential decay of the Morse potential with distance implies that when the pendula escape the immediate vicinity of the opposing pendula, the Morse potential and the consequent perturbation are effectively zero and the linear solution becomes nearly exact. Hence, the error estimate is conservative in that it utilizes only the fact that epsilon is small, and not the fact peculiar to the biomolecule model that in a large region of phase space the forcing term is also small. For these reasons, the numerically observed time-scale over which the reduced model for the biomolecule chain remains accurate is greatly increased over the rigorous estimate.

8.2.4 Efficient actuation of conformation change

The foregoing analysis indicates that the process of conformation change can be well-modelled by a low-dimensional model with a single degree of freedom. Next, we investigate the possibility of inducing this interesting flipping behavior using a low-powered local control.

Consider a pendulum chain in the stable equilibrium position. The minimum possible energy of a perturbation that induces flipping can be computed analytically and is

$$E_{\min} = N\epsilon \left(e^{-a[2-x_0]} - 1 \right)^2 \approx 0.143$$

and corresponds to each pendulum being kicked with the same initial velocity. With this lower energy limit for reference, we have investigated the robustness of the zipped state by numerically measuring the flipping time as a function of energy for random and structured perturbations. The results shown in Figure 8.5 indicate that the tendency for flipping is strongly influenced by the structure of the initial perturbation and confirm the results found in [Mezic 2006].

Higher mode numbers refer to sinusoidal perturbations of higher harmonic frequencies. We also include the energy required to induce flipping with random perturbations (labeled “noise”), as well as perturbations consisting of the displacement

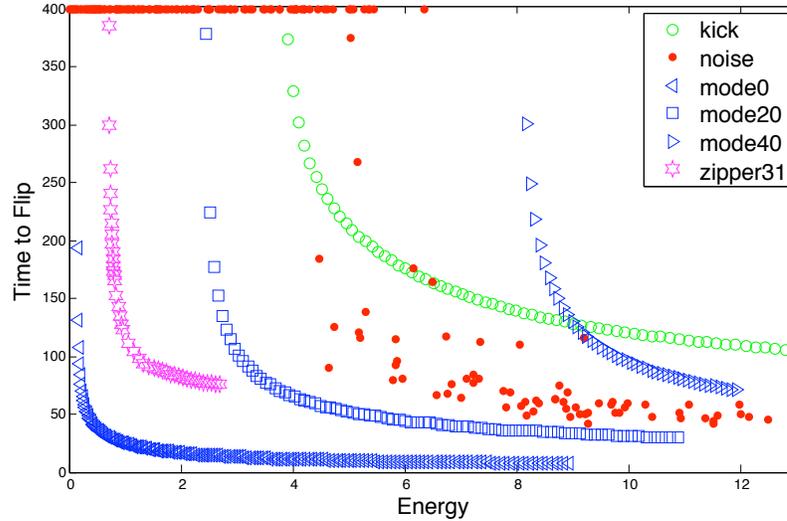


Figure 8.5: The time required for the chain to flip is plotted here as a function of the energy for various types of initial conditions. Green circles indicate initial conditions in which a single pendulum is displaced from the equilibrium position. Red dots indicate random initial conditions with gaussian distribution about the equilibrium. If flipping was not observed for this perturbation, a marker was placed at the 400 time mark at the top of the figure. The blue series in triangles and squares correspond to initial perturbations consisting of sinusoidal displacements with frequency indicated by the mode number. Magenta stars represent efficient “zippers” imparting 31 small kicks of varying strength as explained in the text. The amount of energy required to robustly induce flipping depends strongly on the frequency content of the initial condition. Less energy is required when the energy is placed in the low order modes. Random initial conditions must have energies above 6.0 to induce flipping whereas zippers require much less energy (0.75) to robustly induce flipping.

of a single pendulum (labeled “kick”). First, we notice that the equilibrium state is robust to very energetic random perturbations; energies in excess of 6.0 are required to robustly induce flipping. Second, flipping can be robustly induced by low-energy structured perturbations that comprise low frequency modes.

The magenta stars in Figure 8.5 refer to the following structured perturbation. Inspired by the unzipping action of polymerase and helicase proteins in DNA, we simulate the action of a hammer that moves along the strand with constant velocity and imparts a small kick to each pendulum as it passes. Despite its very low power requirement and localized interaction, this structured perturbation efficiently and

robustly induces global flipping motion by exciting the zeroth order collective mode. The action of the zipper comprises 31 small kicks (energy = 0.0245) providing a total energy of 0.75 to the pendulum chain. In contrast, the strand remains closed for random kicks that are two orders of magnitude larger than the individual kicks provided by the hammer. Hence, the pendulum chain has the interesting property that the conformation is robust against noise, yet allows for robust actuation of global conformation change through a small structured local control.

