Stochastic and Collective Properties of Nonlinear Oscillators

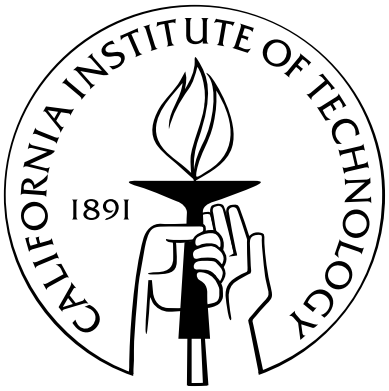
Thesis by

Oleg Kogan

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2009

(Defended December 16, 2008)

Dedicated to the memory of my mother.

Acknowledgements

I would like to thank my advisor and teacher Michael Cross for his patience and countless chats about physics through all these years. I always found these conversations to be very insightful and would like to hope that they had an influence on my own development as a scientist. I always aspire to attempt to see problems in a similarly intuitive and broad point of view. I also thank Mark Dykman for his "unofficial mentorship" and support throughout my many visits to MSU and for many also insightful, long phone conversations that we have had about physics. I also thank Michael Khasin for useful discussions. Part II of this work would not have been possible without the clever renormalization ideas of Gil Refael and his general expert-level understanding of renormalization group and critical phenomena. I feel lucky to have come across these three very strong physicists and look forward to potential future collaborations. The work described in part II would also have been much more difficult without the input from Jeffrey Rogers, especially in the numerical simulations. On the personal, but not any less important side, I thank my father for making it possible for me to be able to go to college and then to graduate school in the first place, while enduring and overcoming the hardships of assimilation in a new country that can only be understood by few. I also thank Sima for her kindness and support. I warmly acknowledge Josephine's presence through these years. Finally, I hold my officemates Janet Scheel and Tony Lee accountable for bringing nonscientific discussions into our office life.

Abstract

Two systems of nonlinear oscillators are considered: (a) a single periodically driven nonlinear oscillator interacting with a heat bath, which may operate in the regime of bistability or monostability, and (b) a one-dimensional chain of self-sustaining phase oscillators with nearest-neighbor interaction.

For a single oscillator we analyze the scaling crossovers in the thermal activation barrier between the two stable states. The rate of metastable decay in nonequilibrium systems is expected to display scaling behavior: the logarithm of the decay rate should scale as a power of the distance to a bifurcation point where the metastable state disappears. We establish the range where different scaling behavior is displayed and show how the crossover between different types of scaling occurs. We map the entire parameter range of bistability and find the regions where the scaling exponents are 1 or 3/2, depending on the damping. The exponent 3/2 is found to extend much further from the bifurcation then where it would be expected to hold as a result of an overdamped soft mode. Additionally, we uncover a new scaling behavior with exponent of ≈ 1.3 that extends beyond the close vicinity of the bifurcation point.

We also study the pattern of fluctuational trajectories in the monostable regime. For nonequilibrium systems, fluctuational and relaxational trajectories are not simply related by time-reversibility, as is the case in thermal equilibrium. One of the consequences of this is the onset of singularities in the pattern of fluctuational trajectories, where most probable paths to neighboring states are far away from each other. This also creates nonsmoothness in the probability distribution of the system in its phase space. We discover that the pattern of optimal paths in equilibrium systems is fragile with respect to the driving strength $\mathcal{F}$, and investigate how the singularities occur

as the system is driven away from equilibrium. As the strength of the driving $\mathcal{F}$ approaches zero, the cusp of the spiral caustic system recedes to larger radius R and the angle of the cusp also decreases. The dependence of R on $\mathcal{F}$ displays two scaling laws with crossovers, where the scaling exponents depend on the damping.

For the one-dimensional chain of nearest-neighbor coupled phase oscillators, we develop a renormalization group method to investigate synchronization clusters. We apply it numerically to Lorentzian distributions of intrinsic frequencies and couplings and investigate the statistics of the resultant cluster sizes and frequencies. We find that the distributions of sizes of frequency clusters are exponential, with a characteristic length. The dependence of this length upon parameters of these Lorentzian distributions develops an asymptotic power law with an exponent of 0.48 ± 0.02. The findings obtained with the renormalization group are compared with numerical simulations of the equations of motion of the chain, with an excellent agreement in all the aforementioned quantities.

Contents

List of Figures

Chapter 1

Overview

Equilibrium statistical mechanics serves as an indisputably powerful and generic tool that paved way for much progress in our understanding of collective phenomena in nature. This success lies in contrast to what is known outside the realm of equilibrium Hamiltonian systems. This includes driven Hamiltonian systems, or systems which are not based on a Hamiltonian at all. Although a great variety of techniques for nonequilibrium systems have been developed [1, 2], a unifying theory does not yet exist. While a big fraction of frontier research in modern physics lies in the quantum regime, understanding the laws that govern nonequilibrium systems, even the classical ones, is also fundamentally significant. At present, the research in this direction lies in the information-gathering stage [3]. A particularly important task is to find generic aspects of such systems. A wide range of problems, from the motion of vortex arrays in superconductors and condensates to the human brain, are examined with the hope of identifying unifying concepts which may eventually pave the way to a more generic understanding of nonequilibrium phenomena.

Nonlinear oscillators model a rich set of natural or man-made systems which can be used to probe fundamental issues in nonequilibrium statistical mechanics. They are also of experimental or practical value. For example, the advent of nanotechnology has brought us to the exciting length and time scales where nonlinearity of nanooscillating devices becomes important [4, 5] (see also appendix A), concomittantly with the rising importance of thermal noise [4] (and references therein). It has been demonstrated that nanoelectromechanical systems (NEMS), when built to operate in the classical

regime are accurately modelled as nonlinear oscillators (see our recent work [6] for instance). Additionally, if the current trend in miniaturization continues, it is very likely that quantum effects, such as tunneling, discretization of energy levels, quantum uncertainties, and entanglement are going to become apparent in the foreseeable future (see for example [7] - [11] for discussion of some of these issues; see also [12] and references therein).

My research at Caltech revolved mostly around two very different systems of nonequilibrium nonlinear oscillators. The first system appears as perhaps the simplest nontrivial example: a Duffing oscillator driven by white noise, which models interaction with the thermal bath, and also by a sinusoidal driving term, which ensures that the oscillator is not allowed to settle into thermal equilibrium. It is the interplay of nonlinearity and nonequilibrium that gives rise to the surprisingly rich effects and even allows us to explore certain fundamental questions in Nonequilibrium Statistical Mechanics.

The nonlinearity is responsible for bistability for some range of parameters: coexistence of two locally attracting steady-states which differ by amplitude and phase (with respect to the driving phase). Due to the noise, there is a probability of escape from one attracting fixed point[1] to another. However, because the oscillator is driven, this problem differs from the more straightforward case of thermal activation in a bistable potential well. In our system, the bistability is not a consequence of a two-well nature of the (time-dependent) potential. This has interesting implications not only on the switching rate between the basins, but also on the probability distribution of the oscillator in an appropriate phase space as well.

The problem of noise-induced escape from a metastable state has been studied in various contexts, from nucleation to chemical reactions to switching in nanomagnets. The escape rate often has the Arrhenius form $W \propto \exp(-R/D)$, where D is the noise intensity (temperature, in the case of thermal noise). Of central interest both for theory and experiment is the activation energy R. Starting with the Kramers paper

[1] In this work we will not deal with complicated attracting structures, such as strange attractors, so the occasional use of the word "attractor" is to be synonymous with attracting fixed point.

[13] much work has been put into calculating R and the prefactor in the escape rate for various physical systems. From this point of view, it is particularly important to reveal generic features of the activation energy R, such as scaling behavior with the system parameters. The onset of the scaling behavior was noticed first for systems close to thermal equilibrium [14, 15]. Scaling occurs also for systems far from thermal equilibrium, which generally do not have detailed balance [16, 17, 18, 19, 20]. A nonequilibrium system may display several types of scaling behaviors and crossovers between different scaling regimes with varying system parameters. The importance of studying scaling behaviors and scaling crossovers in a nonequilibrium system comes from the idea that scaling exponents may be generic and apply to a wide range of systems [21]. Chapter 3 is devoted specifically to the crossovers between different scaling regimes in the nonlinear oscillator.

On the experimental front, several types of scaling behavior of the switching activation energy have been observed in recent years for such different experimental systems as modulated nanoelectromechanical systems (NEMS) [4, 22, 23], and microelectromechanical (MEMS) oscillators [24] (various sources of noise in the nanomechanical systems have been thoroughly described in [25]), Josephson junctions [26], optically trapped atoms [27], and superconducting resonators [28]. All of these systems are modelled as Duffing oscillators. Not only is this scaling interesting on its own, but it also underlies various applications, an important example being quantum measurements [26].

The lack of detailed balance is fundamentally important here: a nonequilibrium system obeying detailed balance is in many respects similar to equilibrium system. One remarkable fact about a nonequilibrium system obeying detailed balance has been proven by Graham and Haken [29], who showed that the probability density function (PDF) over the phase space coordinates $\vec{r}$ can be expressed in the Boltzmann-like form $\exp[-\phi(\vec{r})/D]$, with a globally smooth effective potential $\phi(\vec{r})$. The other, closely-related manifestation of detailed balance is that the most probable fluctuational trajectory to a given state is related to the fluctuation-free (relaxational) trajectory back to the stable state by time reversal.

The detailed balance may be fragile with respect to the driving, as indeed happens in the case of the sinusoidally driven Duffing oscillator. We show that the detailed balance is broken at any nonzero value of the driving strength. The question remains, however, whether the smoothness of $\phi(r)$ is fragile? Chapter 4 is devoted to this question and to detailed discussion of detailed balance in our particular system. We found that in the limit of vanishing noise strength, the tails of $\phi(\vec{r})$ become piecewise differentiable at an infinitesimal driving strength, and we have found the specific mechanism by which this happens. It turns out that nonsmooth features of $\phi(\vec{r})$ arrive from infinity and move in closer to the attractor of the oscillator as the strength of the driving is increased. The distance of these singularities from the attractor versus the driving strength displays power laws with crossovers. Again, this is closely-related to the pattern of fluctuational trajectories: the most probable paths to *neighboring* states are *far* away from each other, and are not simply-related to the relaxational paths by time-reversal.

The $\phi(\vec{r})$ is the action of the auxiliary problem in classical mechanics that comes out of the WKB treatment of the Fokker-Planck equation for the PDF. Nondifferentiability of $\phi(\vec{r})$ occurs as a result of multivaluedness of the action [30, 31]: only one of the branches must be chosen in the limit $D \to 0$. Such discontinuities are related to the objects called caustics - the envelopes of trajectories. Here they are envelopes of most probable trajectories from an attracting fixed point to a point $\vec{r}$ (in analogy with Quantum Mechanics they may be called "classical trajectories"). Chapter 4 will explore the development of nonsmooth features of the PDF as the driving strength is increased from zero. The formation of regions where the action is multivalued, and the resultant rich phenomenology exhibited by ϕ through caustics and switching lines is ultimately related to the nonintegrability of the classical mechanics that comes out of the WKB treatment. Understanding the conditions required to maintain integrability in the absence of detailed balance remains at frontier of mathematical physics [32]. This problem is of fundamental relevance also because the aforementioned mechanism of development of nonsmooth effective potentials is believed to be generic in nonequilibrium systems described by finite number of degrees of freedom [30, 31].

In summary, even very simple systems lacking detailed balance display a rich phenomenology, and are not well understood; the conditions under which they retain properties of systems with detailed balance, and hence are qualitatively similar to equilibrium systems is one appealing line of research. It would also be very interesting to consider spatially extended systems and to explore noise-induced transitions as a mechanism for pattern selection.

While part I of this dissertation is devoted to nonequilibrium effects in a single nonlinear oscillator, part II of the thesis is devoted to nonequilibrium collective effects in a 1-dimensional chain of nonlinear oscillators. One of the most exciting aspects of modern physics is the investigation of collective phenomena in many-body problems. Spontaneous synchronization in networks of interacting nonlinear oscillators with different, often randomly distributed frequencies is one of the best understood collective phenomena in nonequilibrium systems [33]. If the interaction between the oscillators is weak, then they will evolve in an uncoordinated fashion. In contrast, with proper connectivity and coupling strength, the population can entrain to the common frequency, a process that can be reflected by an increase of an appropriately defined order parameter. This collective behavior has been observed in diverse systems including neural networks [34, 35, 36], Josephson junctions [37, 38, 39], lasers [40], electronic circuits [41], and a range of biological systems, including sleep cycles [42], animal gaits [43], and biological rhythms in general [44]. Synchronization, as one of the simplest examples of collective phenomena in nonequilibrium systems, has also been the focus of theoretical work [45].

Useful insights into the phenomenon of synchronization are obtained by analyzing the idealized case of all-to-all coupling in the limit of a large number of oscillators. In one of the most common models each oscillator is described by a single variable, phase, that advances at a constant rate in the absence of coupling. This model is exactly soluble using a mean-field approach, and it provides a framework for understanding synchorinzation [45, 46, 47]. In this model, the synchronized phase in which a finite fraction of oscillators evolve at a common frequency appears above a critical coupling

strength. Many features of this phase, such as the growth of the fraction of locked oscillators as a function of the coupling strength, can be calculated.

In many physical systems finite-range interactions depict a more realistic picture. The simplest of these interactions are nearest-neighbor couplings [48, 49]. In general, finite-range coupled systems are much harder to analyze, and, as a result, there are few theoretical tools that produce a quantitative description of the collective motion. Much of our understanding of systems of coupled oscillators with short-range interaction rests on numerical solutions of the governing equations. Yet, such systems are especially important from the fundamental point of view, for it is an amazing fact of nature that short-range coupling can give rise to collective phenomena on a much longer, even macroscopic length scale. Understanding of the mechanism of this process for equilibrium systems is the great achievement of the theory of critical phenomena [50] and deserves to be better understood for nonequilibrium systems as well, particularly in non-Hamiltonian models.

In part II we analyze a one-dimensional chain of nearest-neighbor coupled phase oscillators with quenched disorder in coupling strengths and intrinsic frequencies. Despite the relative simplicity of the model, it can exhibit complex behavior since randomness competes with a tendency to establish macroscopic order: Strogatz and Mirollo [51, 52] nicely showed that a nearest-neighbor chain cannot exhibit extensive synchronization. Nevertheless, the chain develops collective structures, albeit at finite length scales. These structures are clusters of common average frequencies, although the instantaneous dynamics of each oscillator in a given cluster can be quite complicated. The size of these clusters grows with increasing coupling strength.

We propose a method alternative to the direct numerical simulation to attack the synchronization problem in 1 dimension. The method we develop is a real-space renormalization group (RG) designed to work well on systems with strong randomness, where both the intrinsic oscillation frequencies and coupling constants are taken from distributions with long tails. The specific example we investigate contains Lorentzian distributions. We apply our RG technique to the problem of frequency clusters and make accurate predictions of various statistical properties of these clusters.

Part I

Aspects of Thermal Activation of Nonlinear Oscillators out of Equilibrium

Chapter 2

Driven Nonlinear Oscillator Interacting with a Heat Bath

This is a background chapter. The concepts developed here will be useful throughout the subsequent chapters. We will be concerned with a single driven Duffing oscillator interacting with the heat bath. The starting point for us will be the Langevin equation description,

$$\ddot{q} + 2\Gamma\dot{q} + \omega_0^2 q + \gamma q^3 = A\cos(\omega_F t) + f(t), \tag{2.1}$$

which is not the most fundamental, but one in which the complicated interaction with the bath has been reduced to a single additive white noise term $f(t)$

$$\langle f(t)f(t')\rangle = 2B\delta(t-t'). \tag{2.2}$$

The "passive" interaction with the environment is modelled here by a simple damping and noise term; the steps that lead to this simplified description form is a long and separate research topic by itself. This reduction has been described for a nonlinear oscillator in [53] for example. The rest of the terms represent Newton's second law for an oscillator in a nonlinear potential and driven via the sinusoidal driving. A wide variety of systems can be modelled as such. The specific case of a doublyclamped beam is worked out in Appendix A, where the steps leading from the PDE for the dynamics of the beam to the Duffing-like equation are shown.

Some of the nontrivial questions that can be asked are: (1) what is the effect of the

noise and what is the probability density function (PDF) $\rho(\vec{r})$ with the presence of the noise term[1] and (2) what are the properties of this nonequilibrium $\rho(\vec{r})$ and how do they differ from the equilibrium $\rho(\vec{r})$? The body of physics literature on this topic is too vast to cite at once; different sources will be cited throughout the discussion. At present, there remain unanswered questions, some of which will be addressed in part I of this thesis. From the mathematical point of view, the classic work describing the effect of noise on dynamical systems is a text by Wentzel and Freidlin, [54]. Another popular text on stochastic methods in physical sciences is [55].

2.1 From Langevin Equation to Lagrangian Manifolds

To address the big questions posed above we will reduce the problem from that of a noisy ODE to a calculation of a certain action of a certain "auxiliary" Hamiltonian. The steps that lead to this auxiliary description are not new, but for clarity and convenience I choose to briefly go through the steps that lead to this description. The sources related to various steps would be cited throughout. Readers wishing to skip over the details, are invited to go directly to the "summary" paragraph with Eqs. (2.36 - 2.40) and read until the end of Sec. 2.1.1, and then skip further until section 2.1.3.

2.1.1 Slow Envelope Dynamics

The aim here is to describe the behavior of the Duffing oscillator in a particular regime when it is weakly driven close to the resonance. In such a regime the oscillator will be nonchaotic, and its dynamic is given by simple harmonic oscillations modulated by the slowly varying modulation. For a more exhaustive survey of the Duffing oscillator, and the other regions of its parameter space one is invited to con-

[1] Given a point in phase space at $\vec{r}$, the quantity $\rho(\vec{r})dV$ is the probability to find a system in the small neighborhood of $\vec{r}$ of volume dV. In the ensemble picture, this gives the fraction of the elements of the ensemble that are found in this small region of phase space.

sult sources such as [56, 57, 58] and references therein. Thus, our first step is to derive the amplitude equations describing the slow modulation of simple harmonic motion due to the influence of the driving, damping and the cubic term in the Duffing equation. This will be done in the special case when the oscillator is driven close to linear resonance, i.e., $\omega_F \approx \omega_0$. The derivation of the amplitude equations of a nonlinear oscillator is described in many sources, all using somewhat different perturbation techniques and somewhat different notations; for example Nayfeh [58, 59] does this with the method of multiple scales and a classic text by Bogoliubov and Mitropolsky [60] employs their own elegant set of perturbation methods (the Mechanics volume of Landau and Lifshitz series [62] contains a typically terse, but a physically insightful discussion of nonlinear oscillations). For completeness I chose to go through the details here employing the averaging perturbation method (this and other perturbation techniques for dynamical systems can be found in [61] for example). The outcome of this procedure appears at the top of page 16.

It proves useful to start by nondimensionalizing Eq. (2.1). Let

$$x = \frac{\gamma^{1/2}}{\omega_0} q, \tag{2.3}$$

$$F = \frac{\gamma^{1/2}}{\omega_0^3} A, \tag{2.4}$$

$$\tilde{f} = \frac{\gamma^{1/2}}{\omega_0^3} f, \tag{2.5}$$

$$\tau = \omega_0 t, \tag{2.6}$$

$$\lambda = \frac{\Gamma}{\omega_0}, \tag{2.7}$$

$$\Omega = \frac{\omega_F}{\omega_0}. \tag{2.8}$$

Note that $q_0 \equiv \omega_0/\gamma^{1/2}$ is the displacement at which the linear and nonlinear terms are equal, so x is a displacement measured in units of this characteristic length x_0. Similarly, $F_0 \equiv \omega_0^3/\gamma^{1/2}$ is the force needed to be applied to extend q to q_0. In terms

of all these parameters, Eq. (2.1) reads

$$\frac{d^2x}{d\tau^2} + \frac{1}{\mathcal{Q}}\frac{dx}{d\tau} + x + x^3 = F\cos(\Omega\tau) + \tilde{f}(\tau). \tag{2.9}$$

The $\mathcal{Q} = \omega_0/2\Gamma$ is the quality factor which is a dimensionless measure of the damping. The correlation properties of the noise are modified as well:

$$\langle \tilde{f}(\tau)\tilde{f}(\tau')\rangle = 2B'\delta(\tau - \tau'), \tag{2.10}$$

where

$$B' = \frac{\gamma}{\omega_0^5}B. \tag{2.11}$$

We now specialize to the case of large $\mathcal{Q}$, which implies a natural smallness parameter $\epsilon \equiv 1/\mathcal{Q}$, in terms of which we can do perturbation theory. In this regime, the width of the resonance curve is much narrower then the resonance frequency. Let us specialize to the physics close to resonance. Set

$$\Omega = 1 + \epsilon\sigma, \tag{2.12}$$

where $\sigma \sim O(1)$. We know that near resonance a linear oscillator's response is $x \sim F/\epsilon$. Let $F = \epsilon^p$ and study what happens as p is varied. We have

$$F \sim \epsilon^p, \tag{2.13}$$

$$\ddot{x}, x \sim \epsilon^{p-1}, \tag{2.14}$$

$$x^3 \sim \epsilon^{3(p-1)}, \tag{2.15}$$

$$\epsilon\dot{x} \sim \epsilon^p. \tag{2.16}$$

Clearly we want $p > 1$ for the linear term to be bigger then the nonlinear one, otherwise the oscillator is in the strongly nonlinear regime and the linear relationship $x \sim F/\epsilon$ does not apply in the first place. We would like the force to be not too large so that the effect of nonlinearity is perturbative, but the linear terms are dominant.

When $p = 3/2$, the cubic, the damping, and the driving terms balance. For $p > 3/2$ the cubic term becomes much smaller then the damping and the forcing terms, which are both ϵ^p. So let $F \sim \epsilon^{3/2}$ (we set $F \equiv F'\epsilon^{3/2}$) and work out the consequences; it becomes clear that the consequences apply to the cases when $F' \ll O(1)$ which is equivalent to letting $p > 3/2$.