The fastest phonons in the pendulum dynamics move along the strand at a rate that is very near 1.0 pendulum per time unit. By increasing the speed with which the zipper moves down the chain well beyond this phonon speed, the total energy required to induce flipping can be made arbitrarily close to the theoretical minimum limit of 0.14. This observation follows from the fact that as the zipper moves faster, the time between kicks to subsequent pendula decreases and hence the perturbation approximates ever more closely the optimal perturbation of kicking each pendulum at the same instant. Thus, extremely (unrealistically) fast zippers can in effect provide the optimal global perturbation. On the other hand, by exploiting resonances in the dynamics, even slow-moving (more realistic) zippers are able to induce flipping with very low power consumption and only local interaction. The zippers referred to in Figure 8.5 move at a rate of only 0.4 pendula per time unit, well below the natural phonon velocity.

The zipper not only induces flipping more efficiently than the other perturbations considered, but also does so robustly when noise is added to the system. Figure 8.6 shows the energy required to induce rapid flipping (flipping occurs in less than 400 time units) when the system is simulated at constant temperature using a Brünger-Brooks-Karplus Langevin integrator with damping coefficient of 5 ns^{-1} [Brünger 1984]. For temperatures above 380K, flipping occurs before 10,000 time units without the action of any external perturbation. Notably, the zipper perturbation is able to robustly induce very fast flipping over a wide range of temperatures. Moreover, if the chain is at a low temperature and rigidly locked in a stable conformational state, conformational change can nevertheless be quickly induced.

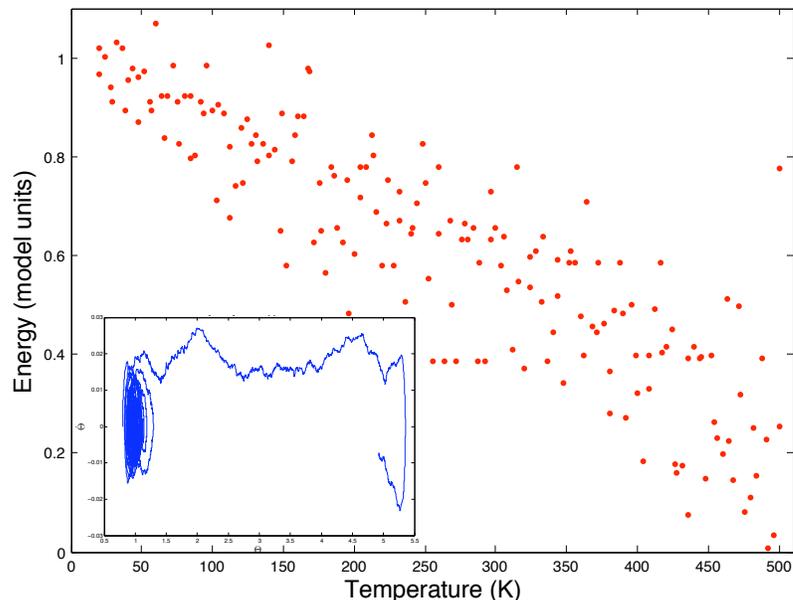


Figure 8.6: When the system is simulated at constant temperature, the zipper robustly and efficiently induces flipping. The energy required to induce flipping in less than 400 time units is shown for 120 simulations over a range of temperatures. The inset shows a sample trajectory induced by the zipper when the temperature is 310K and the zipper is activated after 2000 time units.

These observations in the simple chain of oscillators model yield insight into how biomolecule conformations can be robust to energetic stochastic perturbations (so that DNA, for example, does not spontaneously divide over a wide temperature range) and yet, with the right “trigger”, it can divide reliably and quickly with a low energy wave-like perturbation. We believe that this is a fundamental dynamical mechanism that biological systems use to utilize natural dynamics in order to robustly perform rapid and reliable conformation changes that are not accidentally caused by random perturbations, and do not require the application of overwhelming force.