Under these conditions the dynamic is approximately that of the lightly damped harmonic oscillator. In general, it has a modulated envelope dynamics due to the beating of the natural and driving frequency, but this beating will decay away and some steady-state amplitude and phase (with respect to the driving) will be reached. Our objective is to calculate this envelope dynamics and the steady-state amplitude and phase to which it relaxes. This is where the nonlinearity plays a role; the linear solutions are known, and the nonlinearity will modify this envelope dynamics and the steady-state. To obtain an equation for the envelope dynamics we follow a common method in the theory of nonlinear oscillations. We make a transformation

$$\begin{aligned} x &= \epsilon^{1/2}\sqrt{4/3}\left(Y\sin\left(\Omega\tau\right) + X\cos\left(\Omega\tau\right)\right), && (2.17)\\ \dot{x} &= \Omega\epsilon^{1/2}\sqrt{4/3}\left(Y\cos\left(\Omega\tau\right) - X\sin\left(\Omega\tau\right)\right) \\ &= \epsilon^{1/2}(1+\epsilon\sigma)\sqrt{4/3}\left(Y\cos\left(\Omega\tau\right) - X\sin\left(\Omega\tau\right)\right). && (2.18) \end{aligned}$$

Note that if the oscillations are simple harmonic, the Y and X are constants. If the motion is not simple harmonic, Y and X are generally functions of time (τ) which represent precisely the modulated envelope dynamics. They can be found by substituting Eqs. (2.17 - 2.18) into the equation of motion, Eq. (2.9) and projecting out the sine or the cosine. We now do just this. First,

$$\begin{aligned} \sqrt{3/4}\ddot{x} &= \Omega\epsilon^{1/2}\left(\dot{Y}\cos\left(\Omega\tau\right) - \dot{X}\sin\left(\Omega\tau\right)\right) - \Omega^2\epsilon^{1/2}\left(Y\sin\left(\Omega\tau\right) + X\cos\left(\Omega\tau\right)\right) \\ &= \epsilon^{1/2}(1+\epsilon\sigma+...)\left(\dot{Y}\cos\left(\Omega\tau\right) - \dot{X}\sin\left(\Omega\tau\right)\right) \\ &- \epsilon^{1/2}(1+2\epsilon\sigma+...)\left(Y\sin\left(\Omega\tau\right) + X\cos\left(\Omega\tau\right)\right), \qquad (2.19) \end{aligned}$$

where we used Eq. (2.12). We now substitute Eqs. (2.17 - 2.19) into Eq. (2.9). Some

terms cancel.

$$\epsilon^{1/2}\left(\dot{Y}\cos(\Omega\tau)-\dot{X}\sin(\Omega\tau)\right)+\epsilon^{3/2}\sigma\left(\dot{Y}\cos(\Omega\tau)-\dot{X}\sin(\Omega\tau)\right)$$
$$-2\epsilon^{3/2}\sigma\left(Y\sin(\Omega\tau)+X\cos(\Omega\tau)\right)+\epsilon^{3/2}\left(Y\cos(\Omega\tau)-X\sin(\Omega\tau)\right)$$
$$+\epsilon^{3/2}(4/3)\left(Y^3\sin^3(\Omega\tau)+3Y^2X\sin^2(\Omega\tau)\cos(\Omega\tau)+3YX^2\sin(\Omega\tau)\cos^2(\Omega\tau)+X^3\cos^3(\Omega\tau)\right)$$
$$=\frac{3^{1/2}}{2}F'\epsilon^{3/2}\cos(\Omega\tau)+\frac{3^{1/2}}{2}f'\epsilon^{3/2} \tag{2.20}$$

We pulled out a factor of $\epsilon^{3/2}$ from the noise function: $\tilde{f}(\tau)=f'(\tau)\epsilon^{3/2}$. This is done without loss of generality: it will simply be reflected in the correlation property of f'. We also notice that the $O(\epsilon^{1/2})$ terms suggest that $\dot{X}$ and $\dot{Y}$ are independent of τ. The higher-order corrections more correctly suggest that these terms are simply small, so on the timescale τ, the functions X and Y vary slowly, reflecting the fact that the beating near the resonance is slow. It will be useful to define below a new timescale $T'=\epsilon\tau$, which is the characteristic timescale of these functions X and Y. We multiply the rest by $\sin(\Omega\tau)$ and integrate over $[\tau,\tau+2\pi/\Omega]$ or we multiply by $\cos(\Omega\tau)$ and also integrate over $[\tau,\tau+2\pi/\Omega]$: this is done while treating the functions X and Y as constants. We also use the following condition

$$\frac{dX}{d\tau}\cos(\Omega\tau)=-\frac{dY}{d\tau}\sin(\Omega\tau), \tag{2.21}$$

which came from differentiating Eqs. (2.17) and setting it equal to (2.18). After this procedure, the following equations for Y and X are obtained (we now switch to the slow timescale T')

$$\frac{dX}{dT'} = \frac{1}{2}\left[-(2\sigma)Y-X+Y(X^2+Y^2)+N_X\right], \tag{2.22}$$
$$\frac{dY}{dT'} = \frac{1}{2}\left[-Y+(2\sigma)X-X(X^2+Y^2)+\frac{3^{1/2}}{2}F'+N_Y\right], \tag{2.23}$$

where

$$N_X = \frac{\Omega}{\pi}\int_{\tau}^{\tau+2\pi/\Omega} f'(\tau')\sin(\Omega\tau')\,d\tau', \tag{2.24}$$

$$N_Y = \frac{\Omega}{\pi}\int_{\tau}^{\tau+2\pi/\Omega} f'(\tau')\cos(\Omega\tau')\,d\tau'. \tag{2.25}$$

We finally make one more round of variable changes. Let $\eta = 1/(2\sigma)$, $T = T'/(2\eta)$, $Y = P/\eta^{1/2}$, $X = Q/\eta^{1/2}$, $\mathcal{F} = \frac{3^{1/2}}{2}F'\eta^{3/2}$, $\mathcal{N}_i = \frac{3^{1/2}}{2}N_i\eta^{3/2}$ and obtain

$$\frac{dQ}{dT} = -P - \eta Q + P(Q^2+P^2) + \mathcal{N}_Q, \tag{2.26}$$

$$\frac{dP}{dT} = Q - \eta P - Q(Q^2+P^2) + \mathcal{F} + \mathcal{N}_P. \tag{2.27}$$

Without noise these would be called the "envelope equations," or the "amplitude equations" with which we will work throughout the entire part I of this thesis. While arriving at them, we went through several rounds of redefining variables, which are summarized for convenience of the reader in Table 2.1. We track down all of those renormalizations and express q and $\dot{q}$ in terms of P and Q and the parameters of the original Duffing model. The result is

$$q = \frac{2^{3/2}}{3^{1/2}}\sqrt{\frac{\omega_0}{\gamma}}\sqrt{|\omega_F-\omega_0|}\left(P\sin(\omega_F t) + Q\cos(\omega_F t)\right), \tag{2.28}$$

$$\frac{dq}{dt} = \frac{2^{3/2}}{3^{1/2}}\omega_F\sqrt{\frac{\omega_0}{\gamma}}\sqrt{|\omega_F-\omega_0|}\left(P\cos(\omega_F t) - Q\sin(\omega_F t)\right), \tag{2.29}$$

and the two parameters in the amplitude equations are

$$\eta = \frac{\Gamma}{\omega_F-\omega_0}, \tag{2.30}$$

$$\mathcal{F} = A\sqrt{\frac{3\gamma}{32\omega_0^3|\omega_F-\omega_0|}}. \tag{2.31}$$

We can also track down all the changes to the noise terms to relate the correlation functions of the original noise to the correlation functions of the noise in the amplitude equations. With details, these calculations are somewhat tedious, so they have been

relegated to Appendix B. The results can be summarized as follows. Given the convention

$$\langle \mathcal{N}_i(T_1)\mathcal{N}_j(T_2)\rangle = 2\mathbf{D}_{i,j}(T_1,T_2), \tag{2.32}$$

where $\mathbf{D}$ is the "diffusion matrix," we have

$$\begin{aligned} \mathbf{D}_{i,j}(T_1,T_2) &= \left(\frac{3\gamma\eta^2}{16\omega_0^3\Gamma^2}B\right)\delta_{i,j}\delta(T_2-T_1) && (2.33)\\ &= D\delta_{i,j}\delta(T_2-T_1), && (2.34)\end{aligned}$$

where

$$D \equiv \frac{3\gamma}{16\omega_0^3(\omega_F-\omega_0)^2}B, \tag{2.35}$$

expressed in terms of the parameters of the natural oscillator defined in the beginning of this section.

Table 2.1: Summary of variable changes

Variable	Change 1	Change 2	Change 3
Position, q	$x = \frac{\gamma^{1/2}}{\omega_0}q$	$X(x,\dot{x})$ $Y(x,\dot{x})$ Eqs. (2.17)-(2.18)	$Q = \eta^{1/2}X, P = \eta^{1/2}Y$
Driving strength, A	$F = \frac{\gamma^{1/2}}{\omega_0^3}A$	$F' = \frac{F}{\epsilon^{3/2}}$	$\mathcal{F} = \frac{3^{1/2}}{2}F'\eta^{3/2}$
Noise, f	$\tilde{f} = \frac{\gamma^{1/2}}{\omega_0^3}f$	$f' = \frac{\tilde{f}}{\epsilon^{3/2}}$	$\mathcal{N}_Q = \frac{3^{1/2}\Omega\eta^{3/2}}{2\pi}\int_\tau^{\tau+2\pi/\Omega} f'(\tau')\sin(\Omega\tau')\,d\tau'$ $\mathcal{N}_P = \frac{3^{1/2}\Omega\eta^{3/2}}{2\pi}\int_\tau^{\tau+2\pi/\Omega} f'(\tau')\cos(\Omega\tau')\,d\tau'$
Time, t	$\tau = \omega_0 t$	$T' = \epsilon\tau$	$T = \frac{T'}{2\eta}$
Damping, Γ	$\epsilon \equiv \frac{1}{Q} = \frac{2\Gamma}{\omega_0}$	-	$\eta = \frac{1}{2\sigma}$
Driving frequency, ω_F	$\Omega = \frac{\omega_F}{\omega_0}$	$\sigma = \frac{(\Omega-1)}{\epsilon}$	periodicity is removed

To sum up, we will now work with the following noisy dynamical system:

$$\dot{Q} = K_Q(Q,P) + \mathcal{N}_Q(T), \tag{2.36}$$

$$\dot{P} = K_P(Q,P) + \mathcal{N}_P(T), \tag{2.37}$$

where

$$K_Q(Q,P) = \frac{\partial g}{\partial P} - \eta Q, \tag{2.38}$$

$$K_P(Q,P) = -\frac{\partial g}{\partial Q} - \eta P. \tag{2.39}$$

And $g(Q,P)$ is a Hamiltonian given by

$$g = \frac{1}{4}(Q^2 + P^2 - 1)^2 - \mathcal{F}Q, \tag{2.40}$$

and η is an effective damping which tends to arrest the amplitude to zero. The $\mathcal{N}_i$s are stochastic terms which will be modelled as white noise with properties defined in Eqs. (2.2), (2.32 - 2.35).

In the absence of noise, this dynamical system may have one, two or three fixed points, depending on the damping η or driving strength $\mathcal{F}$. The amplitude versus $\mathcal{F}$ at fixed η are given by curves such as in Fig. 2.1 that exhibit multivaluedness for $\eta < \eta_c$. The upper and lower branches correspond to amplitudes of attracting fixed points (or attractors), while the middle branch is the amplitude of a fixed point with one stable and one unstable eigenvector (a saddle). Similar curve exists for phases of fixed points versus $\mathcal{F}$. An initial condition (Q_0, P_0) will evolve into one of the attractors. In the (Q,P) space, there is a separatrix between the basins of attraction of each of the two attractors. This separatrix is formed from a set of initial conditions that flows into the saddle via its attractive eigendirection, as demonstrated in Fig. 2.2. For $\eta < \eta_c$, one of the attracting fixed points coalesces with the saddle via the saddle-node bifurcation at the endpoints of the trivalued region. These endpoints are denoted by $\mathcal{F}_B$ and there are two of these: the low-field $\mathcal{F}_B^l$ and the high-field

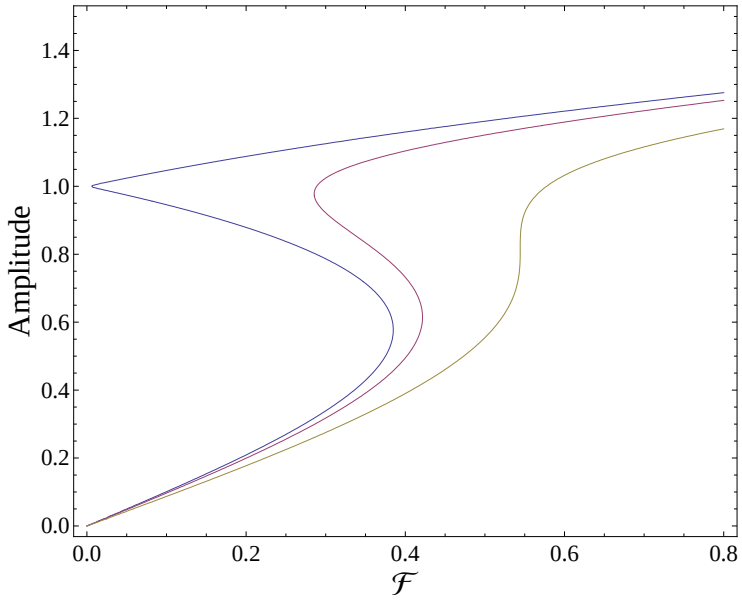


Figure 2.1: Amplitude of fixed points versus $\mathcal{F}$ for three different values of damping: $\eta = 0.01\eta_c$, $\eta = 0.5\eta_c$, and $\eta = \eta_c = \frac{1}{\sqrt{3}}$, left to right.

$\mathcal{F}_B^h$. As $\mathcal{F}$ is slowly varied,[2] the system will adiabatically follow one of the attracting fixed points until this fixed point disappears by colliding with an unstable fixed point in the saddle-node bifurcation. Upon further varying $\mathcal{F}$ the system will settle into the only other remaining fixed point. Note that reversing the change in $\mathcal{F}$ will result in the hysteretic behavior. Incidentally, there are interesting effects that arise when $\mathcal{F}$ is varied fast. I explored this recently in [63].

The entire η-$\mathcal{F}$ parameter space is represented in Fig. 2.3. The closed triangular-shaped domain is the region in which there are three solutions. If the parameter range is in the bistable regime the noise can overthrow the system from one basin to another. It is critically important to note that at any finite η, bistability will exist only at a nonzero $\mathcal{F}$. Therefore the problem of transition rates between two stable attractors is a nonequilibrium problem! Noise-induced transitions from low amplitude to high amplitude branch of 2.1 will be called "down $\rightarrow$ up transitions," and those from high amplitude to low amplitude branch, will be called "up $\rightarrow$ down transitions." In the

[2]Slow variation means that the rate at which the position of the quasi-fixed point changes as $\mathcal{F}$ is varied is much smaller then the rate of relaxation to the attractor determined by its eigenvalues at that instantaneous value of $\mathcal{F}$ and η.

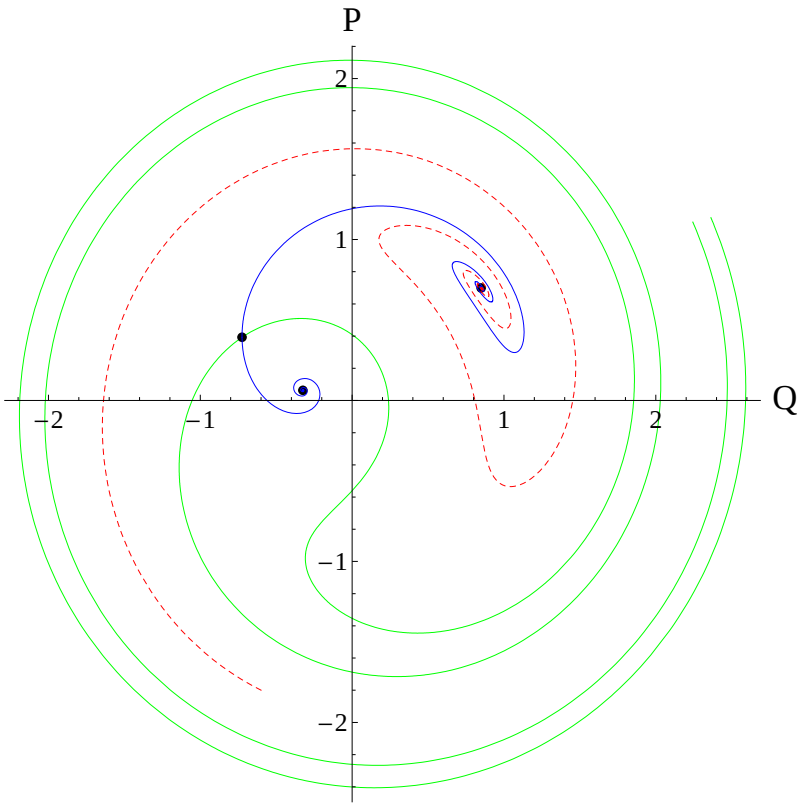


Figure 2.2: The dynamics in the P-Q space for a particular values of parameters $\eta = 0.3\sigma_c \approx 0.173$ and $\mathcal{F} = 0.3$. The fixed points are shown as black solid dots. The left-most one is the saddle, which lies on the separatrix between two basins. The other curves shown are trajectory from the saddle to each of the attractors (blue spiraling curves) and some other typical trajectory (dashed curve).

limit of very small noise, when D is the smallest parameter in the system, there is a characteristic time for these switching events. In this limit, the switchings are rare events because the system has to fight the much stronger tendency of an oscillator to return to one of the fixed points. The escape process is cumulative and can be thought of as a diffusion process in the deterministic drift field $\vec{K}$ in Eqs. (2.36)-(2.37). A mathematical description of this process is captured by the Fokker-Planck equation which we describe next. Intuitively, a system placed at one of the attractors at time t_0 will wonder around this attractor and occasionally find itself far enough away from it to reach and overstep the separatrix. Due to the strong deterministic current, the most likely subsequent scenario will be the quick relaxation of the system toward the

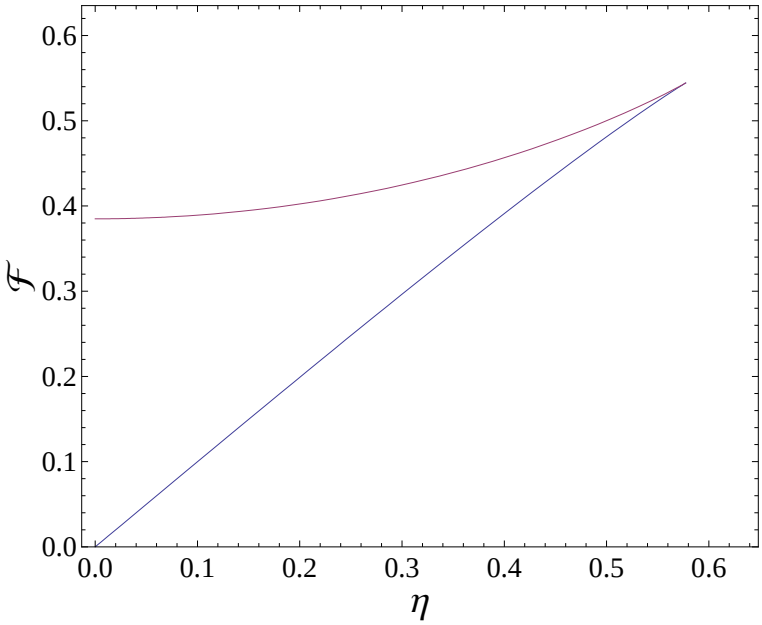


Figure 2.3: The region of bistability in the η-$\mathcal{F}$ parameter space is the closed region between the $\mathcal{F}_B^h$ (top) and $\mathcal{F}_B^l$ (bottom) curves. These curves meet at the cusp point at $\left(\eta_c = \frac{1}{\sqrt{3}}, \mathcal{F}_c = \sqrt{\frac{8}{27}}\right)$. The width of the hysteresis at zero η is $\sqrt{\frac{4}{27}}$. Outside of the bistability region there is always only one fixed point which is an attractor.

attractor on this new side of the separatrix, completing the switching event.

2.1.2 Fokker-Planck equation and the Path Integral solution

The probability density function (PDF) $\rho(Q, P)$ of reaching (Q, P) from some initial (Q_0, P_0) satisfies the Fokker-Planck equation (FPE), see [30] for example:

$$\frac{\partial \rho}{\partial t} = -\vec{\nabla} \cdot \left(\vec{K}\rho - D\vec{\nabla}\rho\right) \tag{2.41}$$

$$= -\vec{\nabla} \cdot \vec{j}. \tag{2.42}$$

This can be obtained by standard derivation methods (see [2] or [64] for example), but it has a simple intuitive interpretation of diffusion in a drift. The $\vec{j}$ is simply a probability current, which has a deterministic component and a diffusive component, much like ink dropped in a current will be carried away by the current, and it will also spread out diffusively. This FPE can be rewritten to look like a Liouville equation

with a scattering term on the right-hand side:

$$\frac{\partial \rho}{\partial T} - \{g, \rho\} = \vec{\nabla} \cdot \left(\eta \vec{r} \rho + D \vec{\nabla} \rho \right) \tag{2.43}$$

The solution can be written as a functional integral

$$\rho(\vec{r}, t | \vec{r}_0) = \int \mathcal{D}\mu \left[\vec{r} \right] \exp \left[-\frac{1}{D} \int_{\vec{r}_0 = \vec{r}(0)}^{\vec{r} = \vec{r}(t)} L \left(\vec{r}(\tau), \dot{\vec{r}}(\tau) \right) \, d\tau \right], \tag{2.44}$$

over all paths [30, 54, 65] where $\vec{r}$ represents a position in the (Q, P)-space, $\mathcal{D}\mu[\vec{\rho}]$ is the functional measure and the Lagrangian is

$$L = \frac{1}{4} \left[\left(\dot{Q} - K_Q \right)^2 + \left(\dot{P} - K_P \right)^2 \right] + O(D). \tag{2.45}$$

The $O(D)$, and an even smaller $O\left(D^{3/2}\right)$ terms will not be considered: we are currently interested in the dominant contribution in the limit of vanishing noise strength. The leading part of L that we consider has a very simple geometric interpretation. The probability density of a given trajectory can be computed by first breaking up this trajectory into small pieces, each piece corresponds to a constant *time interval* Δt. Consider the $k$the piece: locally the drift field $\vec{K}$ is constant. Suppose the probability of being found at the beginning of this kth piece at time t_k is known (i.e., the PDF is a delta function of unit strength). Over the course of the subsequent Δt, the probability density will spread out and it will also drift. We can then use an elementary solution to the diffusion problem to calculate the probability of being found at the end of the kth piece after time Δt, i.e.,

$$\rho_k \sim \exp \left[\frac{\left(\vec{r} \Delta t - \vec{K} \Delta t \right)^2}{4 D \Delta t} \right], \tag{2.46}$$

where $\vec{r}$ and $\vec{K}$ are evaluated at position of the beginning of kth piece, $\vec{r}_k$. Finally, the probability of reaching $\vec{r}$ from $\vec{r}_0$ is the product $\Pi_k \rho_k$: the product of exponentials gives an exponential of a sum. We also see that since the difference between $\vec{r}$ and

$\vec{K}$ is just the noise, the leading-order part of the action represents the integral over the noise realization which allows reaching $\vec{r}$ and minimizes $\int \left|\vec{\mathcal{N}}(T)\right|^2 dT$, which can be interpreted as the "noise energy," or the work the system needs to do against the stochastic force. Thus, the sum over trajectories can also be represented as the sum over noise realizations.