8.3 Visualizing transport in the reduced model

A single pendulum oscillating freely in the Morse potential is a one degree of freedom system whose phase portrait is shown in Figure 8.2(a). The shapes of the trajectories

in the full coupled-pendulum system, when projected onto the average variables as shown in Figure 8.2(b), resemble the trajectories of the single pendulum system except for the important difference that the single pendulum is an integrable system in which no *spontaneous* coherent flipping event can occur; the pendulum trajectory is either inside the homoclinic trajectory and never flips, or is outside the homoclinic trajectory and flips repeatedly. Therefore, the homoclinic trajectory in the single pendulum case defines a boundary of the conformation basin, and we say that a *spontaneous* flipping event in the full pendulum chain occurs when a trajectory in the average projection starts inside the region defined by the homoclinic trajectory, escapes it, and crosses into the other half plane by passing through the angle π radians.

These definitions afford us a precise manner in which to visualize transport from one conformation basin to the other. In the spirit of the FTLE-LCS method, the visualization method uses a measure of separation of nearby trajectories to find the boundaries of the regions in the flow that will undergo conformation change. Very simply, the visualization process involves four steps:

1. Construct a uniform grid of initial conditions in the phase space of the average angle in the reduced model.
2. Record whether each initial condition lies inside or outside a conformation basin as delineated by the homoclinic trajectory for the single pendulum.
3. (a) Advect each initial condition *forward* from time t_0 to time $t_0 + T$ and record whether the integrated position lies inside or outside a conformation basin.
 (b) Mark with blue each initial condition in the grid that has a neighboring initial condition that started in the same region but ended in a different region, and hence has separated from its neighbor (thus indicating points of forward separation).
4. (a) Advect each initial condition *backward* from time t_0 to time $t_0 - T$ and record whether the integrated position lies inside or outside a conformation basin.

- (b) Mark with red each initial condition in the grid that has a neighbor position that started in the same region but ended in a different region after backward integration (thus indicating points of backward separation).

These four steps can be repeated for different initial times t_0 producing an animation of the time-dependent manifolds of separation. This method is of course motivated by the FTLE-LCS approach for measuring separation between nearby trajectories. In this application, however, since we are specifically interested in transport from one pre-defined set to another, we define separation using the boundary of these sets rather than the FTLE.

Before an initial condition in the reduced phase space of the average angle can be advected, it must first be *lifted* to the full $2N$ -dimensional phase space. The initial condition in the reduced space determines $\hat{\theta}_0$ and $\dot{\hat{\theta}}_0$ exactly; however, the higher order components are free to be chosen arbitrarily and hence the lifting procedure is not unique. Presently, we compare the transport structures present when lower order modes are activated with the transport structures obtained when higher order modes are activated in order to make connections with the activation energies recorded in Figure 8.3. After lifting to the full space, the initial condition is then integrated using either the full system of equations or the reduced model.

Sample results obtained from this visualization procedure using the reduced order model are shown in Figure 8.7 and reveal manifolds in the flow that govern transport. For clarity of this initial presentation, the initial condition used in the visualization process was lifted to the full space by adding energy to only the first mode yielding a periodic time-dependence in the reduced model. Just as for the perturbed pendulum of Chapter 4, the intersection of the blue and red boundary surfaces define *lobes* that mediate transport in the flow. Indeed, the manifolds reveal that transport occurs via the now familiar mechanism of lobe dynamics that attends a perturbed homoclinic trajectory.

In Figure 8.8, snapshots of the trajectories inside two of the lobe regions taken for different values of t_0 have been suitably colored to indicate explicitly the action of lobe dynamics in this example. We observe that trajectories colored red are captured

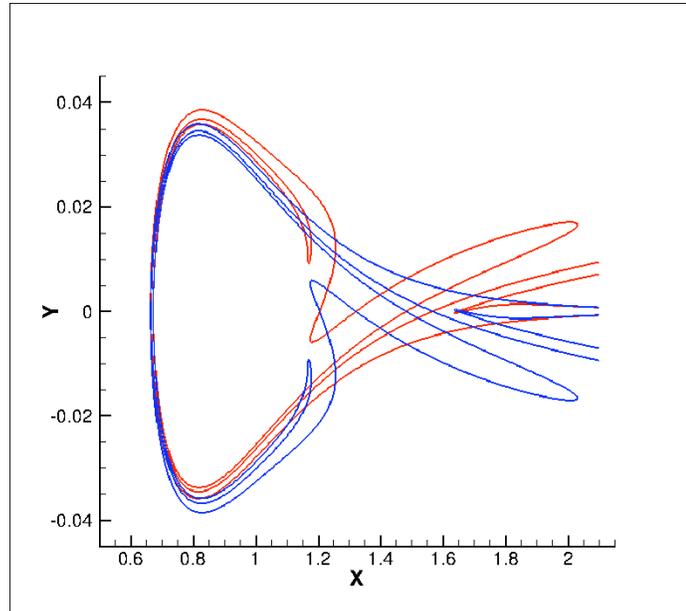


Figure 8.7: Surfaces that govern transport in the flow of the reduced model for the parameter values provided in 8.2.1 are visualized here using the method described in the text. Blue surfaces separate regions that will undergo conformation change from those that will not when trajectories are integrated *forward* in time. Conversely, red surfaces separate regions that will undergo conformation change from those that will not when trajectories are integrated *backward* in time. Together the surfaces reveal the familiar structure of a perturbed homoclinic tangle. This visualization is computed for the case when the perturbation is created by adding energy to the first mode only.

into the conformation basin, while those colored blue escape from the conformation basin and undergo flipping. Regions in which lobes overlap indicate trajectories that undergo complex itineraries, i.e., drifters in an overlapping region may first be captured into the conformation basin and then subsequently escape. Increasing the integration time T reveals more and more intersections and corresponding families of possible itineraries precisely as shown by Smale in the Horseshoe map description of the homoclinic tangle [Smale 1967].

Straightforward linear analysis of the full system reveals that the equilibrium point

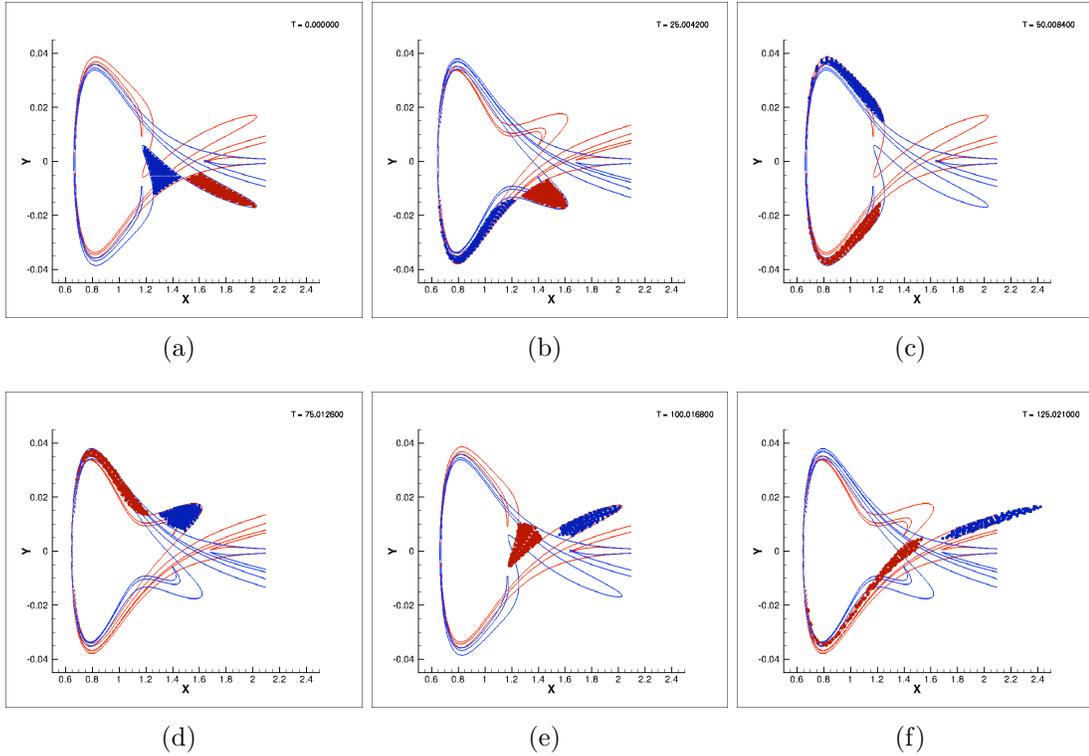


Figure 8.8: The action of lobe dynamics associated with the perturbed homoclinic trajectory in Figure 8.7 is easily observed by visualizing the time-dependent manifolds of greatest separation. The intersections of the manifolds define lobes that are either entrained or detrained from the conformation basin. In these snapshots, the trajectories visualized in red have recently flipped and are captured into the conformation basin. Conversely, the blue trajectories begin within the conformation basin and then escape and undergo flipping.

at $(\pi, 0)$ has $2N$ eigenvalues:

$$\lambda_{\pm}^p := \pm \sqrt{2 \cos\left(\frac{2\pi p}{N}\right) - 2 + \epsilon\mu} \quad p = 0, \dots, N-1, \quad (8.11)$$

where the real number μ is defined by

$$\mu := 2a(1 - e^{-a(2-x_0)})e^{-a(2-x_0)}. \quad (8.12)$$

The reasonable assumption that the Morse potential equilibrium distance x_0 is less than twice the pendulum length guarantees that μ is positive.