A rigorous and general derivation by Graham can be found in [66] (for a concise summary, see Section III of Graham's 1985 paper [30]), which also includes the omitted $O(D)$ and $O\left(D^{3/2}\right)$ terms, as well as a generalization to the case of a nonunit diffusion matrix; it also mentions prior works of Onsager and Machlup who first proposed a path integral solution to a PDF, albeit restricted to the Gaussian process. A complimentary set of sources are those by Dykman [67] and [68] as well as in Dykman and Smelyanskiy [69], all of which have a more physically intuitive discussions then those of Graham, but they gloss over some mathematical details which can be found in the works by Phythian [70] and Jouvet and Phythian [71]. The classics on path integrals by Feynman and Hibbs [72] and Schulman [73] also serve as useful references in setting up the path integral.

One can reexpress all paths as variations around a path that minimizes the exponent. We will call this the "optimal path[3]." Then

$$\begin{aligned} \rho(\vec{r},t|\vec{r}_0) &= \rho'(\vec{r},\vec{r}_0,t)\exp\left[-\frac{1}{D}min\int_{\vec{r}_0=\vec{r}(-t)}^{\vec{r}=\vec{r}(0)} L\left(\vec{r}(\tau),\dot{\vec{r}}(\tau)\right)\,d\tau\right] \\ &= \rho'(\vec{r},\vec{r}_0,t)\exp\left[-\frac{S(\vec{r})}{D}\right], \end{aligned} \tag{2.47}$$

where the prefactor ρ' comes from performing path integrals over variations around the optimal path. This prefactor is generally nonexponential: we can motivate it by approximating the functional integral by a finite number of regular integrals, each of which is over a variation and *in general* converges to some finite quantity. There are notable exceptions, particularly at the cusp of the pair of caustics (Chapter 4) as recently pointed out by Maier and Stein in [19]. We will only be concerned with the

[3]In analogy to the path integral formulation of Quantum Mechanics, this can be called the classical path.

exponential contribution, because it will dominate transition rates in the limit of weak noise. The problem of prefactors for continuous systems without detailed balance remains open. The formulation of the type in Eq. (2.47), where $S(\vec{r})$ is expressed as the sum over trajectories $\vec{r}(t)$ is known as the Wentzell-Freidlin formulation. A different formulation, where the functional integration is also performed over the noise trajectories and over a Lagrange multiplier that relates noise trajectories and trajectories $r(t)$ can be found in [67]-[72]. Clearly, the Wentzell-Freidlin functional can be obtained when noise and the lagrange multiplier trajectories are either integrated out or a saddle-point type method is used.

We are concerned with the distributions and escape rates in the limit of vanishing noise, $D \to 0$. Clearly, the escape rates W from one basin to the other approach ∞ in this limit.[4] In an ensemble of experiments where all systems are initially positioned at some $\vec{r}_0$, the distribution $\rho(\vec{r}, t|\vec{r}_0)$ quickly establishes a quasistationary form on the time scale $\tau_{rel} \ll t \ll W^{-1}$, where τ_{rel} is the relaxational time scale dictated by the noise-free dynamics. In the limit $D \to 0$ we can replace this time range by $\tau_{rel} \ll t < \infty$. Then, we can say that in the limit $D \to 0$ the quasistationary distribution $\rho(\vec{r})$ is given by $\lim_{t\to\infty} \rho(\vec{r}, t|\vec{r}_0)$ (note that $D = 0$ and the limit $D \to 0$ are not the same). This limit is indeed independent of $\vec{r}_0$ and hence represents the steady-state probability distribution [30]. Mathematically, $-D \ln [\lim_{t\to\infty} \rho(\vec{r}, t|\vec{r}_0)]$ is the action of a trajectory from $\vec{r}_0$ to $\vec{r}$ that takes an infinitely long time. With no finite time constraint to reaching $\vec{r}$, the most likely trajectory from $\vec{r}_0$ to $\vec{r}$ first follows a relaxational path to the attractor of $\vec{K}$ (i.e. the noise-free path governed by $\dot{\vec{r}} = \vec{K}$); this costs zero action, and takes an infinitely long time, so all the memory of $\vec{r}_0$ is lost, because all such trajectories to reach $\vec{r}$ will pass through the attracting fixed point. Notice also from Eq. (2.45) that $\dot{\vec{r}} = \vec{K}$ corresponds to zero-momentum

[4]Note that there are two contexts in which the word "escape" appears. (1) escape from an attractor to a point $\vec{r}$: this most likely happens via an optimal (or classical) trajectories and moreover, the energy of this trajectory is 0 in the case of stationary states. These may also be called fluctuational trajectories, as opposed to the relaxational ones. (2) escape from one attractor to another, as used in this sentence and in Sec. 2.1.4: this most likely happens through a *special* optimal trajectory which crosses the separatrix through the saddle (see below). Reaching a saddle normally involves a large deviation from the attractor: in our work the noise is assumed to be the smallest parameter in the theory, so this is always a rare event.

"auxiliary" dynamics of the Lagrangian L. The Hamiltonian associated with this Lagrangian L is

$$\mathcal{H} = \left(p_Q^2 + p_P^2\right) + p_Q K_Q + p_P K_P. \tag{2.48}$$

This $\mathcal{H}$ is often called the "auxiliary" Hamiltonian, to avoid confusion with g. The points $(Q_{FP}, P_{FP}, p_Q = 0, p_P = 0)$ are fixed points of the auxiliary dynamical system formed from $\mathcal{H}$. Therefore, the remaining part of the path, from the attractor to $\vec{r}$, also takes an infinite time because it escapes the fixed point of the auxiliary dynamics; however, unlike the first part of the path, it will cost a finite action.

In reality, the quantity $-D \ln\left[\lim_{t\to\infty} \rho(\vec{r}, t|\vec{r}_0)\right]$ is the *leading-order* term in the expansion of the action in powers of D and the resulting ρ represents stationary distribution. The corrections in powers of D would take into account the fact that W^{-1} is finite. In short, the leading-order exponential part (the stationary part) of the quasistationary distribution is given by

$$\rho(\vec{r}) \sim \exp\left[-\frac{1}{D} min \int_{\vec{r}_a(-\infty)}^{\vec{r}=\vec{r}(0)} L\left(\vec{r}(\tau), \dot{\vec{r}}(\tau)\right)\, d\tau\right]. \tag{2.49}$$

2.1.3 Auxiliary Classical Mechanics

To calculate $\rho(\vec{r})$ up to this accuracy in D, we need in principle to solve Euler-Lagrange equations to find a path that leads from the attractor to point $\vec{r} = (Q, P)$ and evaluate the action S along this path. The action $S(\vec{r})$ serves as a generalized potential. Note that there are infinitely many *optimal* (or classical) trajectories from an attractor to the point $\vec{r}$; they all differ by their initial momenta, but only those trajectories that have zero energy arrive to and emanate from fixed points (in 4D) and thus take an infinite time. In fact, another way to see that only zero energy (and hence infinite time) trajectories are related to stationary $\rho(Q, P)$ is to substitute $\rho(Q, P) = \rho'(Q, P) \exp\left(-S(Q, P)/D\right)$ into the stationary FPE [74], and obtain

$$\left(\frac{\partial S}{\partial Q}\right)^2 + \left(\frac{\partial S}{\partial P}\right)^2 + \frac{\partial S}{\partial Q} K_Q + \frac{\partial S}{\partial P} K_P = 0. \tag{2.50}$$

This is just a Hamilton-Jacobi equation, $\mathcal{H}(Q, P, \frac{\partial S}{\partial Q}, \frac{\partial S}{\partial P}) = 0$, and it describes dynamics on the $\mathcal{H} = 0$ manifold. Now, since $\vec{K}(Q_{FP}, P_{FP}) = 0$, the momenta at all the FPs must also be zero.

The Hamiltonian EOM that follow from $\mathcal{H}$ form an auxililary 4D dyanmical system with two momenta, p_Q and p_P and two spatial variables Q and P. This system is

$$\dot{Q} = 2p_Q - \eta Q + \frac{\partial g}{\partial P} = 2p_Q - \eta Q - P + (Q^2 + P^2), P \tag{2.51}$$

$$\dot{P} = 2p_P - \eta y - \frac{\partial g}{\partial Q} = 2p_P - \eta P + Q - (Q^2 + P^2)Q + \mathcal{F}, \tag{2.52}$$

$$\dot{p}_Q = \eta p_Q - p_Q \frac{\partial^2 g}{\partial Q \partial P} + p_P \frac{\partial^2 g}{\partial Q^2} = (\eta - 2QP)p_Q + (3Q^2 + P^2 - 1)p_P, \tag{2.53}$$

$$\dot{p}_P = \eta p_P - p_Q \frac{\partial^2 g}{\partial P^2} + p_P \frac{\partial^2 g}{\partial Q \partial P} = -(Q^2 + 3P^2 - 1)p_Q + (\eta + 2QP)p_P. \tag{2.54}$$

We notice immediately that the deterministic 2D dynamics takes place on the $\vec{p} = 0$ plane. In fact, $2\vec{p} = \frac{1}{2}(\dot{\vec{x}} - \vec{K}) =$ noise, so in the absence of noise, the only possible dynamics lie on the $\vec{p} = 0$ plane. Any fluctuational (as opposed to relaxational) dynamics will be lifted out of this plane. As already mentioned, the fixed points (FP) of this dynamics are $(Q_{FP}, P_{FP}, 0, 0)$ where (Q_{FP}, P_{FP}) are the FP of the original 2D dynamics. It can also be easily shown that eigenvalues of each of these auxiliary FP are $(\lambda_1, \lambda_2, -\lambda_1, -\lambda_2)$ where λ_1 and λ_2 are eigenvalues of the FP of the original dynamics. The $\vec{p} = 0$ plane forms the stable manifold of each of the fixed points. A trajectory that escapes from an attracting fixed point (now using the first meaning word escape, see footnote 4) will lie on a certain manifold (surface) in the auxiliary (Q, P, p_Q, p_P) space. In the vicinity of the attracting fixed point, this manifold is tangent to the plane spanned by two unstable eigenvectors of this FP. Hence, the manifold on which the fluctuational trajectories lie is conventionally called the unstable manifold. The stable and unstable manifolds are known as Lagrangian Manifolds [30, 74, 75, 76]. Calculation of trajectories that escape the attractor will involve placing initial conditions on this unstable manifold close to the attracting FP. This procedure is described in detail in the following Chapter. Some analytical aspects of the Classical Mechanics associated with $\mathcal{H}$ in chapter 4.

2.1.4 Escape Rates

To calculate the escape rate [65, 67] from one attracting fixed point to the other (here I am making use of the second meaning of the word escape), in principle, one must compute the line integral of $\vec{j}(l) \cdot \hat{n}(l)$ over the separatrix, where $\vec{j}$ is the probability current, $\hat{n}$ is the unit normal to the separatrix and l is the arc length along the separatrix. The $\vec{j}$ must come from a nonstationary distribution, otherwise the net flow from the basin would be zero in view of divergence theorem (see Eq. (2.42) with zero left hand side). For example, it may come from the slowly-decaying quasistationary distribution. We have seen that the leading order term in D of such a distribution is given by the stationary distribution in Eq. (2.49). Moreover, it may be seen from the saddle-point method [30], for example, that the line integral over the separatrix will pick up an exponential contribution given by the minimum value of the integrand along the separatrix. In fact, it was shown by Dykman and Kryvoglaz [65] that of all the classical trajectories emanating from the attractor and leading to the separatrix around its basin of attraction, the path that leads to the saddle point, if it lies on the separatrix (as it does in our system), will have the minimum action. We call this action S^*. Furthermore, in performing the line integral over the separatrix, one can factor out $\exp(-S^*/D)$ and perform the integration, which will give some number. In other words, up to the exponential factor, the escape rate W will simply be given by

$$W \sim \exp(-S^*/D), \tag{2.55}$$

which also happens to be the exponential factor in $\rho(\vec{r}_s)$. The preexponential factors (of either the distribution $\rho(\vec{r})$ or escape rate W) are not easy to obtain but they are also the least important in the limit of low D!

In summary, in the limit of weak noise, the rate of escape W from one attracting fixed point to another is given by $W \approx \exp(-R/D)$ where $R = S^*$ is the action along the trajectory which connects the attracting fixed point to the saddle and satisfies $\mathcal{H} = 0$.

2.2 Analytical Limits

Both limits discussed below are physically analogous to the two types of limits considered by Kramers [13]: an overdamped limit and a low-damping limit. One major difference is that our Hamiltonian $g(Q, P)$ does not have an explicit separation of kinetic and potential energies. The results in Chapter 3 will be compared with the analytical limits derived in this section.

2.2.1 Close to Bifurcations

At either bifurcation point (see Fig. 2.1) which we called $\mathcal{F}_B^l$ or $\mathcal{F}_B^h$, the saddle and one of the attracting nodes of $\vec{K}$ merge together at (Q_B, P_B) in a saddle-node bifurcation. As $\mathcal{F}$ approaches the bifurcating value, $\mathcal{F}_B$, a saddle and a node approach each other, the attractive eigenvector of the saddle and one of the attracting eigenvectors of the node align parallel to each other, and the corresponding attracting eigenvalues of both FP approach the same negative value. Also, the repulsive eigenvector of the saddle and the other attracting eigenvector of the node become equal and the corresponding eigenvalues both tend to zero. This process is depicted in the cartoon in Fig. 2.4.

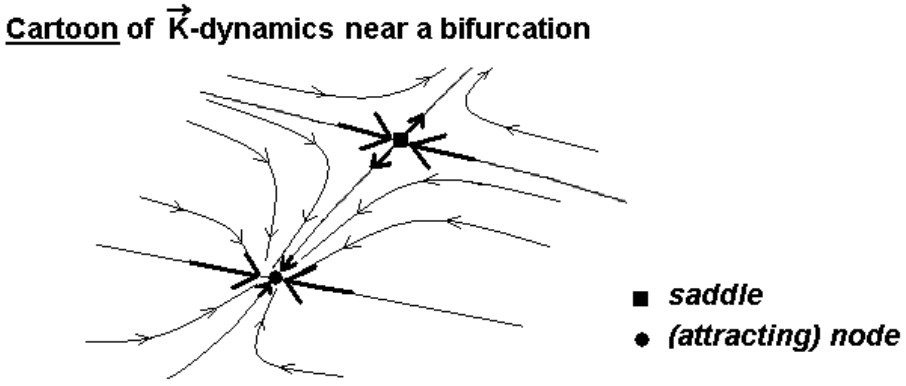


Figure 2.4: Cartoon of the $\vec{K}$-flow near a saddle-node bifurcation.

Thus, close to a bifurcation, there develops a soft mode: a narrow region connecting the saddle with an attractor along which the $\vec{K}$-motion is slow. Due to this slowing down, the soft mode forms a path of least resistance along which a large noise-

induced fluctuation away from the attractor is most likely to take place. To analyze this system we rewrite $\vec{K}$ in terms of parameter $\delta = |\mathcal{F} - \mathcal{F}_B|$ and variables ($f = Q - Q_B, s = P - P_B$); it turns out to be unnecessary to make a linear transformation to eigencoordinates of $\vec{K}$ at $(Q_B, P_B, \mathcal{F}_B)$. In terms of these new coordinates, $\vec{K}$ transforms to the following form:

$$\dot{f} = -2\eta(f - a_B s) + P_B f^2 + 2Q_B sf + 3P_B s^2 + f^2 s + s^3 + \mathcal{N}_Q \quad (2.56)$$

$$\dot{s} = \pm\delta - 3Q_B f^2 - 2P_B fs - Q_B s^2 - f^3 - s^2 f + \mathcal{N}_P \quad (2.57)$$

Here the + sign applies for the low field (high amplitude) bifurcation and

$$a_B = \eta^{-1}(1 \pm 2\sqrt{1 - 3\eta^2})/3. \quad (2.58)$$

Notice that the position of the fixed points sets a characteristic scaling of both s and f variables to be $\propto \delta^{1/2}$. From this we see that s is a slow variable while f is fast - on its timescale, s appears approximately frozen. To lowest order in δ, we treat s as completely frozen, in which case f relaxes to $a_B s$ plus corrections of order δ. In such adiabatic limit, the dynamics of s is then given by

$$\dot{s} \approx \pm\delta - b(\eta)s^2 + \mathcal{N}_P, \quad (2.59)$$

where

$$b(\eta) = 3Q_B a_B^2 + 2P_B a_B + Q_B, \quad (2.60)$$

($b > 0$ for the low-field bifurcation and $b < 0$ for the high-field bifurcation). Notice that the characteristic timescale for the variable s, characterized by the eigenvalues around the fixed point, scales as $\delta^{1/2}$, while the relaxational timescale for the variable f scales as 2η which is not assumed to be particularly small, so indeed s is a slow variable. The equation describing the dynamics of the slow variable, Eq. (2.59), can

be seen as an equation for an overdamped particle in a cubic potential U given by

$$U = \mp\delta s + \frac{b}{3}s^3, \tag{2.61}$$

while the distance between the saddle and the attractor is $\sqrt{\delta/|b|}$; thus the potential barrier for the slow coordinate is

$$\Delta U = \frac{4\delta^{3/2}}{3|b|^{1/2}(\eta)}. \tag{2.62}$$

This is the well-known $\xi = 3/2$ activation scaling. A 1-dimensional Fokker-Planck equation for s can be written, a stationary solution for $\rho(s, \tilde{\eta}, \eta)$ obtained, and the escape rate calculated in the standard way. This quantity will have the form $W = W'(\eta)\exp(-\Delta U/D)$. This type of analysis for a multidimensional nonlinear dynamical system was first performed by Dykman and Krivoglaz in 1980 [16]. These authors also considered the escape problem near the double-bifurcation point in which the top and bottom branches in Fig. 2.3 meet. Such an adiabatic approach is valid only sufficiently close to the bifurcation point. When $\eta \ll 1$, "sufficiently close" means that $\delta \ll \eta^3$. In this regime,

$$R \approx \begin{cases} \frac{4\sqrt{2}}{3}\eta^{1/2}\delta^{3/2} & \text{low field bifurcation,} \\ \frac{4}{3^{1/4}}\eta\delta^{3/2} & \text{high field bifurcation.} \end{cases} \tag{2.63}$$

A more systematic treatment of this regime will allow us to extract corrections to the adiabatic limit and derive the $\delta \ll \eta^3$ criterion will be considered in Section 3.2.1.2.

2.2.2 Low Effective Damping Limit

When the effective friction η is low, the dynamics of $\vec{K}$ is approximately Hamiltonian: the timescale for the decay of energy is much larger then the timescale to make one cycle on the contour of approximately fixed energy. This separation of timescales breaks down very close to a bifurcation point (refer back to Figs. 2.1 and 2.3), so as long as $\eta \neq 0$, there will always be a small fraction of the hysteresis displaying

a $\xi = 3/2$ scaling. Aside from this very narrow region, the separation of timescales allows one to turn a 2-variable FPE into a 1-variable FPE for the diffusion of *energy*, by averaging out the fast part of the dynamics. The noise strength D is proportional to η^2 (see Eq. 2.33), so the FPE in the stationary regime can be rewritten as

$$\{g, \rho\} = -\eta\vec{\nabla} \cdot \left(\vec{r}\rho + D'\eta\vec{\nabla}\rho\right). \tag{2.64}$$

Note that when $\eta = 0$, *any* function of the form $\rho(g(Q, P))$ will be a solution to this FPE. Therefore, we can expand

$$\rho(Q, P) = \rho^{(0)}(g(Q, P)) + \eta\rho^{(1)}(Q, P) + O(\eta^2), \tag{2.65}$$

and integrate over a contour of constant $g(Q, P)$. We will sometimes refer to such theory as "quasi-Hamiltonian" theory. After some algebra, this leads to:

$$\frac{d}{dE}\left(B(E)\rho^{(0)}(E) + D'\eta D(E)\frac{d\rho^{(0)}}{dE}\right) = 0, \tag{2.66}$$

where $B(E)$ and $D(E)$ are given by

$$B(E) = \int\int_{g(Q,P)=E} dQ\, dP, \tag{2.67}$$

$$D(E) = \int\int_{g(Q,P)=E} \nabla^2 g(Q, P)\, dQ\, dP. \tag{2.68}$$

This approach has been followed by Dmitriev and D'yakonov in [77] An alternative approach followed by [21] is to calculate an average rate of energy decay due to friction and an effective diffusion coefficient for the energy drift due to noise. These would turn out to be precisely the $B(E)$ and $D(E)$ respectively, hence Eq. (2.66) is the FPE for the energy. Yet an entirely different approach has been undertaken in the paper by Dykman [65]. Returning back from D' to D, the activation rate is $W \propto \exp(-R/D)$ where

$$R = -\eta \int_{E_a}^{E_s} \frac{B(E')}{D(E')}\, dE'. \tag{2.69}$$

Here E_a is the energy of attractor (it is strictly speaking a center at zero η but becomes an attracting fixed point for finite η), and E_s is the energy of the saddle (it remains a saddle even at zero η). To compute $R_{up\to down}$ the relevant "attractor" is the one with larger amplitude and to compute $R_{down\to up}$ the relevant "attractor" is the one with smaller amplitude (see Fig. 2.1).[5] We plot both of these R/η in Fig. 2.5. It is worth mentioning that in steady-state the ratio of rates gives the ratio

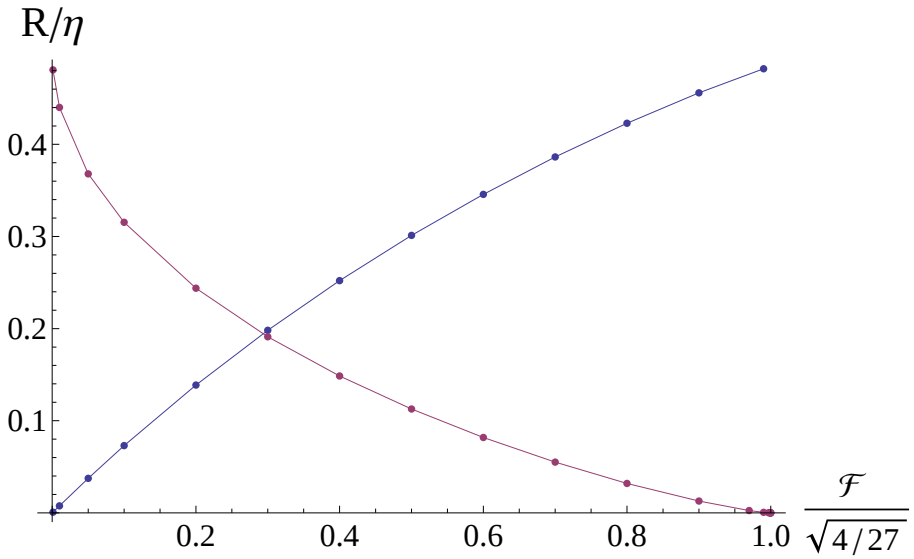


Figure 2.5: Activation barriers in the limit of $\eta \to 0$ given by Eq. (2.69). The curve that increases from left to right corresponds to $R_{up\to down}$ and the curve that decreases from left to right corresponds to $R_{down\to up}$.

of occupation probabilities of high or low amplitude states. Thus the $\mathcal{F}$ at which $R_{down\to up} = R_{up\to down}$ is the point at which the most probable states change. Both are experimentally relevant quantities and have in fact been measured; see recent NEMS experiment of Aldrige and Cleland [22] for example.