When epsilon is identically zero, one pair of eigenvalues lies at the origin, while all the remaining eigenvalues fall along the imaginary axis in complex conjugate pairs. For small non-zero epsilon, the zero eigenvalues are perturbed off the origin along the real axis yielding a symmetric pair of real eigenvalues. Precisely when

$$0 \leq \epsilon < \frac{2 - 2 \cos\left(\frac{2\pi}{N}\right)}{\mu} \quad (8.13)$$

the set of eigenvalues contains only one real pair and the equilibrium point is a rank 1 saddle, i.e., the linear stability of the equilibrium point is of the type saddle \times center₁ \times ... \times center _{$N-1$} . Transport in systems with rank 1 saddles has been well-studied with regard to low-dimensional applications in celestial mechanics and molecular dynamics. The results of Koon et al [Koon 2000] show that transport in systems with rank 1 saddles is governed by tubes defined by the invariant manifolds of the periodic orbit: trajectories inside the tube defined by the stable manifold, for example, will cross the saddle and be transported inside the tube from one region to another. For the parameters given for the coupled oscillator system in section 8.2.1, the inequality in condition (8.13) is satisfied whenever N is less than 24,000 and hence for $N = 200$ the equilibrium point associated with all the pendula in the flipped position at π radians is indeed a rank 1 saddle. The visualization procedure thus yields the interesting insight that transport via tubes in the full many degrees of freedom system corresponds to lobe dynamics in the projection onto the coarse average coordinates in the reduced model. This observation is currently being pursued further by the author and collaborators as a method for understanding transport in a wide range of molecular systems, such as the atomic cluster described in [Yanao 2007], for example.

We now proceed to use this method for visualizing transport structures to study the effect of the frequency content of perturbations on flipping. Figure 8.9 compares the transport structures obtained when the perturbing energy is initially placed in a range of modes and the resulting initial conditions are then advected using both the reduced and full models. In each of these plots, the amount of energy placed into the perturbing mode is 0.75 energy units – an amount equal in size to the energy required

by the zipper presented in section 8.2.4 to cause a trajectory at the equilibrium point to flip. Blue curves represent the boundary of the flipping region and are determined by integrating the initial conditions forward in time. Recall that this blue curve is not a trajectory, but rather a surface that separates trajectories that will flip from those that will not. In particular, the region enclosed by the blue curve contains initial conditions that will not flip. The gray curve is the homoclinic trajectory for the single pendulum in the Morse potential that defines the basin from which a flipping trajectory must escape. Also, for reference, the black “sawtooth” trajectory of the zipper is shown indicating the manner in which the zipper causes an initial condition at the equilibrium point (represented by a red dot) to flip. Since the equilibrium point in each of the plots falls inside the blue separatrix, none of the perturbations (except for the zipper) are able to induce flipping of trajectories starting at the equilibrium point at this particular energy level.

A study of the plots in Figure 8.9 yields several observations that in many ways serve as a summary of the chief results of this chapter. These observations are listed here:

- Qualitatively, *the reduced model captures the transport structures in the projected phase space remarkably well.* For the case when the initial energy is placed in mode 10, the reduced model does not provide the correct amplitude of the lobes, yet the frequency of the lobes and the overall proportions of the area delineated by the separatrix are very similar. For higher order modes all the relevant features are captured correctly and the differences between the reduced model and full model become visually indistinguishable. A single degree of freedom model obtained through averaging or truncation methods fails to capture flipping entirely. The time-dependence in the reduced model encodes the effect of the initial conditions of the full system so as to include the effect of higher order modes in inducing flipping.
- Perturbations consisting of higher order modes are less efficient at inducing flipping since the lobe structures are much smaller and the region inside the blue

separatrix that represents trajectories that do not undergo flipping is enlarged. For the case when the perturbing energy is placed in mode 40, for example, we see that no flipping is induced whatsoever. This observation serves as a visual representation of the fact that the conformation basin is impervious to noisy perturbations with high frequency content as has been previously observed in section 8.2.4. *The biomolecule chain is robust to noisy perturbations.*