Close to bifurcation points

$$R \approx \begin{cases} 2\eta\delta & \text{low field bifurcation [65]}, \\ \frac{4}{3^{1/4}}\eta\delta^{3/2} & \text{high field bifurcation [77]}, \end{cases} \tag{2.70}$$

[5]For these down $\to$ up transitions there are two disjoint contours of g at a given energy E. The smaller contour immediately around the low-amplitude center needs to be chosen in the calculation of $B(E)$ and $D(E)$ in Eqs. (2.67) and (2.68).

Of course these equations in particular, and the whole low-damping method (or quasi-Hamiltonian method) of this subsection breaks down sufficiently close to bifurcation points when η is finite (because the trajectories are no longer approximately Hamiltonian trajectories), and must be replaced by the strong-damping method (or quasi-1-dimensional method) of the previous subsection. As $\eta \to 0$ the interval of δ over which the latter is applicable shrinks (as a reminder, δ is the deviation of $\mathcal{F}$ from the value of $\mathcal{F}$ at which there is a saddle-node bifurcation point). There must be a crossover from one regime to the other at some δ. Indeed, comparison of Eqn. (2.70) with Eqn. (2.63) reveals that there must be a crossover from a $\xi = 1$ regime to $\xi = 3/2$ regime as $\mathcal{F}$ approaches $\mathcal{F}_B^l$ for up $\to$ down transitions near the low-field bifurcation point. This is the part of my research which is presented in the next chapter.

2.3 Power Law Scalings of Escape Barriers

The aim of this section is to discuss conditions when a generic system is expected to exhibit the scaling exponents 3/2 and 1. The discussion may appear somewhat detailed, but for the most part, we summarize recent work of Dykman, Schwartz and Shapiro [21] in which the scaling exponent at nonzero η was related to the geometry of homoclinic orbits at zero η.

We just saw that the two analytical limits, the low-η limit and the overdamped limit close to bifurcation points exhibit their own scaling laws. In general, the inquiry into scaling exponents is an interesting one because these exponents may be universal. For example, the 3/2 scaling is ubiquitous, and it also appears in the original work of Kramers [13]. This is a result of a universal structure of a dynamical system near the saddle-node bifurcation point. When $\eta \neq 0$, sufficiently close to the saddle-node bifurcation, the dynamics near the attractor and the saddle is generically overdamped [78, 79]. This leads to the soft mode 1-dimensional geometric picture offered in Section 2.2.1, which leads to the cubic potential, and finally the exponent 3/2. Hence, $\xi = 3/2$ is a genetic property in a system with finite damping and close enough to the saddle-node bifurcation.

Further from the bifurcation, the overdamped soft mode does not exist. We have seen that for sufficiently low η, the overdamped $\xi = 3/2$ crosses over to $\xi = 1$ (close to $\mathcal{F}_B^l$) or remains $\xi = 3/2$ (close to $\mathcal{F}_B^h$). As recently proved by Dykman, Schwartz and Shapiro [21] (DSS), these are in fact also *generic* scaling laws in the underdamped vicinity of a saddle-node bifurcation. DSS considered two scenarios which may happen when η *equals* zero. As $\delta \to 0$, the center and the saddle may merge, or they may remain separate. DSS named the first type of *Hamiltonian* bifurcation local and the latter nonlocal. The two situations dictate the form of the Hamiltonian g.

Let us switch to coordinates (q, p) such that when $\eta = 0$, both the center and the saddle fixed points lie on the $p = 0$ axis. In the case of a local Hamiltonian bifurcation, in the vicinity of $(q = 0, p = 0, \delta = 0)$, g is given by $g_L(q, p, \delta) = \frac{1}{2}p^2 + U_{cub}(q)$ where $U_{cub}(q) = -\frac{1}{3}q^3 + \tilde{\eta}q$. The center lies inside a small homoclinic loop at the energy of the saddle. DSS proved that given this form of the Hamiltonian g, addition of noise and damping leads to $R(\delta) = A(\eta)\delta^{3/2}$ (in the underdamped regime). But we know that as as the system approaches the bifurcation even closer, it becomes overdamped, a soft mode sets up and the method of analysis presented in Section 2.2.1 leads to $R(\delta) = B(\eta)\delta^{3/2}$. Therefore, assuming smooth dependence upon δ, it must be that $A(\eta) = B(\eta)$. In other words, the $\xi = 3/2$ power law in this situation holds beyond where it was expected to hold due to a soft mode theory! In the Duffing system, a nonlocal bifurcation takes place at $\mathcal{F}_B^h$. In the numerical experiments described below, we found that the $\xi = 3/2$ power law holds near $\mathcal{F}_B^h$ over a *finite* range of δ even as $\eta \to 0$. This $\xi = 3/2$ law at zero η has been predicted by [77], but the new work of DSS gives it a new explanation. We elaborate upon this in Section III.B.

In the case of a nonlocal Hamiltonian bifurcation, g is given by $g_{NL}(q, p, \delta) = \frac{1}{2}p^2 + \delta U(q)$, where $U(q)$ is independent of δ and has a local minimum and maximum. The homoclinic loop of such systems has the property that only its size in the p-direction shrinks to zero as $\delta \to 0$, while the distance between the center and the saddle inside the loop remains finite. Introduction of damping into these types of systems will not change the separation between the saddle and the attractor until the system has come close enough to the bifurcation: $\delta \sim \eta^2$. DSS proved that given

this form of the Hamiltonian g, addition of noise and damping leads to $R(\delta) \propto \delta$ (in the underdamped regime). Therefore, at finite η we expect a $\xi = 1$ scaling moderately close to the bifurcation, but further away then the region of overdamped $\xi = 3/2$ scaling. In the Duffing system, a nonlocal bifurcation takes place at $\mathcal{F}_B^l$. It is important to stress that all saddle-node bifurcations at *finite* η are local in the sense that the attractor and the saddle merge as $\delta \to 0$. The terms local or nonlocal only make sense at zero η. We must also mention that a nonlocal bifurcations may take place in many forms. For example, in a Duffing oscillator, the homoclinic loop has a horseshoe shape at $\eta = 0$ which becomes thinner while remaining extended as $\delta \to 0$. DSS found a change of variables which maps from g to g_{NL}.

Chapter 3

Scaling Crossovers

In this chapter we study properties of crossovers between different scaling regimes. Each scaling regime and the corresponding scaling exponent may have a specific physical explanation; we have seen two examples: the overdamped soft mode close to the bifurcation or the separation of the damped and the hamiltonian timescales further away from the bifurcation. In such cases an analytical approximation may be possible, as in fact it was in the aforementioned cases. These analytical approximations often take the form of a perturbation series. The regime of crossovers between different scaling exponents is also interesting, however, it is less likely to afford an analytical approximation because this is precisely where the perturbation theory applicable in each regime must fail. Perturbation series may still be useful in estimating the characteristic beginning point or the characteristic end point of a crossover region. To do this one may develop the perturbation series in each scaling regime to one higher order and look at the values of δ where the correction to the leading order becomes comparable in magnitude to the leading order itself. We will in fact do this below, but clearly, this approach is not a rigorous, but one designed to make an estimation. Thus in understanding crossovers, the power of numerics is very useful and is an appropriate first step in trying to understand crossovers between scaling regimes.

We would like map out the $(\eta, \mathcal{F})$ parameter space (cf. Fig. 2.3) for regions in which each scaling regime holds and for locations of crossover regions between these scaling regimes. This will tell us the characteristic location of a crossover and the characteristic width of the crossover region. In what follows we first outline the

numerical method used, then present the results and discuss the findings.[1]

3.1 Numerical Method

Numerical calculations of escape rate are based on direct evaluation of S^*, the quantity defined in Sec. 2.1.3. The algorithm for doing so is as follows. (1) Choose parameters η and $\mathcal{F}$. (2) Calculate repulsive eigenvectors (i.e., those whose eigenvalues have positive real part) of the FP ($Q_{attr}, P_{attr}, p_Q = 0, p_P = 0$) of the auxiliary system. These eigenvectors span the tangent plane to the unstable manifold of this attractor. (3) Position a locus of initial conditions on a small circle that lies on this tangent plane and centered around this FP and evolve each of these initial conditions according to the Hamiltonian equations with Hamiltonian $\mathcal{H}$ from Eqn. (2.48). The angle around this circle, φ, serves as a parameter that enumerates a trajectory. [2] (4) From this set of trajectories, find the one which leads into the saddle, with a special trajectory parameter φ^*. Because the method involves "shooting" many trajectories to see which one hits the saddle, this method is sometimes called "the shooting method". (5) Evaluate the action along this trajectory. For each η, this procedure would be repeated for many values of $\mathcal{F}$ between $\mathcal{F}_B^l$ and $\mathcal{F}_B^h$.

In practice, step (4) may require several rounds of bracketing. In each round, we shoot a given number of trajectories in some range of φ, starting from the 2π range in the first round. For each trajectory we calculate the arc length on the separatrix between the saddle and the point of intersection of that trajectory with the separatrix. We identify a φ for which the trajectory came closest to the saddle and then evolve another set of trajectories in a small range of trajectory parameters around this φ. After more and more rounds of such procedure, we obtain a trajectory that hits the separatrix closer and closer to the saddle. In general, finding a trajectory that connects an attractor with the saddle is difficult because it lies on the intersection of

[1] A preprint of our paper related to this chapter can be found in [80].

[2] It turns out to be not the ideal way to parametrize trajectories, since the local pattern of spiraling trajectories may be elliptical, so different values of φ may actually lie on the same trajectory. This made this method a little inconvenient, but it could still be used. A different technique is to sample initial conditions situated on a straight line in between two windings of a spiral.

the unstable manifold of the attractor and the stable manifold of the saddle in the 4D auxiliary space [74]. [3] A small deviation of φ on opposite side of φ^* will produce trajectories that lie on opposite sides of this intersection and will diverge from each other. This effect is most pronounced closer to bifurcations because the motion along the optimal trajectory becomes much slower then the diverging motion away from it - such is the structure of eigenvalues of the attractor-saddle pair close to the bifurcation. This polarity takes place only close to bifurcation points, but it persists farther away from bifurcation points when η is larger. Therefore, in the small-η regime, finding a trajectory between an attractor and the saddle is difficult (i.e., requires many rounds of bracketing) only very close to bifurcation points, but in the large-η regime, finding such a trajectory is difficult over a larger part of the hysteresis.

We will now present results of the computation of S^*, scaling laws, and crossovers. It is sometimes easier to present results in terms of

$$x \equiv \frac{\delta}{\mathcal{F}_B^h - \mathcal{F}_B^l}, \tag{3.1}$$

the "reduced" δ. For each friction η and each type of transition (there are two: up $\rightarrow$ down and down $\rightarrow$ up), we scanned the hysteretic region at multiple values of x, and for each x we followed the shooting procedure, and where necessary, the bracketing method. We first summarize our findings and then discuss the details for each type of transition separately.

[3] The question of *existence* of this intersection, to our knowledge, is not a fully understood one.

3.2 Results

We first summarize the main results and then discuss them in detail. For the up $\rightarrow$ down transitions, we identified three scaling regimes of S^* versus $\mathcal{F}$ at fixed η: the power law with exponent 3/2 close to the bifurcation, followed by a regime with the scaling exponent 1 further from the bifurcation, followed by a nonpower law regime. The regions of the $(\eta, \mathcal{F})$ parameter space for each regime are mapped out, showing the *disappearance* of the 3/2 regime as $\eta \rightarrow 0$. We point out the consistency of this behavior with generic predictions of DSS. We also find that for finite η the exponent 3/2 holds over a much wider range of parameters than where it would be expected to hold based on the overdamped 1-dimensional reduction close to bifurcation points.

For the down $\rightarrow$ up transitions, we identified a new intermediate scaling with the exponent of ≈ 1.3, which follows the power law with exponent 3/2 close to the bifurcation; the nonpower law regime again takes place furthest from the bifurcation. The regions of the $(\eta, \mathcal{F})$ parameter space for each regime are mapped out, showing the *persistence* of the 3/2 regime as $\eta \rightarrow 0$ (again underscoring the nonoverdamped nature of this exponent). These observations are also consistent with the generic predictions of DSS.

3.2.1 Up $\rightarrow$ Down Transitions

3.2.1.1 Numerical Results

A plot of S^*/η versus x, Figs. 3.1, 3.2 reveals several features. At low η two regimes clearly stand out: one with $\xi \approx 3/2$ (regime A), and one with $\xi \approx 1$ (regime B). Further from the bifurcation, the $\xi = 1$ scaling law breaks down and S^*/η versus x displays a nonpower-law behavior. For various values of η, the data in the regime A fits well to a power law with ξ which is up to 3% below the theoretical value of 3/2. In general, precise extraction of the exponent 3/2 requires approaching very close to the bifurcation. It appears that the discrepancy was highest in those cases when the bifurcation was approached less closely (relative to the crossover point) or

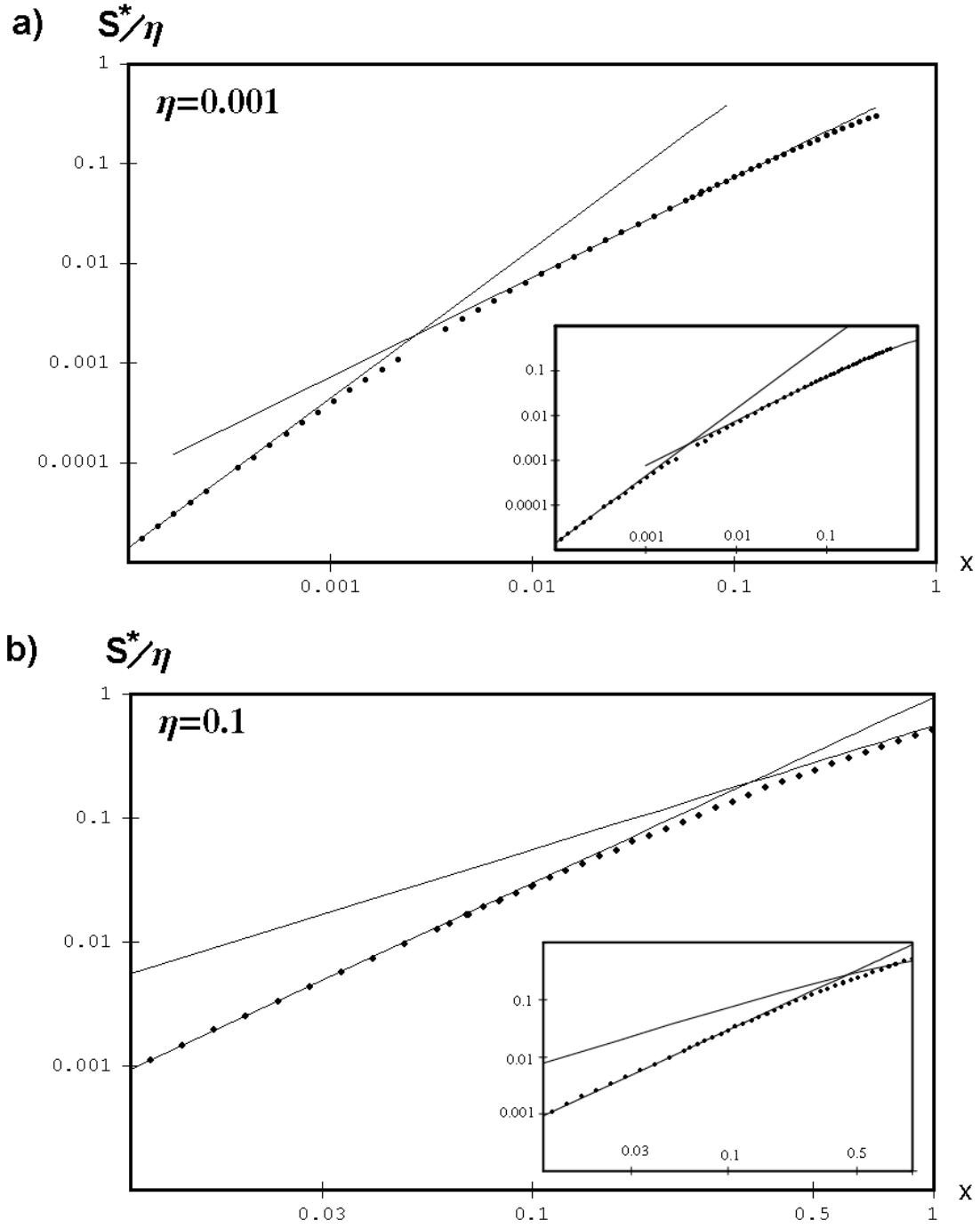


Figure 3.1: Main plots: $S^*(x)/\eta$ (dots) and *fits* (lines) for up $\rightarrow$ down transitions. The fit to the $\xi = 3/2$ regime was made by analyzing $\ln(S^*/\eta)$ versus $\ln x$ and a fit to the $\xi = 1$ regime was made by analyzing S^*/η versus x. Inserts: $S^*(x)/\eta$ (dots) and *theory* (lines) based on Eq. (2.63) for the $\xi = 3/2$ regime and on the quasi-Hamiltonian theory, Eq. (2.69).

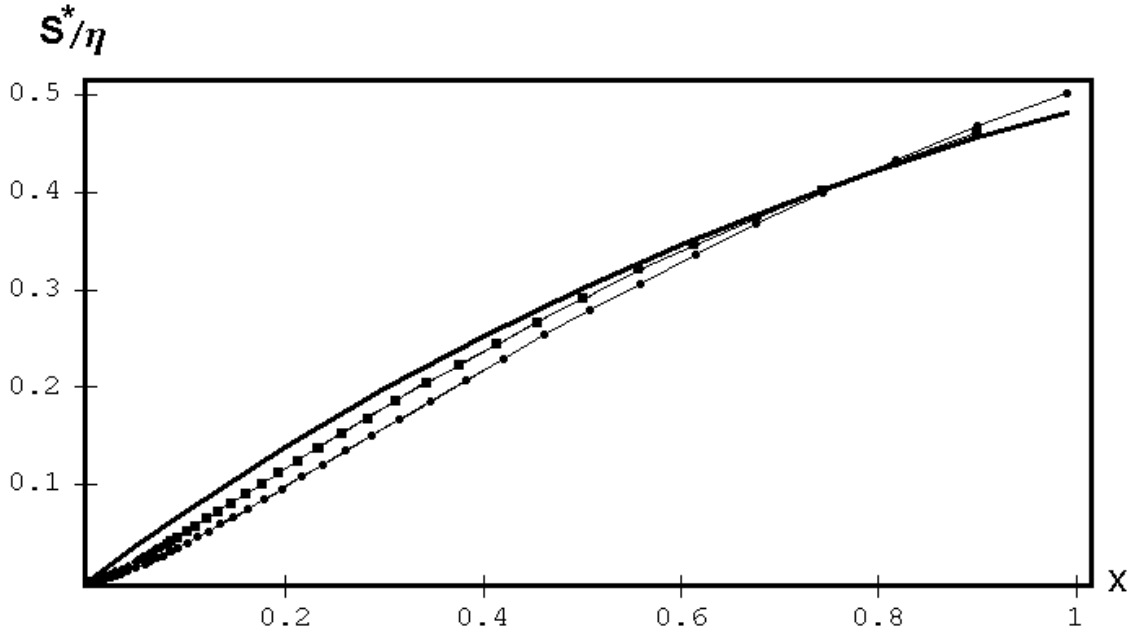


Figure 3.2: $S^*(x)/\eta$ on a linear scale. The thick black line represents the zero damping theory, Eq. (2.69), the same as one of the curves in Fig. 2.5. The other curves represent numerical calculations of S^*: squares for $\eta = 0.03$ and circles for $\eta = 0.06$. Note the linear regime at finite η is delayed until larger $\mathcal{F}$, giving way to the 3/2 regime prior to it.

less reliably. The reliability of the approach is hampered at larger η by the strong divergence of trajectories explained in the previous section.

The pattern of escape trajectories for various trajectory parameters is shown in Fig. 3.3 (see also appendix C for a survey of other interesting features of trajectories). We plot several (~ 10) optimal escape trajectories for various values of $\mathcal{F}$ and η as thin dotted lines (see footnote 4 of Chapter 3 to recall the two meanings of the word "escape"; here the first meaning is implied). Thick dotted lines denote the separatrix. In each case, only one of the trajectories leads to the saddle.

A crossover between the $\xi = 3/2$ and $\xi = 1$ regime is defined to be such x at which fits to the respective regions intersect. A crossover between the $\xi = 1$ and a nonpower law regime was defined to be the point at which the fit to $\xi = 1$ regime came closest to the $\ln(Q^*/\eta)$ versus $\ln(x)$, see Fig. 3.1. Collecting data for S^* versus x for various η we were able to map out the entire bistability region for locations of

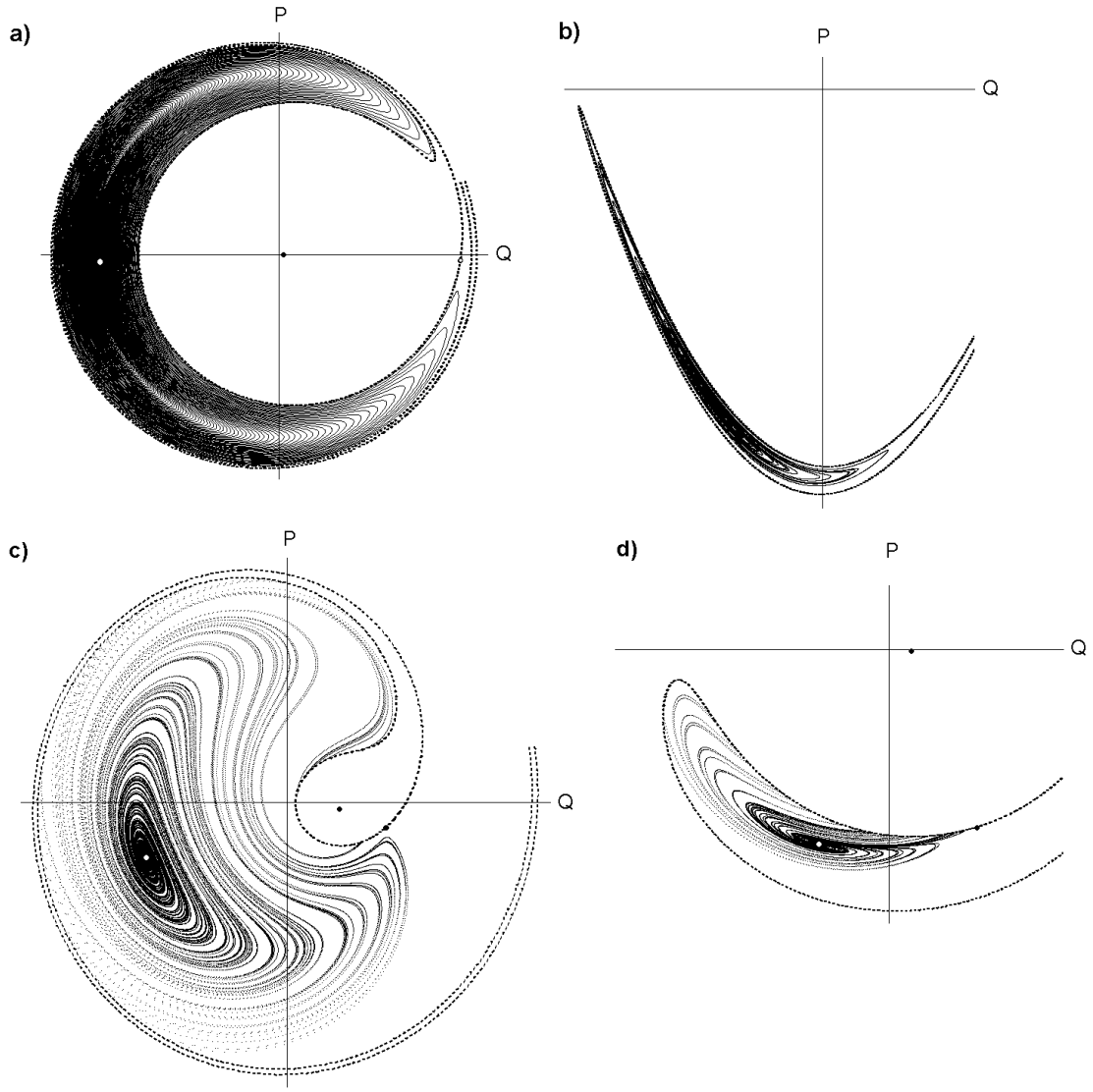


Figure 3.3: (a),(b) Up $\rightarrow$ down escape trajectories for $\eta = 0.001$; (a) at 7% of the hysteresis (i.e. $x = 0.07$), (b) at 0.001% of the hysteresis. (c),(d) Up $\rightarrow$ down escape trajectories for $\eta = 0.1$; (c) at 80% of the hysteresis, (d) at 0.001% of the hysteresis.

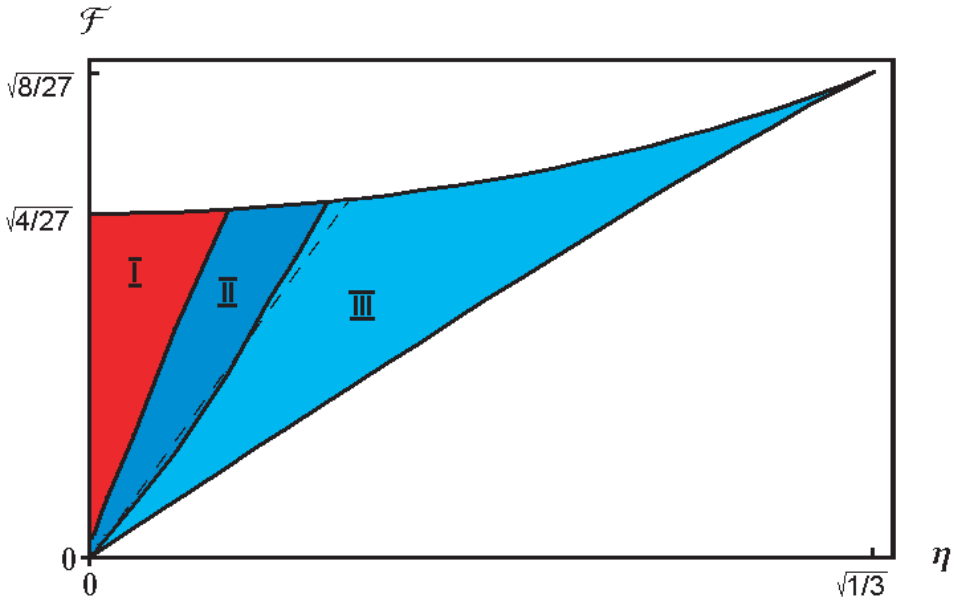


Figure 3.4: Regions of different scaling behaviors for up → down transitions obtained by scanning $\mathcal{F}$ at fixed η for multiple values of η. The meaning of the three scaling regions is described in the text. The dashed line is a simple prediction of the crossover between the $\xi = 1$ and $\xi = 3/2$ regime expressed by Eq. (3.3).

various scaling regimes and crossovers between them. The computations were made down to $\eta = 10^{-3}$. The result of this work is depicted in Fig. 3.4. Three distinct regions have been identified; region I: nonpower-law regime, region II: $\xi = 1$ regime, and region III: $\xi = 3/2$ regime. We remind the reader that in Fig. 3.4 for up → down transitions and in Fig. 3.9 for down → up transitions, the mapping was performed by scanning $\mathcal{F}$ while holding η fixed. Note in Fig. 3.4 that as $\eta \to 0$, the linear scaling is found between $x = 0$ and some finite x. At small, but nonzero η, the region of $\xi = 3/2$ subsequently grows. These results are in accord with the theorem of DSS which states that close to a bifurcation which at *zero* η becomes nonlocal (i.e., the low $\mathcal{F}$ bifurcation in our system), the underdamped regime (regime B) at *nonzero* η is expected to give $\xi = 1$. Since the underdamped portion of the hysteresis extends all the way to the bifurcation as $\eta \to 0$, we expect the linear scaling to extend all the way to the bifurcation in this limit.[4]

[4]In practice we could go down to $\eta = 0.001$. In principle, strictly at $\eta = 0$ escape trajectories do not exist; they only have meaning in the limit $\eta \to 0$.

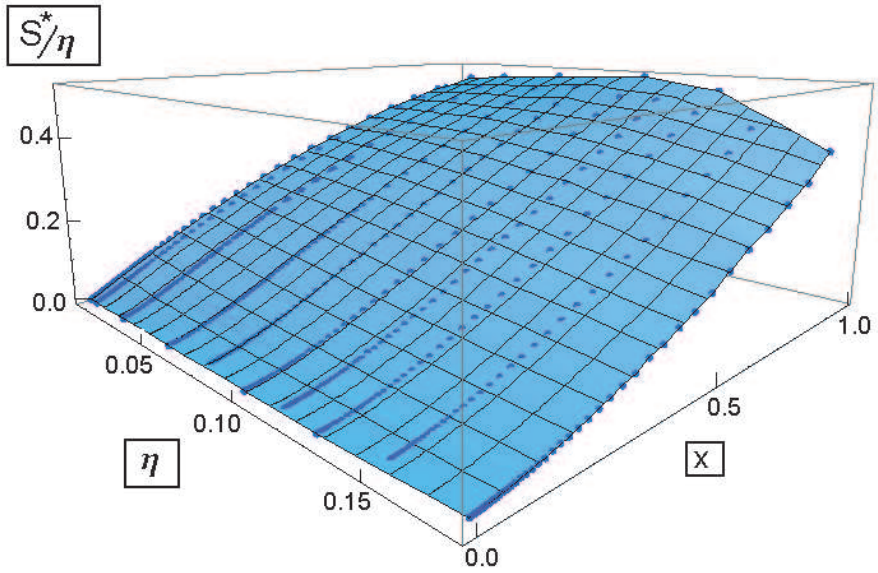


Figure 3.5: S^*/η versus x and η up to $\eta = 0.175$ which is about 30% of the hysteresis. Fig. 3.1 are two slices of this surface.

The boundaries between the different scaling regions I, II and III are purely conventional - they have been defined in some convenient way, but do not correspond to anything specifically physical. To emphasize this point, it helps to construct a 3-dimensional plot of S^*/η versus x and η - Fig. 3.5. This surface is smooth and large parts of it happen to fit well to power laws. The regions when this is possible appear to be much bigger then the regions at which the $\xi = 1$ and $\xi = 3/2$ scaling were supposed to be applicable on the basis of approximations discussed in Section 2.2. A natural question is why this is so? To shed more light on this issue, let us digress on the crossover from the $\xi = 3/2$ to $\xi = 1$ in more detail.

3.2.1.2 Discussion of the $\xi = 3/2$ to $\xi = 1$ Crossover

We now discuss the various ways in which the crossover from the $\xi = 3/2$ to the $\xi = 1$ regime may be defined. Notice that in Fig. 3.1 such crossover was defined as the point of intersection of the fits to the respective regimes. Thus, to predict this crossover we simply equate the asymptotic in the $\xi = 3/2$ regime, $R = \frac{4\sqrt{2}}{3}\eta\delta^{3/2}$ to the asymptotic

in the $\xi = 1$ regime $R = 2\eta\delta$ (this is most accurate as $\eta \to 0$). The result is

$$\delta = \frac{9}{8}\eta, \tag{3.2}$$

which corresponds to

$$\mathcal{F} = \mathcal{F}_B^l + \frac{9}{8}\eta, \tag{3.3}$$

($\mathcal{F}_B^l$ is itself a function of η). The crossover predicted this way is depicted by the dashed line in Fig. 3.9. The fact that this method of predicting a crossover works well tells us that the asymptotic expressions are correct. It does not teach us about the crossover region itself: if it is a very wide region for example, the crossover defined by equating the fits or predicted by equating the asymptotics is not a particularly physically meaningful quantity. The width of the crossover region, for example, is a more physically interesting quantity. Alternatively, we may also define the following two crossovers: a "lower crossover" (LC) defined as a characteristic value of $\mathcal{F}$ (or δ or x) above which the $\xi = 3/2$ regime begins to fail and an "upper crossover" (UC) defined as a characteristic value of $\mathcal{F}$ (or δ or x) below which the $\xi = 1$ regime begins to fail. The width of the crossover region is the difference between these two values.

Physically, we expect the LC to take place where the soft mode picture fails. Therefore, a good *characteristic* estimate of this will be such δ at which the real and imaginary parts of eigenvalues become equal. This can be shown to happen for $\delta = \eta^3/2$, suggesting the cubic scaling. To predict the LC more accurately we turn back to the type of analysis described in section 2.2.1. Let us rewrite Eqs. (2.56)-(2.57) for the nonadiabatic part of the fast variable $A = f - a_B s$. Near the low-field bifurcation, these will have the form

$$\dot{A} = -a_B\delta - 2\eta A + a_1 s^2 + a_2 As + a_3 A^2 + a_4 s^3 + a_5 s^2 A + a_6 sA^2 + a_7 A^3 + n_x - a_B n_y, \tag{3.4}$$

$$\dot{s} = \delta + b_1 s^2 + b_2 As + b_3 A^2 + b_4 s^3 + b_5 As^2 + b_6 A^2 s + b_7 A^3 + n_y. \tag{3.5}$$

All coefficients a_i and b_i are functions of η. The position of the fixed points sets the characteristic scaling of A to be δ and of s to be $\delta^{1/2}$ (a generic feature of saddle-node

bifurcations), so we define new variables: $s = \delta^{1/2} s'$ and $A = \delta A'$. Moreover, the η-dependencies of the coefficients a_i and b_i near the low-field bifurcation are such that when η is small compared to 1, further rescaling of s' and A' as $s' = \eta^{1/2}\sigma$ and $A' = \eta^{-2}\phi$ yields a single dimensionless parameter $\mu = \delta/\eta^3$; such reduction is approximate, but becomes asymptotically accurate for smaller η.[5] In terms of μ, σ and ϕ we have

$$\frac{d\phi}{d\tau} = \left(-1 - 2\phi + \frac{3}{2}\sigma^2\right) + \mu^{1/2}\left(\sigma\phi + \sigma^3\right) + \mu\left(3\phi\sigma^2 - \frac{\phi^2}{2}\right) + \mu^{3/2} 3\sigma\phi^2 + \mu^2\phi^3 + N_\phi, \tag{3.6}$$

$$\frac{d\sigma}{d\tau} = \mu^{1/2}\left(1 - \frac{\sigma^2}{2}\right) + \mu\left(\sigma\phi - \sigma^3\right) + \mu^{3/2}\left(\frac{3}{2}\phi^2 - 3\sigma^2\phi\right) - \mu^2 3\sigma\phi^2 - \mu^{5/2}\phi^3 + N_\sigma, \tag{3.7}$$

where $\tau = \eta T$. With the help of the definition of σ and Eq. (2.58) we note that the zeroth-order part of the fast variable $a_B s$ is given by $\delta^{1/2}\eta^{-1/2}\sigma$ and the correction due to nonadiabaticity of the fast variable is $A = \delta\eta^{-2}\phi$. We also see from Eqs. (3.6)-(3.7) that when $\mu \ll 1$, both σ and ϕ are $O(1)$. Hence comparison of the two terms says that when $\mu \ll 1$, *nonadiabatic part/adiabatic part* $\sim (\delta\eta^{-2})/(\delta^{1/2}\eta^{-1/2}) = \mu^{1/2} \to 0$ as $\mu \to 0$. In other words, in this regime the nonadiabatic part becomes unimportant and the problem reduces to 1 dimension defined by the noisy dynamics of the slow variable. To do this reduction explicitly, we first notice from Eqs. (3.6)-(3.7) that at lowest order, the variable σ can be considered constant while $\phi(\sigma) \approx \frac{3}{4}\sigma^2 - \frac{1}{2} +$ nonadiabatic corrections. This $\phi(\sigma)$ is then substituted into Eq. (3.7). The resulting expression can only be considered up to $O(\mu)$ since the higher-order terms must also include the nonadiabatic corrections to ϕ to be complete [note that if the $\mu^{1/2}$-terms in Eq. (3.7) contained any ϕ terms, the complete expression within the adiabatic approximation would only extend to $O(\mu^{1/2})$, not $O(\mu)$]. The barrier height in the

[5] From our definitions, $\omega_F = \omega_0\left(1 + \frac{1}{Q\eta}\right)$. Recall that in order for equations (2.36)-(2.37) to apply, we must be working in the regime when $\omega_F = \omega_0(1 + \text{small number})$. Therefore, if the limit $\eta \to 0$ is taken, it is implied that the limit $Q \to \infty$ is taken first. Moreover, wherever the limit $\delta \to 0$ is taken, it is taken last, because we talk about approaching the bifurcation at *constant* η. However, since η is always less then $\eta_c \sim O(1)$, the η is already small (compared to 1) and so wherever the statements about "small" η are made, they do not imply the formal $\eta \to 0$ limit.

resultant 1-dimensional potential is

$$\Delta U = \frac{4\sqrt{2}}{3}\mu^{1/2} + \sqrt{2}\mu^{3/2}. \tag{3.8}$$

Again, the correction is within the adiabatic approximation; a nonadiabatic correction would be higher order in μ. Taking into account that the N_σ has a diffusion constant $D_\sigma = \frac{D}{\delta\eta^2}$, we have

$$R = \frac{4\sqrt{2}}{3}\eta^{1/2}\delta^{3/2} + \sqrt{2}\frac{\delta^{5/2}}{\eta^{5/2}} \tag{3.9}$$

[compare with Eq. (2.63)]. We see again that the correction term grows to be dominant when $\mu \sim 1$ (i.e., when $\delta \sim \eta^3$), but this is precisely when the nonadiabatic part becomes comparable to the adiabatic part, and it turns out to be meaningless to treat the problem as effectively 1-dimensional. Our analysis taught us that in the regime when $\delta \ll \eta^3$, the activation barrier R will scale as $\delta^{3/2}$ to a good approximation. Outside the regime of $\delta \ll \eta^3$ the current analysis only allows us to conclude that if ξ remains to be 3/2, the physics behind this scaling is not the physics of a 1-dimensional overdamped soft mode: the escape problem is not 1-dimensional there.

To answer the question of what happens outside of the $\delta \gg \eta^3$ regime we chose to compute the most likely escape path numerically and to calculate S^* on that path. We looked at the difference between the zeroth-order theory, $R_0 = \frac{4\sqrt{2}}{3}\eta^{1/2}\delta^{3/2}$ and the numerically computed S^*. An example of this for $\eta = 0.01$ is shown in Fig. 3.6, but to be consistent with out plots, we display $R_1/\eta \equiv (S^* - R_0)/\eta$ versus x (i.e., $S^*(\delta)$ and $R_0(\delta)$ were reexpressed to be $S^*(x)$ and $R_0(x)$ and both were divided by a constant $\eta = 0.01$, which of course does not affect the slope of the fit versus x).

This procedure was repeated for several values of η and in each case, a fit to the function $\propto x^p$ was made at low values of x. The first few points at lowest x where the scatter was large were excluded; although the numerical calculations of S^* do not have an inherent randomness, the data exhibits a scatter at low x due to the difficulty of hitting the saddle and the discreteness of the time steps (the trajectory accelerates exponentially in the initial stages). As before, we did not use a systematic criterion

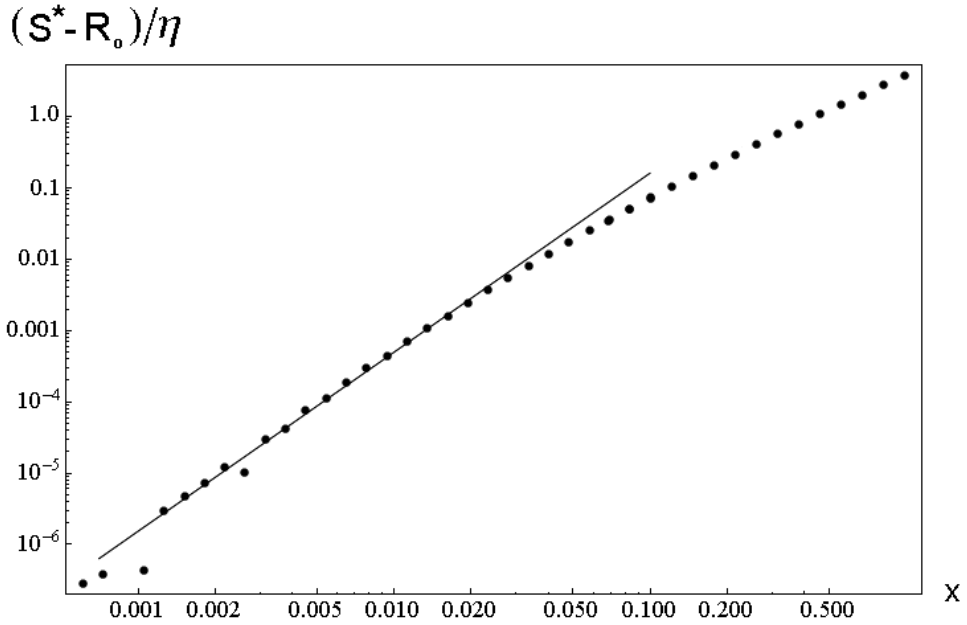


Figure 3.6: $R_1(x)/\eta \equiv (S^*(x) - R_0(x))/\eta$ plotted for $\eta = 0.01$. The straight line indicates a fit within the range spanning from $x \approx 0.0013$ to $x \approx 0.019$; the slope of the fit for this particular value of η is ≈ 2.51.

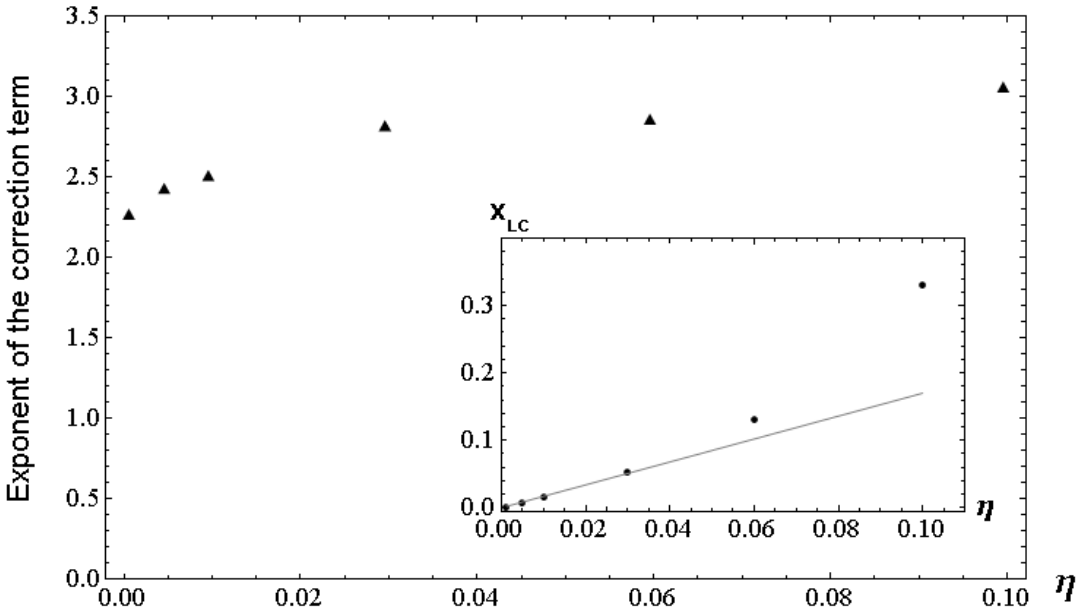


Figure 3.7: Main plot: the exponent p obtained from the fit of $R_1(x)/\eta$, plotted versus η. Insert, the position of LC defined as that x at which $0.2R_0(x)/\eta = R_1(x)/\eta$.

for the high end of the range of x over which the fit was made in accordance with the assumption that the value of the fitted slope approaches an asymptotic figure when the high end of the fitting range is not large. Values of p found this way are shown in Fig. 3.7. As $\eta \to 0$, p seems to approach ≈ 2.2, not far from $p = 5/2$ predicted above. However, the crossover LC defined as that x at which $0.2R_0(x)/\eta = R_1(x)/\eta$ scales *linearly* with η at small η, not cubically, as shown in the insert of Fig. 3.7.

The question of why the $\xi = 3/2$ scaling holds over a much larger region of parameter space then expected based on the 1-dimensional soft mode picture remains unanswered, and forms a good challenge problem for future work. One potential hypothesis is that in the region where the 1-dimensional soft mode picture does not hold true, the action along the MPEP actually grows only over a small portion of the MPEP where the motion is essentially 1-dimensional. We provide a plot of the action versus the distance along the trajectory, $S(\ell)$ and the derivative $dS/d\ell$ for a particular value of parameters deep in region III, Fig. 3.8. The derivative plot clearly shows that the action grows fastest in the last winding of the trajectory, but this is not sufficient to conclude that the dynamics in that part of the trajectory can somehow be described as effectively 1-dimensional.