- The zipper is an efficient and robust method for inducing conformation change that utilizes only localized perturbations. Certainly, perturbations with low frequency content are efficient at unzipping the chain; however, these perturbations have global structure. Hence, *conformation change in the biomolecule chain can be robustly induced via a low-energy, localized structured perturbation.*

In summary, then, we have introduced a strategy for deriving an approximate reduced equation for the evolution of a coarse variable in a high degree of freedom system in a way that carefully includes the influence of fine-scale dynamics, and have applied it to a simplified model for biomolecules. It has been shown that the basic mechanism of global conformation change can be understood using a single degree of freedom dynamical system that is acted on by an explicitly computable “subgrid” time-dependent forcing that represents the effect of the higher order modes on the main mode. We have also demonstrated that conformations are robust to large random perturbations, yet conformation change can be robustly induced by the application of a small, local structured perturbation. Visualization of transport in the coarse variables of the reduced model indicates that one channel through which conformation change occurs for a large class of perturbations is the process of lobe dynamics associated with the perturbed homoclinic trajectory of a rank 1 saddle.

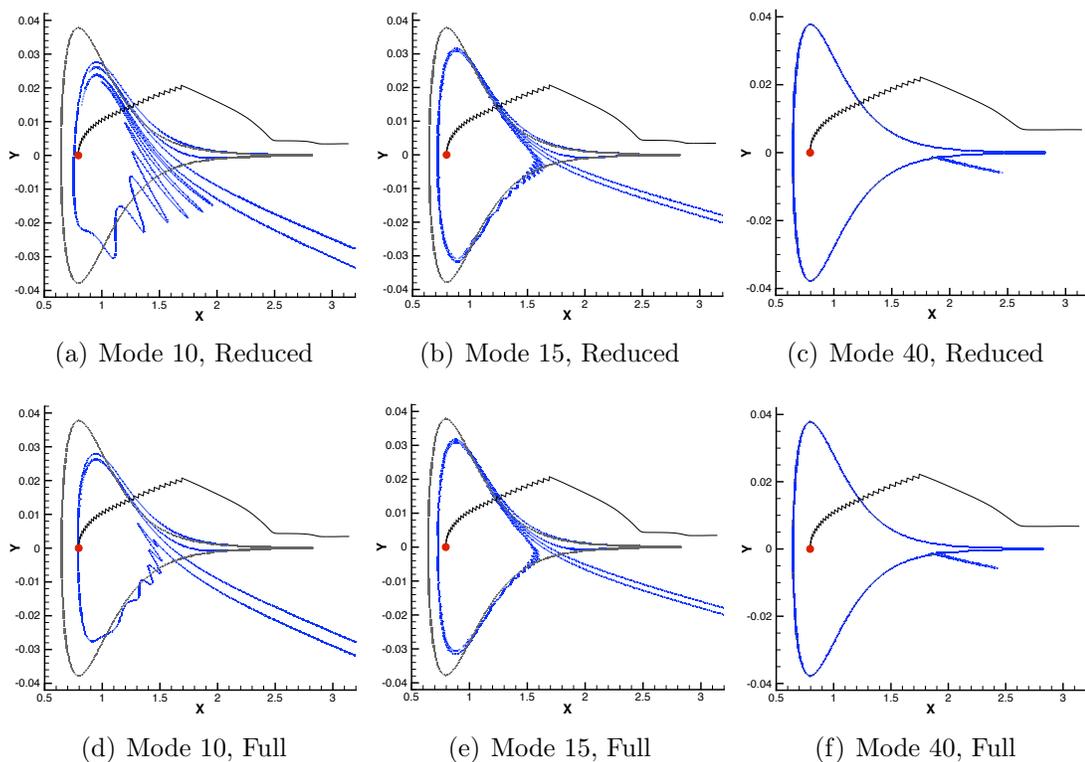


Figure 8.9: As described in more detail in the text, transport structures in the phase space of the average variables for the biomolecule are compared here for both the full and the reduced models. A review of the plots leads to the following observations: 1.) For perturbations in mode 10, the reduced model predicts incorrect amplitudes for the lobe structures, yet the overall geometry and topology of the flipping region is well captured. 2.) For perturbations in higher order modes, the transport structures in the reduced model are visually indistinguishable from those of the full model. 3.) Perturbations consisting of lower order modes are more efficient at inducing flipping motion. 4.) The zipped conformation of the chain is robust against high frequency perturbations. 5.) The zipper induces flipping efficiently while using only a localized perturbation.