Analogously to the lower crossover (LC), the upper crossover (UC) can be defined as a characteristic value of the driving field below which the linear scaling begins to break down. To predict this point, we must again compare the lowest and one higher order terms in the theory for $R(\delta)$ or $R(x)$ that applies in the $\xi = 1$ regime. The effect of a nonzero η was clearly demonstrated in Fig. 3.2. The lowest-order in dissipation (quasi-Hamiltonian) theory, Eq. (2.69) indeed predicts linear scaling for low values of δ (or x), which according to Eq. (2.70) is given by $2\eta\delta$. The region of linear scaling persists at larger η, but due to the existence of the 3/2-scaling region, this linear part has been pushed over to larger δ, and hence R takes lower values then those predicted by the zeroth-order theory (see Fig. 3.2). Chinarov, Dykman, and Smelyanskiy [81] worked out dissipative corrections to the quasi-Hamiltonian theory to next order in η. Their result, translated to our notation reads: $R = \eta\,(2\delta - \pi\eta)$ (compare with Eq. 2.70). The factor $\pi\eta^2$ is just that correction which takes this

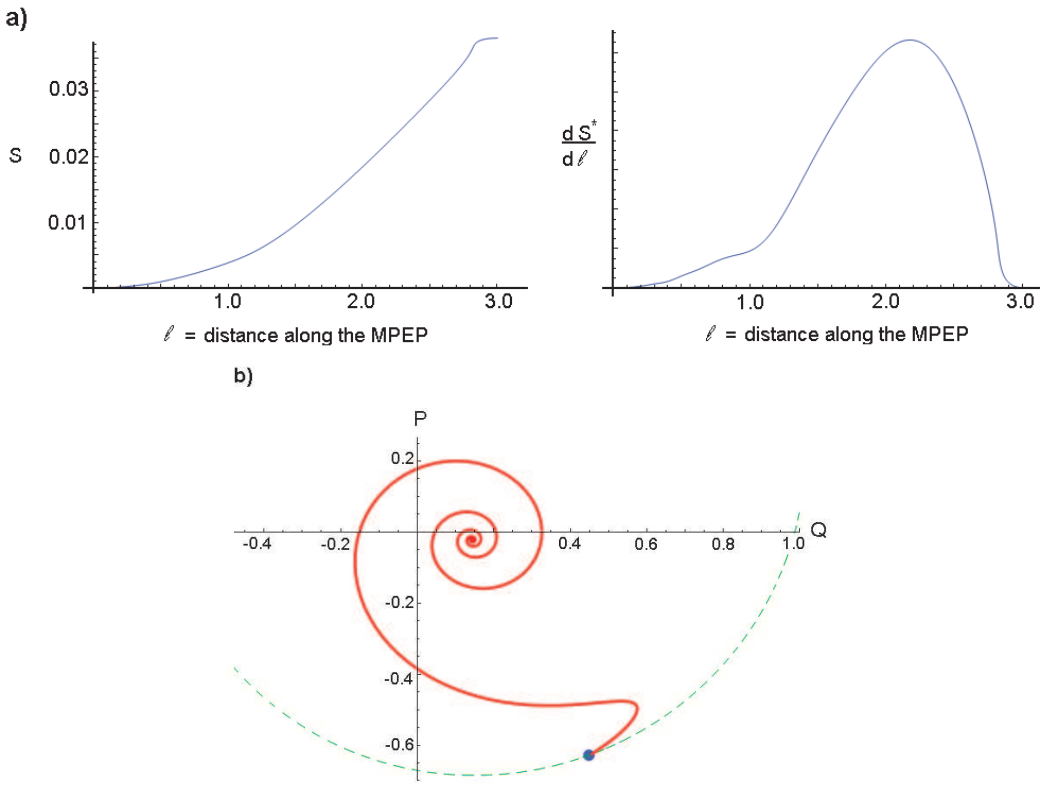


Figure 3.8: (a) $S(l)$ and its derivative for particular parameters in region III: here $\eta = 0.15$ and $\mathcal{F} = 10\%$ of the hysteresis, corresponding to $\mu \approx 5$; the soft mode picture is expected to hold only for $\mu \ll 1$. (b) The plot of the MPEP along which $S(l)$ in (a) is computed; dashed curve denotes separatrix.

lowering into account. We expect the linear scaling to start breaking down when the correction reaches some fraction r of the zeroth-order term, i.e., when $\delta_{UC} = \frac{\pi}{2r}\eta$ for low η.

To extract δ_{UC} from the data we followed the same procedure used to extract δ_{LC}: we subtracted the zeroth-order part $R_0(x)$ from $S^*(x)/\eta$. However, recall what $S^*(x)$ does: it has an initial regime with the leading exponent 3/2, followed by a linear regime for higher x, followed by a nonpower law regime for higher x still. On the other hand, R_0 starts out linear: it does not have the 3/2 for the up $\rightarrow$ down transitions (for down $\rightarrow$ up transitions, even R_0 has a 3/2 regime). Therefore, the function $(S^*(x) - R_0(x))/\eta$ grows for low x, slows down for intermediate x, and increases for yet larger x past the linear regime. Note that in the intermediate regime

the function $S^*(x)/\eta$ continues to grow because the exact slope in the linear regime is slightly different from the zeroth-order result! Thus, because $(S^*(x) - R_0(x))/\eta$ does not plateau in the intermediate, linear regime, it is hard to make quantitative predictions. We attempted to extract an average value in the intermediate regime, although is is somewhat ambiguous because it involves a conventional choice of the interval over which this averaging is performed. Nevertheless, equating this correction to the zeroth-order term yields δ_{UC} which does depend linearly on η.

3.2.2 Down → Up Transitions

The scaling regimes for the down → up transitions were also mapped out over the entire bistability region. The product of this mapping is shown in Fig. 3.9. The computations were made down to $\eta = 0.005$.

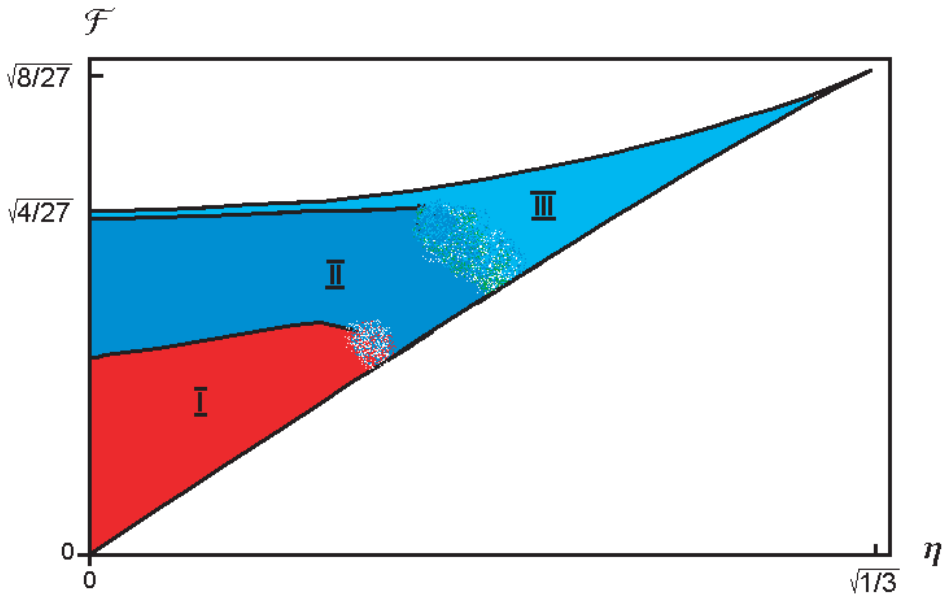


Figure 3.9: Regions of different scaling behaviors for down → up transitions obtained by scanning $\mathcal{F}$ at fixed η for multiple values of η. Region I corresponds to a nonpower-law regime. Region II corresponds to the $\xi \approx 1.3$ regime. Region III corresponds to the $\xi = 3/2$ regime. Fuzzy regions correspond to parts of the $(\mathcal{F}, \eta)$ plane which proved very difficult to map reliably due to the high divergence of trajectories.

A portion of the surface $S^*/\eta(x, \eta)$ is presented in Fig. 3.11 and two cuts at a fixed η

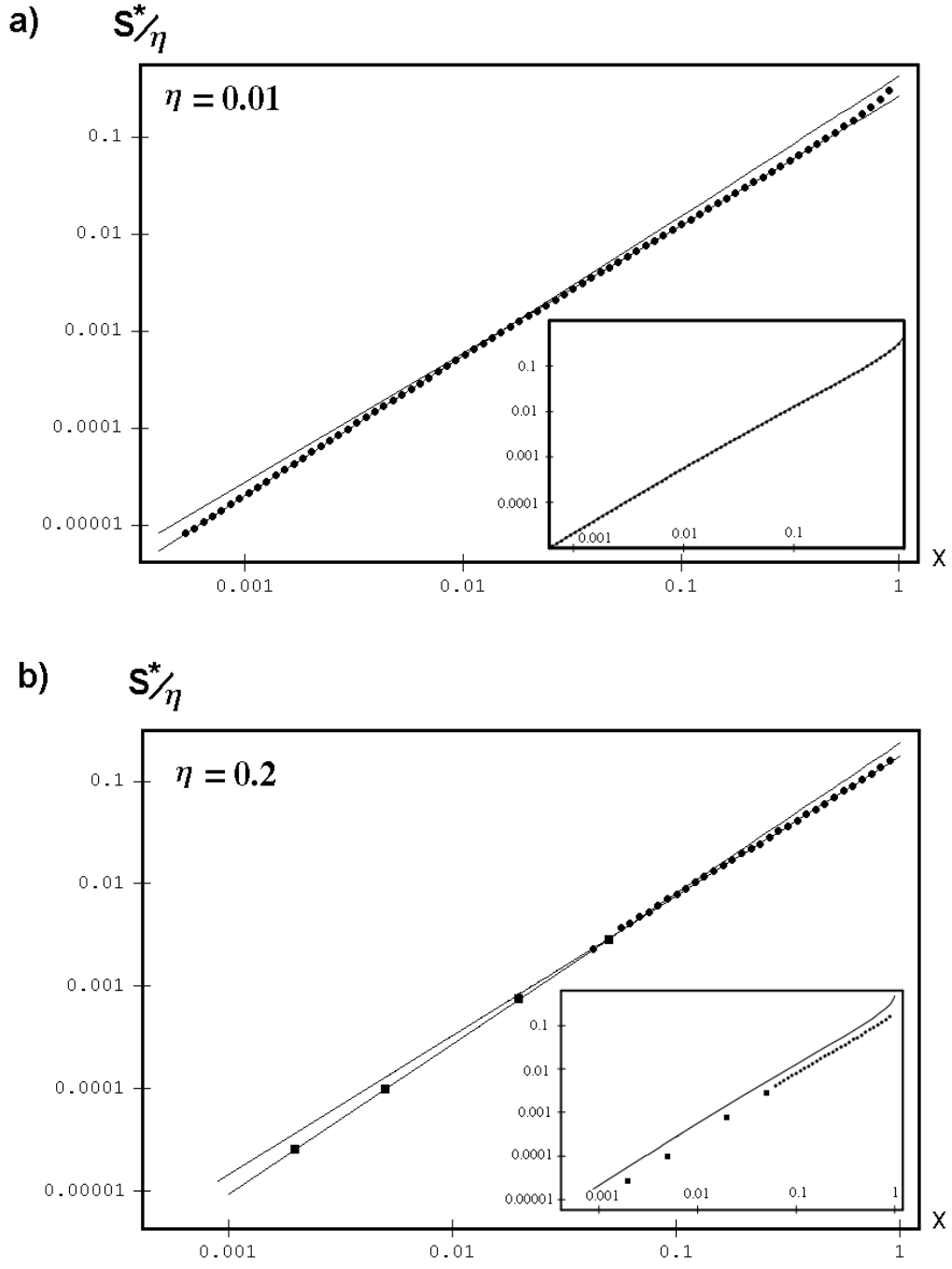


Figure 3.10: Main plots: $S^*(x)/\eta$ (dots) and *fits* (lines) for down $\rightarrow$ up transitions. Fits to the $\xi = 3/2$ and $\xi \approx 1.3$ regimes was made by analyzing $\ln(S^*/\eta)$ versus $\ln x$. In the region where the simple shooting method was unreliable for hitting the saddle due to the high divergence of trajectories, a bracketing method was used. The results from using this method are denoted by squares in plot (b). Inserts: $S^*(x)/\eta$ (dots) and *theory* (lines) based on the quasi-Hamiltonian theory, Eq. (2.69). Note: this theory predicts the $\xi = 3/2$ scaling, as was first pointed out by [77]; in the light of the recent work by [21] this was related to the nonlocality of the bifurcation at zero η.

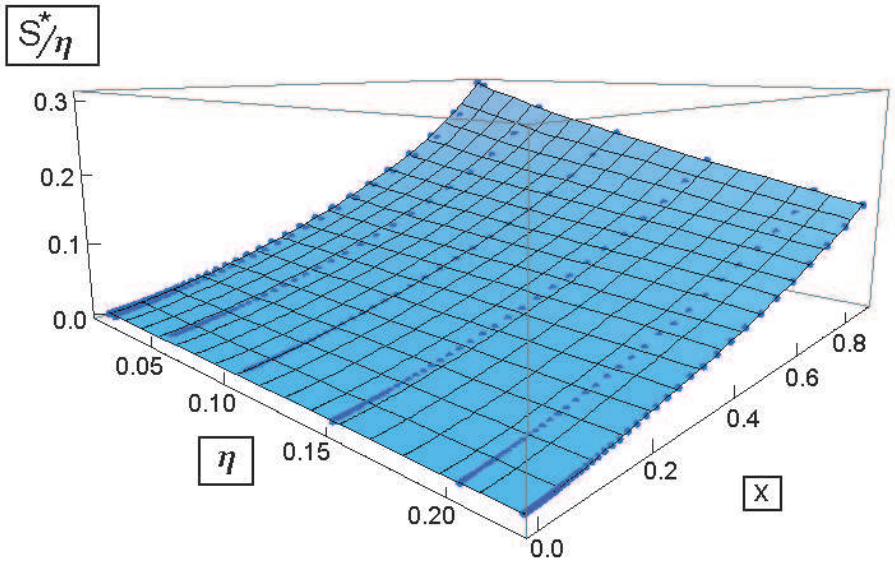


Figure 3.11: S^*/η versus x and η up to $\eta = 0.22$ which is about 33% of the hysteresis. Fig. 3.1 are two slices of this surface.

appear in Fig. 3.10. The low amplitude basin does not exhibit the nonlocal geometry discussed in Sec. 2.3 upon approach to the bifurcation at zero η, and correspondingly, the linear scaling is not found in the down $\rightarrow$ up transitions at any η (see Fig. 3.10), which is in accord with the DSS theorem. Instead, the $\xi = 3/2$ scaling gives way to a different scaling relationship further from the bifurcation, with an exponent $\xi \approx 1.3$. We currently do not have a physical explanation for the origin of such exponent; there may not be such an explanation, and 1.3 may simply be a power-law interpolation of 1 or 2 decades of numerical data. Note that region III in Fig. 3.9 (the $\xi = 3/2$ region) extends all the way to zero η - this is unlike the case of up $\rightarrow$ down transitions, where the $\xi = 3/2$ region disappears as $\eta \rightarrow 0$. These results are also in accord with the theorem of DSS which states that close to a bifurcation which is local at $\eta = 0$, the ξ is expected to be 3/2 in the underdamped regime at nonzero η. Since the underdamped portion of the hysteresis extends all the way to the bifurcation as $\eta \rightarrow 0$, we expect the 3/2 scaling to extend all the way to the bifurcation in this limit. At finite η, the region of the overdamped 3/2 scaling and this underdamped 3/2 scaling merge smoothly into each other. Had we assumed that $\xi = 3/2$ is due only to the overdamped soft mode,

we would follow the nearly adiabatic analysis of the previous section, and we would find that for low η there exists another dimensionless parameter $\nu = \delta^{1/2}\eta^{-2}$. When $\nu \ll 1$ such analysis predicts $\xi = 3/2$. When $\nu \sim 1$, the reduction to 1 dimension due to the adiabaticity becomes impossible and predictions based on this method can not be made. The numerical observation that ξ remains 3/2 beyond $\nu \ll 1$ confirms that the 3/2 scaling is not a consequence of reduction to an overdamped 1-dimensional soft mode.

In the previous section we offered three definitions of the 3/2-to-1 crossover. In the case of up $\rightarrow$ down transitions, we do not have a theory for the $\xi \approx 1.3$ scaling regime, so the crossover from the 3/2 scaling can be defined here only as a point at which the 3/2 scaling begins to break. Since the 3/2 scaling is predicted by the quasi-Hamiltonian theory, we look at next order corrections due to Chinarov, Dykman, and Smelyanskiy [81]. Translating the results of that work to our notation says that close to the bifurcation $R \approx \frac{4}{3^{1/4}}\eta\delta^{3/2} + 1.06\frac{4}{3^{1/4}}\eta^3 \left|\ln \eta\right| \delta^{3/2}$ (compare with Eq. 2.70). Evidently, the correction is to simply shift $R(\delta)$ by a constant offset (see Fig. 3.10 (b) where this effect is clearly pronounced throughout the entire $R(\delta)$). Because the correction is so small for small η, its effect does not become significant until a rather large η. Equating the zeroth-order term to the correction term predicts this value of η to be approximately $1.5 > \eta_c$. Numerically, this delayed effect is observed (see Fig. 3.9): the crossover does not seem to depend on η significantly until $\eta \approx 0.4\eta_c$. However, we state this with caution because for η larger than $\approx 0.4\eta_c$ we were not able to extract the position of the crossover accurately due to the high divergence of trajectories (the fuzzy region in Fig. 3.9).

Chapter 4

Consequence of Breaking of Detailed Balance

In this chapter we focus on the probability density function (PDF) in the (Q, P) phase space. We first bring to light the notion that a steady-state nonequilibrium system that satisfies detailed balance behaves essentially like an equilibrium system. It is the breaking of detailed balance that causes drastic differences to take place. We show that in our oscillator system, the detailed balance is broken at any finite value of the driving $\mathcal{F}$, so in this system the lack of detailed balance is equivalent to the lack of equilibrium; in other words the detailed balance is fragile in our system. We then present our results on the formation of nondifferentiable features in the PDF in the limit of vanishing noise strength[1]and find that the smoothness of the the PDF is also fragile. Section 4.1 introduces the meaning of detailed balance for Fokker-Planck models, states the consequences of detailed balance, proves an interesting relationship between the fluctuational and relaxational trajectories in the presence of detailed balance and finally lists the possibilities that might occur if detailed balance *does not* hold. Section 4.1 may seem somewhat formal, but the necessary information related to subsequent sections is given in its last several paragraphs. The section that follows, Sec. 4.2, studies the listed consequence of breaking of detailed balance specifically for our nonlinear oscillator.

[1]The resulting physics of is qualitatively analogous to the zero wavelength limit in optics (ray optics) or to the quantum tunneling in limit of $\hbar \to 0$.

4.1 Detailed Balance

There is a long history of research aimed at generalizing the methods of equilibrium statistical physics. One of the central questions is the feasibility of a nonequilibrium system to attain a Boltzmann-like distribution[2]

$$\rho(\mathbf{r}) = Ae^{-\phi(\mathbf{r})}. \tag{4.1}$$

Depending on the problem, ϕ may also depend on some measure of effective temperature or noise strength. Narrowing the discussion to Fokker-Planck models, the necessary and sufficient condition for the existence of such a ϕ appears to be the presence of the detailed balance with respect to a time reversal transformation [29], as was shown by Graham and Haken (GH) in 1971.

Consider a 2-point distribution function $\rho(\mathbf{r_2}, \mathbf{t_2}, \mathbf{r_1}, \mathbf{t_1})$. Under a time-reversal transformation, the variables $\mathbf{r}$ transform in some way:

$$t \rightarrow \tilde{t} = -t \tag{4.2}$$

$$\mathbf{r} \rightarrow \tilde{\mathbf{r}} \tag{4.3}$$

For example, when $\mathbf{r} = (\mathbf{x}, \mathbf{p})$, i.e. position and momentum, then $\tilde{\mathbf{x}} = \mathbf{x}$ and $\tilde{\mathbf{p}} = -\mathbf{p}$. Detailed balance means the following [2]:

$$\rho(\mathbf{r_2}, \mathbf{t_2}, \mathbf{r_1}, \mathbf{t_1}) = \rho(\tilde{\mathbf{r}}_1, \mathbf{t_2}, \tilde{\mathbf{r}}_2, \mathbf{t_1}). \tag{4.4}$$

For physical processes such as atomic collisions for example, the criteria for this condition come from the fundamental equations of motion. For us, the discussion is restricted to the cases when ρ evolves according to FPE (the two-point distribution function can be written as a product of a one-point distribution and a conditional transition probability distribution, both of which obey FPE). Consider a FPE of the form Eq. (2.41) with D replaced by a spatially dependent diffusion matrix $\mathbf{D}$ (assumed

[2]Due to the abundance of vectors and indices, vectors will be represented by boldface characters in this chapter.

symmetric) with components D_{ij} (GH also factor out 1/2 from it). According to GH, if detailed balance holds *and* we require the solution to the stationary FPE of the form $e^{-\phi(\mathbf{r})}$, then certain potential conditions will hold. To state these conditions, it helps to introduce a few definitions. First, let

$$\begin{aligned} \tilde{K}_i(\mathbf{r}) &= \epsilon_i K_i(\tilde{\mathbf{r}}(\mathbf{r})), & (4.5)\\ \tilde{D}_{ij}(\mathbf{r}) &= \epsilon_i \epsilon_j D_{ij}(\tilde{\mathbf{r}}(\mathbf{r})), & (4.6)\\ \epsilon_i &= \begin{cases} +1 \text{ if } \tilde{r}_i = r_i, \\ -1 \text{ if } \tilde{r}_i = -r_i \end{cases} & (4.7) \end{aligned}$$

(no summation over indices is implied). The $\tilde{\mathbf{K}}$ and $\tilde{\mathbf{D}}$ represent transformations to $\mathbf{K}$ and $\mathbf{D}$ under time reversal, just as $\tilde{\mathbf{r}}$ represents transformation to $\mathbf{r}$ under time reversal. Moreover, we may break the $\mathbf{K}$ into the part that changes sign under the transformation (the reversible part, $\mathbf{R}$) and the part that does not change sign (the irreversible part, $\mathbf{I}$): $\mathbf{K} = \mathbf{R} + \mathbf{I}$. Then

$$\begin{aligned} I_i &= \frac{1}{2}(K_i(\mathbf{r}) + \tilde{K}_i(\mathbf{r})), & (4.8)\\ R_i &= \frac{1}{2}(K_i(\mathbf{r}) - \tilde{K}_i(\mathbf{r})), & (4.9) \end{aligned}$$

Given these definitions, the first potential condition is a condition of the diffusion matrix and reads

$$D_{ik}(\mathbf{r}) = \tilde{D}_{ik}(\mathbf{r}). \tag{4.10}$$

The second condition dictated by the detailed balance says that the irreversible part of the probability current must be zero:

$$I_i \rho(\mathbf{r}) - \frac{\partial}{\partial r_k}\left(D_{ik}\rho(\mathbf{r})\right) = 0 \tag{4.11}$$

When $\rho = Ae^{-\phi}$ is substituted into this equation, we get the second potential condition:

$$\frac{\partial \phi}{\partial r_l} = D^{-1}_{li}\left(\frac{\partial D_{ik}}{\partial r_k} - I_i\right) \tag{4.12}$$

This serves as an equation for a smooth ϕ.[3] Finally, the third potential condition reads

$$0 = \frac{\partial R_i}{\partial r_i} - R_i \frac{\partial \phi}{\partial r_i}. \tag{4.13}$$

With the help or Eq. (4.1), this can be recast as

$$\frac{\partial (R_i \rho)}{\partial r_i} = 0, \tag{4.14}$$

which says that the reversible part of the drift must support the time-independent $\rho(\mathbf{r})$. All together these are conditions on the diffusion matrix and the drift field. The converse also holds true: if the potential conditions hold, then there is detailed balance *and* there exists a smooth ϕ obtained from Eq. (4.12) by quadrature. This is remarkable given that the system may actually be driven away from equilibrium![4]

We now derive an interesting consequence of detailed balance. To do this, it is helpful to first note the effect of generalizing a scalar D to a matrix on the Classical Mechanics described in Sections 2.1.3 and 2.1.2. Let $\mathbf{D} = D\mathbf{d}$ - we factored out the noise strength D, just a scalar control parameter. Then the Lagrangian becomes

$$\begin{aligned} L &= d^{-1}_{\mu\nu} \left(\dot{r}^\mu - K^\mu(\mathbf{r})\right) \left(\dot{r}^\nu - K^\nu(\mathbf{r})\right) && (4.15) \\ &= (\dot{\mathbf{r}} - \mathbf{K})^{\mathbf{T}} \mathbf{d}^{-\mathbf{1}} (\dot{\mathbf{r}} - \mathbf{K}), && (4.16) \end{aligned}$$

(now sum over repeated indices), the Hamiltonian in Eq. (2.48) generalizes to

$$\begin{aligned} \mathcal{H}(\mathbf{r}, \mathbf{p}) &= d_{\mu\nu} p^\mu p^\nu + K_\mu(\mathbf{r}) p^\mu && (4.17) \\ &= \mathbf{p}^{\mathbf{T}} \mathbf{d} \mathbf{p} + \mathbf{K}(\mathbf{r}) \mathbf{p} && (4.18) \end{aligned}$$

(recall that $\mathbf{r}$ stands for the vector (Q, P) in our problem). In the case of a unit matrix $\mathbf{d}$ we recover the old expressions for L and $\mathcal{H}$. The auxiliary equations of

[3] From Eq. (4.12) it follows that $\frac{\partial \phi}{\partial r_i \partial r_j} = \frac{\partial \phi}{\partial r_j \partial r_i}$.

[4] The potential conditions were known before GH, but it was their work [29] that explained the physical meaning behind these conditions to be detailed balance.

motion read

$$\dot{r}^k = \frac{\partial \mathcal{H}}{\partial p^k} = 2d_{k\mu}p^\mu + K_k(\mathbf{r}), \tag{4.19}$$

$$\dot{p}^k = -\frac{\partial \mathcal{H}}{\partial r^k} = -\frac{\partial d_{\mu\nu}}{\partial r_k}p^\mu p^\nu - \frac{\partial K^\mu}{\partial q_k}p_\mu, \tag{4.20}$$

and finally the momenta are related to the action S in the usual way: $\mathbf{p} = \nabla S$. The distribution is still given by

$$\rho(\mathbf{r}) \sim exp\left[-\frac{S(\mathbf{r})}{D}\right]. \tag{4.21}$$

We compare this to the definition of ϕ in Eq. (4.1) and conclude that $S = D\phi$, so $\nabla S = D\nabla\phi$. We now specialize to the case of spatially independent d. Then Eq. (4.12) tells us that $\mathbf{p}(\mathbf{r}) = -\mathbf{d}^{-1}\mathbf{I}(\mathbf{r})$. We substitute this into EOM for $\dot{\mathbf{r}}$, Eq. (4.19), and obtain

$$\dot{r}_k = -2I_k(\mathbf{r}) + K_k(\mathbf{r}). \tag{4.22}$$

In other words,

$$\dot{\mathbf{r}}_{FL} = -\tilde{\mathbf{K}}(\mathbf{r}). \tag{4.23}$$

The subscript "FL" is to remind that this equation describes a *fluctuational* trajectory. On the other hand, the *relaxational* trajectory obeys

$$\dot{\mathbf{r}}_{REL} = \mathbf{K}(\mathbf{r}). \tag{4.24}$$

Therefore, the flow field that describes the fluctuational trajectory (or escape trajectory) is the negative of time-reversed flow field that describes the relaxational trajectory! In our particular system, Eqs. (2.36 - 2.40), the diffusion matrix is $D_{ij} = D\delta_{ij}$ with constant D, and in the next section we will see explicitly how this conclusion holds true by calculating analytically the relaxational and the fluctuational trajectories when detailed balance holds.

The next logical question is, what are the consequences of violation of detailed balance? Specifically, what happens to the PDF $\rho(\mathbf{r})$? It is reasonable that in general $\rho(\mathbf{r})$ can still be written down in a Boltzmann-like form; we see this from the path-integral solution to the FPE: unless the fluctuations around the optimal (classical) path contribute to the exponent, the distribution will still have the form $\rho(\mathbf{r}) = \rho'(\mathbf{r}) \exp[-S(\mathbf{r})/D]$. However, the prefactor may now depend on $\mathbf{r}$ and it may even be singular [82] and the action S may in general be a multivalued function of $\mathbf{r}$ if there exists more then one escape trajectory reaching this $\mathbf{r}$ from $\mathbf{r_0}$. Generically, these trajectories will yield a different value of the action S_1 and S_2, and in the limit of small noise $\exp[-S_1/D]$ may drastically differ from $\exp[-S_2/D]$ although the potential becomes smoothed out at a finite noise strength [30]. Thus, in the limit of vanishing noise strength one must always choose that trajectory which leads to the least value of the action, a procedure that may lead to a nonsmoothness (or nondifferentiability) of ϕ with respect to $\mathbf{r}$. This effect was described by several authors; Graham and Tél in [30] appear to be one of the earlier authors to describe this particular mechanism of formation of nondifferentiability of ϕ. They point out a crucial geometric fact of the auxiliary dynamical system: in order for a projection of two trajectories in the $\mathbf{r} - \mathbf{p}$ phase space to meet at one point in the $\mathbf{r}$ space, the Lagrangian manifolds on which these trajectories live must "fold over": the trajectories cannot intersect on the manifold, in virtue of the existence and uniqueness of solutions of dynamical systems, so the only way for their projections unto the $\mathbf{r}$ plane to intersect is for the manifold to fold over. Dykman, Millonas and Smelyanskiy further clarify the geometric details of this in [31] (one is highly encouraged to have a look at Fig. 2 in this work for a cartoon explanation of the relationship between folds, caustics, switching lines and trajectories reflecting off of caustics). Relating to Catastrophe theoretic classification of folds in Lagrangian Manifolds (Appendix 12 of [76] and [83, 84]), they note that such foliation *generically* happens in a cusp manner, also known as Whitney's tuck. Because the intersections of neighboring trajectories happen at the folds, the projections of the folds onto the $\mathbf{r}$ space map out a pair of caustics (see [31] or Maier and Stein's

paper [82]).[5] Dykman, Millonas and Smelyanskiy also describe a generic form of the caustic near a saddle fixed point of the deterministic system (i.e. of $\mathbf{K}$). Although the caustics are not experimentally observable, the "switching line," which is a curve at which $S_1 = S_2$ [30, 31, 85, 86] is the feature which should be observable in experiments (see references in [31]). The nondifferentiability of $S(\mathbf{r})$ takes place along switching lines.

In the 1980s Graham and Tél undertook the study of caustics in various models and made an attempt to place conditions on the structure of the FPE which would entail smooth potentials (assuming detailed balance is broken). Their search for such conditions centered on the search for integrability criteria [32], although they state that integrability is not a sufficient condition, because it does not prevent multivaluedness. Yet, integrability in their point of view was sufficient to prevent infinite foliation of Lagrangian manifolds, a phenomenon also known as wild separatricies [30, 87]. A general condition that ensures integrability of the auxiliary system associated with the FPE Hamiltonian $\mathcal{H}$ appears elusive. In the words of Graham and Tél [30], "methods testing for integrability on certain hypersurfaces in phase space, like $E = 0$, would be better suited for the problem at hand, then tests for integrability in the entire phase space, but methods of this type remain to be developed".

Most of the efforts to describe caustics appear to rely upon the existence of a fixed point or a limit cycle. Among more recent works, the formation of caustics in the presence of a saddle fixed point has been investigated by Maier and Stein in [82]. This work studies conditions when the caustic is attached to the saddle and when it breaks off as a control parameter in the model is varied. The case when the system contains a limit cycle has also been understood in great detail by [88] for example. All of Graham's works cited here appear to focus on models with limit cycles as well. To my knowledge, a work by Dykman, Millonas and Smelianskiy [31] appears to be the only paper in which the situation without any limit sets (aside from the attractor from which the trajectories escape) is considered. The system they describe happens to be the Duffing oscillator in the *monostable* regime (they also consider the bistable

[5] In mechanics or ray optics, a caustic is an envelope of intersecting trajectories or rays.

regime where there is a saddle; they study conditions for detachment of the caustic from the saddle). Indeed, as described in Chapter 1, for $\mathcal{F} < \mathcal{F}_l$ there exists only one single attracting fixed point. They write down the normal form for $\rho(\mathbf{r})$ in the vicinity of the cusp point - the point where the pairs of caustics meet. What is not discussed in this work is the evolution of the caustic system as the driving is turned up, and most importantly, the mechanism of the formation of the caustic pair. The questions are, is there a finite threshold $\mathcal{F}$ at which the caustic appear and what serves to nucleate the caustics as $\mathcal{F}$ is turned up. We answer these questions here.

We make several surprising discoveries: (i) the caustics do exist; they appear in pairs as expected on the grounds of general catastrophe theory [31, 83, 84] and appendix 12 of [76]; (ii) there does not appear to be a threshold $\mathcal{F}$ for the appearance of a caustic; (iii) the radial distance R (not to be confused with $\mathbf{R}$ above) to the cusp point as a function of $\mathcal{F}$ appears to be a power law: $R \sim \mathcal{F}^{-p}$. Thus, as $\mathcal{F}$ is turned on from zero the caustic appears at infinity and is brought to smaller radii for larger $\mathcal{F}$. We now present these findings, which are mostly numerical, and describe our (so far largely unfruitful) attempts to explain them analytically.

4.2 Fragility of Detailed Balance and Formation of Caustics in the Monostable Regime of the Oscillator

Before moving on, it is instructive to check if and when our system represented by Eqs. (2.36 -2.40) satisfies detailed balance [i.e., satisfies the potential conditions and admits a smooth $\phi(\mathbf{r})$ or $S(\mathbf{r})$]. First, note that the variables of the Duffing oscillator have the following obvious time-reversal properties: $\tilde{x} = x$ and $\dot{\tilde{x}} = -x$. According to the definition of Q and P, Eqs. (2.28)-(2.29), these will have the following time-reversal properties: $\tilde{P} = -P$ and $\tilde{Q} = Q$. Then, using the definition of K_Q and K_P, Eqs. (2.26)-(2.27) and the definitions in Eqs. (4.5)-(4.8) we get

$$
\begin{aligned}
I_Q &= -\eta Q, && (4.25)\\
I_P &= -\eta P, && (4.26)\\
R_Q &= -P + P(Q^2 + P^2), && (4.27)\\
R_P &= Q - Q(Q^2 + P^2) + \mathcal{F}. && (4.28)
\end{aligned}
$$

Given this, we construct the potential conditions. The first potential condition, Eq. (4.10), is satisfied automatically. The second potential condition, Eq. (4.12) says

$$
\begin{aligned}
\frac{\partial \phi}{\partial Q} &= -\frac{1}{D} I_Q = \frac{\eta Q}{D}, && (4.29)\\
\frac{\partial \phi}{\partial P} &= -\frac{1}{D} I_P = \frac{\eta P}{D}. && (4.30)
\end{aligned}
$$

This is indeed an equation for ϕ with the simple solution

$$
\phi = \frac{\eta}{2D}(Q^2 + P^2). \tag{4.31}
$$

And the third potential condition expressed in Eq. (4.13) says

$$
\begin{aligned}
0 &= \frac{\partial R_Q}{\partial Q} - R_Q \frac{\partial \phi}{\partial Q} = 2PQ + \frac{\eta PQ}{D}\left[1 - (Q^2 + P^2)\right], && (4.32)\\
0 &= \frac{\partial R_P}{\partial P} - R_P \frac{\partial \phi}{\partial P} = -2PQ - \frac{\eta PQ}{D}\left[1 - (Q^2 + P^2)\right] - \frac{\eta \mathcal{F} P}{D}. && (4.33)
\end{aligned}
$$

Adding the two equations we get

$$
\frac{\eta \mathcal{F} P}{D} = 0. \tag{4.34}
$$

Thus, detailed balance holds only when $\mathcal{F} = 0$, in other words the detailed balance is fragile!

The auxiliary EOM were explicitly stated in Sec. 2.1.3 [Eqs. (2.51)-(2.54)]. Some-

times, it will prove useful to work in polar coordinates, so in terms of the radial coordinate r and the angular coordinate θ, the Lagrangian is

$$L = \frac{1}{4}\left[(\dot{r}+\eta r)^2 + \left(r\dot{\theta} - r + r^3\right)^2\right] - \frac{\mathcal{F}}{2}\left[\left(r\dot{\theta} - r + r^3\right)\cos\theta + (\dot{r}+\eta r)\sin\theta\right] + \frac{\mathcal{F}^2}{4},$$

the Hamiltonian is

$$\mathcal{H} = p_r^2 + \frac{p_\theta^2}{r^2} - \eta r p_r + p_\theta(1-r^2) + \mathcal{F}\left(p_r\sin\phi + \frac{p_\theta}{r}\cos\theta\right) + \frac{\mathcal{F}^2}{2}, \tag{4.35}$$

and EOM are

$$\begin{aligned}
\dot{r} &= 2p_r - \eta r + \mathcal{F}\sin\theta, && (4.36)\\
\dot{\theta} &= \frac{2p_\theta}{r^2} + (1-r^2) + \frac{\mathcal{F}}{r}\cos\theta, && (4.37)\\
\dot{p}_r &= \frac{2p_\theta^2}{r^3} + \eta p_r + 2rp_\theta + \mathcal{F}\frac{p_\theta}{r^2}\cos\theta, && (4.38)\\
\dot{p}_\theta &= -\mathcal{F}\left(p_r\cos\theta - \frac{p_\theta}{r}\sin\theta\right). && (4.39)
\end{aligned}$$

4.2.1 Exact Solution at Zero $\mathcal{F}$

We can exploit the angular symmetry of the situation: when $\mathcal{F} = 0$, the Hamiltonian g is rotationally symmetric, $g = \frac{1}{4}(Q^2+P^2-1)^2$, so the 2D flow is symmetric and the set of escape trajectories must also be symmetric![6] This is where polar coordinates are most useful. Because of this symmetry, $p_\theta = const$: this is seen directly from Eq. (4.39) and represents conservation of angular momentum, but this *const* must be 0 due to the fact that the trajectory emanates from the fixed point. Because of this, and the fact that the motion takes place on the $\mathcal{H} = 0$ surface, Eq. (4.35) reveals that $p_r = 0$ (relaxational paths) or $p_r = \eta r$ (fluctuational paths). Now, $\frac{\partial S}{\partial q_i} = p_i$, so

[6] To remind the reader, this refers to the most likely trajectories along which the fluctuation from the attracting fixed point to a point (Q, P) takes place.

the action is $S(r) = \eta\frac{r^2}{2} = \eta\frac{Q^2+P^2}{2}$, and $p_Q = \eta Q$, $p_P = \eta P$. Therefore,

$$\dot{Q} = \eta Q + \frac{\partial g}{\partial P}, \tag{4.40}$$

$$\dot{P} = \eta P - \frac{\partial g}{\partial Q}. \tag{4.41}$$

This is consistent the prediction we made earlier from general considerations in Sec. 4.1. When our nonlinear oscillator satisfies detailed balance, the escape trajectories are obtained from relaxational trajectories with a *reversed sign of friction* [indeed, this is $\tilde{\mathbf{K}}(\mathbf{r})$, Eq. (4.23), for this particular system]. We also learn that the distribution is given (aside from a normalization factor) by

$$\rho(Q,P) \propto \exp\left(-\frac{\eta}{2D}(Q^2+P^2)\right), \tag{4.42}$$

which we already derived (Eq. (4.31)) via the more general method.

For the record, I write out the zeroth-order ($\mathcal{F} = 0$) solution to Eqs. (4.36)-(4.37) explicitly because we will need it below. It satisfies $\dot{r} = \eta r$ and $\dot{\theta} = 1 - r^2$ (remember that $p_\theta = 0$ for $\mathcal{F} = 0$).

$$\theta(r) = \theta_0 + \frac{1}{\eta}\left[\ln\left(\frac{r}{r_0}\right) - \frac{r^2 - r_0^2}{2}\right] \tag{4.43}$$

or in terms of time,[7]

$$r(t) = r_0 e^{\eta t}, \tag{4.44}$$

$$\theta(t) = \theta_0 + t - \frac{r_0^2}{2\eta}\left(e^{2\eta t} - 1\right). \tag{4.45}$$

[7] I switch notation from T to t for now, and it will be understood that t is not the original time in the Duffing equation, but the slow time in the amplitude equations, formerly known as T.

Or

$$\begin{aligned} Q(t) &= e^{\eta t}\left[Q_0\cos\psi(t) - P_0\sin\psi(t)\right], && (4.46)\\ P(t) &= e^{\eta t}\left[Q_0\sin\psi(t) + P_0\cos\psi(t)\right], && (4.47)\\ \text{where } \psi(t) &= t - \frac{Q_0^2+P_0^2}{2\eta}\left[e^{2\eta t}-1\right]. && (4.48) \end{aligned}$$

The solution is designed to arrive to point (Q_0, P_0) at time $t = 0$. We will want to think of the potential ϕ as a function of the field point (Q_0, P_0) and then drop the subscript.

4.2.2 An Aside

We can actually extract the result given by Eq. 4.42 from more physical arguments directly for the Duffing oscillator - this will underline the usefullness of potential conditions discussed in prior sections. In the absence of driving, the Duffing oscillator is descried by $\ddot{q} + 2\Gamma\dot{q} + \omega_0^2 q + \gamma q^3 = f(t)$, Eq. (2.1). Clearly, in thermodynamic equilibrium at temperature T (not to be confused with the slow time), q and $\dot{q}$ obey the Boltzmann distribution

$$\rho(q,\dot{q}) \propto \exp\left(-\frac{\frac{\dot{q}^2}{2} + \omega_0^2\frac{q^2}{2} + \gamma\frac{q^4}{4}}{kT}\right). \tag{4.49}$$

Now we can switch to P and Q via the transformation in Eqs. (2.28)-(2.29). The $\frac{\dot{q}^2}{2} + \omega_0^2\frac{q^2}{2}$ part gives

$$\frac{\eta}{2} \times \frac{8}{3\eta}\frac{\omega_0^3}{\gamma}(\omega_F - \omega_0)\left[P^2 + Q^2\right] + O\left(\frac{\epsilon}{2\eta}\right), \tag{4.50}$$

where the $O\left(\frac{\epsilon}{2\eta}\right)$ term comes form the fact that $\dot{q}$ contains ω_F in its definition (Eq. (2.28)) while q^2 has an ω_0^2 in front of it. The γq^4 term is of the form

$$\frac{\eta}{2} \times \frac{2}{\eta}\frac{\omega_0^2}{\gamma}(\omega_F - \omega_0)^2 \times O(P^4, Q^4, ...). \tag{4.51}$$

Thus (quartic term)/(quadratic term) is of order $\frac{\omega_F-\omega_0}{\gamma\omega_0} = \frac{\epsilon}{2\eta}$. Thus,

$$\rho(Q,P) \propto \exp\left(-\frac{\eta}{2D'}(Q^2+P^2) + O\left(\frac{\epsilon}{2\eta}\right)\right), \tag{4.52}$$

where

$$D' = \frac{3\gamma}{16\omega_0^3}\frac{1}{(\omega_F-\omega_0^2)} \times (2\Gamma kT). \tag{4.53}$$

This equals D if $B = 2\Gamma kT$ which is just an instance of a Fluctuation-Dissipation relationship. The $O\left(\frac{\epsilon}{2\eta}\right)$ correction comes from the fact that the displacement is assumed small and therefore q samples mostly the linear part of the potential. It may appear that because of this $O\left(\frac{\epsilon}{2\eta}\right)$, the ρ obtained with this method does not correspond exactly to the ρ obtained above, which did not mention $O\left(\frac{\epsilon}{2\eta}\right)$. However, that derivation of ρ was based upon the slow amplitude equations, Eqs. (2.26)-(2.27) which were obtained under the assumption that displacement q is small and samples mostly the linear parts of the Duffing potential. In other words, in the amplitude equations the ϵ-corrections have also been neglected.

What we have just seen is that if a Boltzmann distribution holds in the phase space of the oscillator, then it also holds in the canonically transformed phase space. From this point of view, it is clear why the new distribution will not remain given by strictly $\exp\left[-\frac{\eta}{2D}\left(Q^2+P^2\right)\right]$: this is because the Boltzmann distribution in phase space does not strictly apply as soon as the time-dependent driving is introduced. By definition, the problem becomes nonequilibrium since the notion of full exploration of phase space does not apply: the system does not have enough time to explore the phase space as the applied field changes it. But what is the picture from the point of view of time-independent amplitude equations, Eqs. (2.26)-(2.27)? This is where the ideas from Sec. 4.1 come in. Suppose we are just given the Fokker-Planck equation without realizing that it came from a Duffing oscillator. The FPE is autonomous, so we can not use time dependence to say that the Boltzmann distribution does not apply, but we can use potential conditions as a "litmus test" for detailed balance and hence for the possibility of a smooth $\phi(Q,P)$. As we just saw, the potential conditions

fail at nonzero $\mathcal{F}$. We will now see how $\phi(Q,P)$ fails to remain smooth in practice. The subject of Sec. 4.2.3 and Sec. 4.2.4 will appear in the upcoming publication [89].

4.2.3 Phenomenology of Caustics in the Monostable Regime

In this subsection we will study the pattern of escape trajectories at ever increasing driving strength $\mathcal{F}$. We recall that the exponential part of the distribution $\rho(Q,P)$ is given by the action of trajectories along which the system is most likely to escape the attracting fixed point (or attractor) to the point (Q,P). These are the "escape trajectories" which form the subject of discussion in this section. The word "escape" refers to the escape from the attractor to (Q,P), not to be confused with the escape from a basin which was relevant in Chapter 3. We also recall that one may refer to these as "fluctuational" trajectories, while the trajectories along which the system is most likely to return from (Q,P), which are just the trajectories of the noise-free dynamics of the oscillator, may be called "relaxational" trajectories.

Unless otherwise stated, the phenomenology will be demonstrated at fixed damping $\eta = 0.4$. When $\mathcal{F} = 0$, the pattern is rotationally symmetric, Fig. 4.1. This fact was proven above in terms of the symmetry of $\rho(Q,P)$, Eq. (4.42), and in terms of the explicit analytical form of the trajectories, Eq. (4.43). All calculations of trajectories in this section are terminated either when a stopping radius is reached or when a trajectory reaches a caustic, as explained shortly. The zeroth-order solution, Eqs. (4.44)-(4.45) suggests that trajectories are accelerating exponentially, which means that to provide an accurate solution valid at larger radii, the time steps must decrease dramatically. The choice of the stopping radius was dictated mostly by the position of interesting features (caustics). As we will see, these interesting features recede to larger radii as $\mathcal{F}$ decreases, and so calculations at very low $\mathcal{F}$ required very small time steps (this would be an ideal place for a variable time-stepping algorithm, but we used a 4th order Runge-Kutta). For a given stopping radius, the time steps were decreased until the results did not change significantly. Then more extensive calculations were performed at this saturated value of time steps.

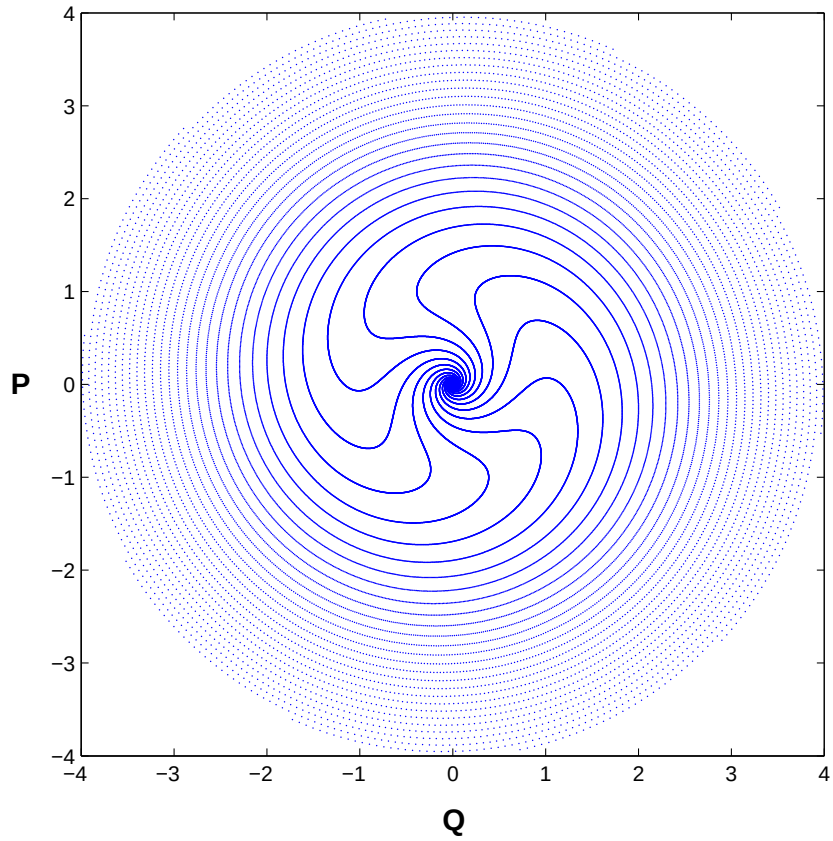


Figure 4.1: The pattern of 10 escape trajectories at $\mathcal{F} = 0$. Each trajectory undergoes approximately 3 revolutions after reversing its direction of rotation at the slow-down radius of 1.

Fig. 4.2 displays the trajectory pattern at $\mathcal{F} = 0.005$, clearly showing the breaking of the rotational symmetry: the density of trajectories becomes nonuniform, and some of them start to "jam" into each other. As $\mathcal{F}$ is increased further, the radius at which this jamming takes place decreases. One way to find out where this happens exactly is to consider pairs of trajectories with a similar neighboring parameter and to observe the position of their intersection. There is a more elegant and powerful way. In our discussion of why projections of trajectories intersect, we mentioned that the reason for this phenomenon is the folding of Lagrangian manifolds in the (Q, P, p_Q, p_P) space. As two trajectories with an infinitesimally different trajectory parameter pass over this fold, their projections will intersect at the fold; excellent cartoons of this can be found in [31] and also in the beginning of [19]. The Lagrangian manifold is the surface

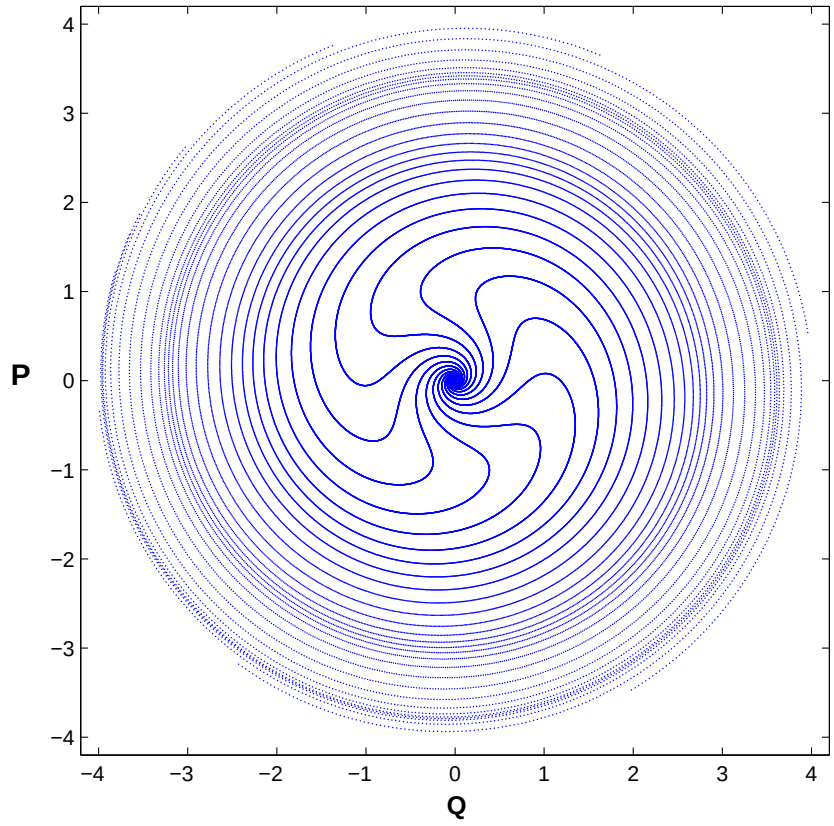


Figure 4.2: The pattern of 10 escape trajectories at $\mathcal{F} = 0.005$.

$\mathbf{p}(\mathbf{r})$, i.e., $(p_Q(Q,P), p_P(Q,P))$. The fold is characterized by such location where the tangent plane to the surface is perpendicular to the (Q,P) plane. This happens when

$$\frac{\partial(Q,P)}{\partial(p_Q,p_P)} = 0, \tag{4.54}$$

in other words, when the 1-to-1 correspondence between $\Delta\mathbf{p}$ and $\Delta\mathbf{r}$ is lost. Alternatively, we can say that

$$J = \frac{\partial(p_Q,p_P)}{\partial(Q,P)} = \infty. \tag{4.55}$$

Thus, knowing $\mathbf{p}(\mathbf{r})$, the caustics can be mapped out at locations where J becomes infinity. In general, the web of caustics may be very complicated, even of infinite complexity in the case of "wild separatricies" [87, 88].[8] The function $\mathbf{p}(\mathbf{r})$ is not

[8]This type of behavior has also been found in our system in the bistable regime. It can be seen in the pattern of trajectories that hit the separatrix: as narrower and narrower ranges of trajectory

known a priori, it is something that we are in fact seeking; only p along a given trajectory can be calculated. However, as the equations of motion for p_Q, p_P, Q and P are integrated in time, it is also possible to simultaneously integrate the equations of motion for second derivatives such as $\frac{\partial^2 S}{\partial Q \partial P} = \frac{\partial p_Q}{\partial P}$, etc. These are called the time-dependent Riccati equations, which are derived and stated by Maier and Stein [74].

$$\dot{S}_{ij} = -\frac{\partial^2 \mathcal{H}}{\partial p_k \partial p_l} S_{ik} S_{jl} - \frac{\partial^2 \mathcal{H}}{\partial r_j \partial p_k} S_{ik} - \frac{\partial^2 \mathcal{H}}{\partial r_i \partial p_k} S_{jk} - \frac{\partial^2 \mathcal{H}}{\partial r_i \partial r_j} \tag{4.56}$$

These equations are not only of practical use; they have a rather deep differential-geometric interpretation [74], which is outside the scope of this thesis. With the help of these and the other four EOM we can track the evolution of J: all together, there are 7 coupled ODEs since the ODE for mixed second partial derivatives of S are the same. The initial conditions are obtained from the knowledge of the surface $\mathbf{p}(\mathbf{r})$ in the vicinity of the attractor, for example, from the knowledge of the eigenvectors.

We now proceed with the demonstration of the effects of increasing the strength of driving. At $\mathcal{F} = 0.008$ the region of increased density of trajectories move in a little bit closer to the origin. This is shown in Fig. 4.3 along with the portrait of the caustic. We see the cusp structure, which is expected on the generic basis (see [31], the Appendix 12 of [76] and [83, 84]).

Another word of technicality is in order. Although caustics are projections of folds in the Lagrangian manifolds unto the (Q, P) plane, the dynamics *on* the manifolds is obviously not trivial. Two trajectories with different shooting parameters may hit a caustic after going through very different dynamics; it is not obvious what each of these two trajectories do before hitting a caustic, if it is destined to do so. In particular, a stopping radius which is not large enough may prevent some trajectories from reaching a caustic. In Fig. 4.3 for example, the portion of the caustic shown lies around the radius of 4, but in mapping out this caustic, we evolved trajectories choosing the stopping radius to be 6.

As $\mathcal{F}$ is increased further, the cusp structure continues to pull in closer to the

parameters are considered, the pattern repeats itself. This fractal structure is believed to be caused by the proliferation of folds, one on top of the other in the vicinity of the saddle.

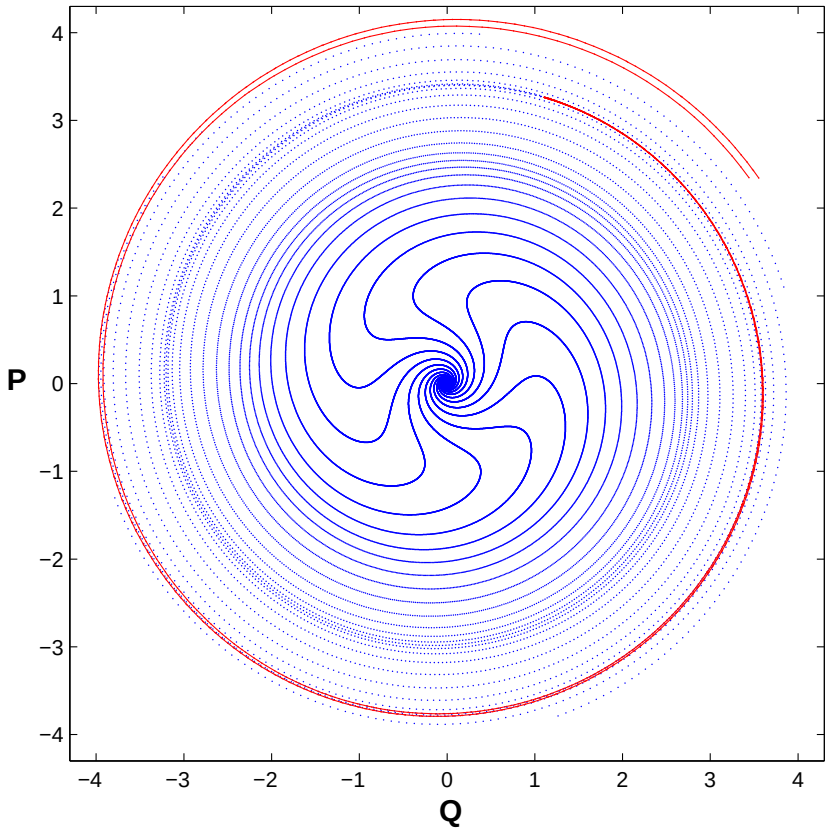


Figure 4.3: The pattern of 10 escape trajectories at $\mathcal{F} = 0.008$ and the pair of caustics originating at the cusp.

origin. Some caustic and trajectory patterns at larger $\mathcal{F}$ will be shown below. For now we summarize the relationship between the radius at which the cusp is located versus $\mathcal{F}$ for two values of η. This is shown in Fig. 4.4.

Two power laws $R \sim \mathcal{F}^p$ with a crossover appear. The nature of the crossover is not obvious, but the discussion below may hint that it may be related to the appearance of a different caustic around the attractor. For $\eta = 0.4$, $p \approx -0.502$ at very low $\mathcal{F}$ and $p \approx -0.386$ at larger $\mathcal{F}$. The second power law appears to hold rather well for $\mathcal{F}$ as large as 0.39 (not shown).[9] For $\eta = 0.1$, the two values of $p \approx -0.456$ at very low $\mathcal{F}$ and $p \approx -0.242$ at larger $\mathcal{F}$. We also observe that the power laws at lower $\mathcal{F}$ are less sensitive to η then the power laws that hold after the crossover, at

[9]If the distance is measured not from the origin, but from the attractor, which does not sit precisely at the origin, the points at larger $\mathcal{F}$ fall off the linear trend.

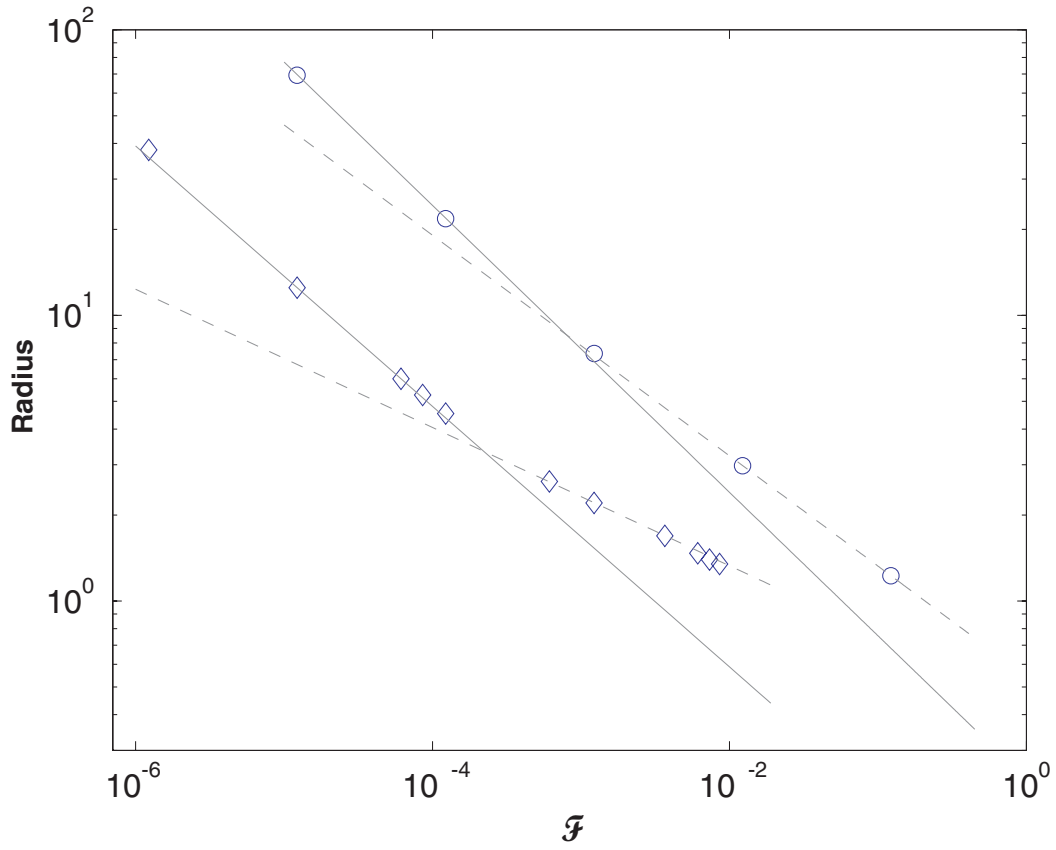


Figure 4.4: Log-log plot representing the radius of the cusp, measured from the origin versus the strength of the driving $\mathcal{F}$ for $\eta = 0.4$ circles and $\eta = 0.1$ diamonds. The solid and dashed lines represent fits to the two power-law regimes.

larger $\mathcal{F}$.

As $\mathcal{F}$ increases, the caustics also evolve to become more complex. Fig. 4.5 conveys a sense of this evolution with increasing $\mathcal{F}$. In each case, the caustics were mapped out by evolving 5000 trajectories with the stoping radius of 6. At $\mathcal{F} = 0.05$ we notice the appearance of the second caustic around the attractor. Its tight spiral structure is seen in greater more detail in panel (b). There are also occasional loose red dots, indicating that in addition to the familiar cusp-type caustic and the new spiral caustic around the origin, there are other, sparsely visited pieces of caustics. At higher $\mathcal{F}$ the spiral caustic around the attractor appears to be the continuation of the inner branch of the familiar pair of caustics, panels (c) and (d). Finally, the last two panels, (e) and (f), depict the situation for $\mathcal{F} = 0.395$, which correspond to the bistable regime:

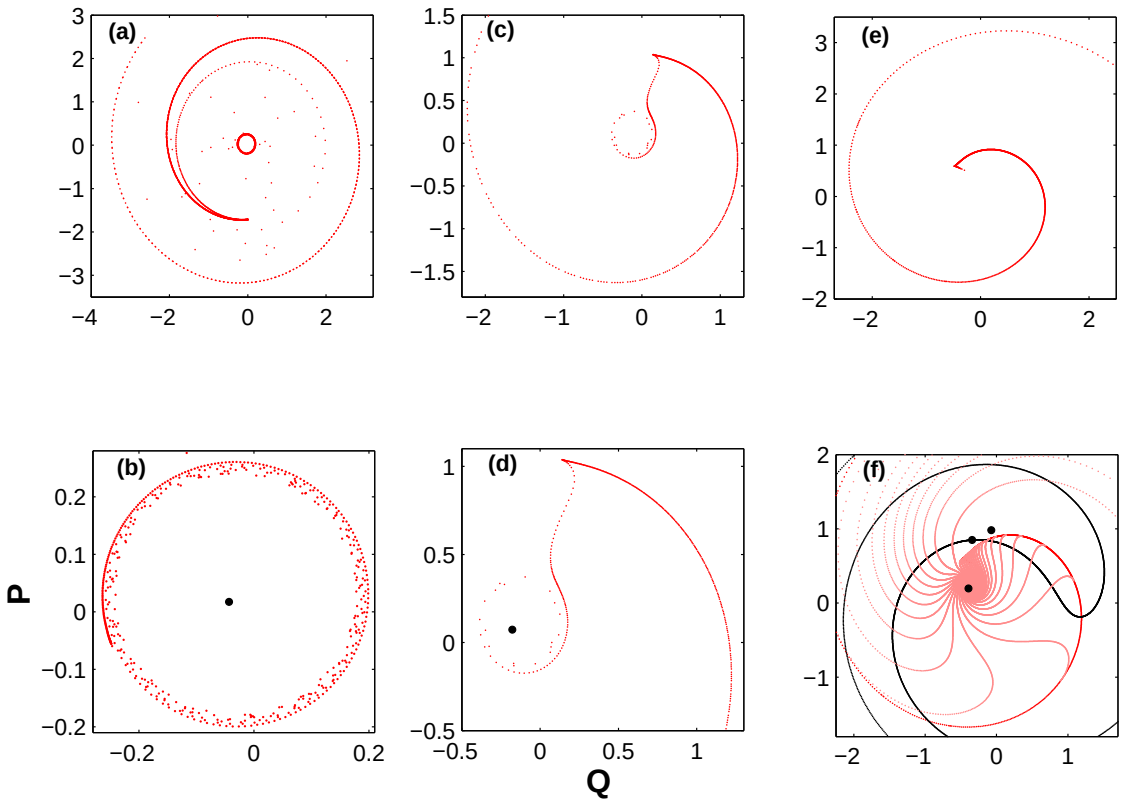


Figure 4.5: Caustics systems at various $\mathcal{F}$. Panels (a) and (b): $\mathcal{F} = 0.05$, panels (c) and (d): $\mathcal{F} = 0.2$, and panels (e) and (f): $\mathcal{F} = .395$, all at $\eta = 0.4$. In panel (f) we also show a set of escape trajectories from one of the attracting fixed points, one for a different trajectory parameter. The fixed points and the separatrix are also shown, with the saddle laying on the separatrix.

two more fixed points, a saddle (in the middle) and another attractor appear. We observe that the familiar caustic does not seem to be connected to the saddle.[10] The phenomenology of the caustic structure and of the pattern of trajectories can be come quite complex at larger $\mathcal{F}$. One may ask, for example, what types of trajectories lead to the circular caustic around the origin and whether these ever interact with the cusp-type caustic that we discussed. Indeed, there is plenty of room for future work to study this phenomenology. On the other hand, our goal here was to understand the features which are expected to be generic, such as the scaling properties of the caustics as $\mathcal{F} \to 0$ and to answer the question whether there is a finite threshold $\mathcal{F}$

[10]Indeed, the ratio of repulsive egenvalues of the saddle is ≈ 9.9 here, so in light of [31], the caustic is not expected to be attached to the saddle for this particular choice of η and $\mathcal{F}$.

at which the caustics appear. This is the main thrust of numerical Sec. 4.2.3 and the analytical Sec. 4.2.4 that follows.

Varying η at fixed $\mathcal{F}$ also reveals a scaling which appears to develop asymptotically in η. This is shown in Fig. 4.6. For the particular value of $\mathcal{F}$ at which the numerical calculations have been performed, $\mathcal{F} \approx 0.0012$, the relationship of R versus η appears to become asymptotic to a power law for larger η. To be more precise the relationship appears to be

$$R = 1 + A\eta^{\tilde{p}} \tag{4.57}$$

where the exponent $\tilde{p} \approx 1.26$ when $\mathcal{F} = 0.0012$. Moreover, since $\eta = \frac{\Gamma}{\delta\omega}$ [where $\delta\omega$ is the deviation of the driving frequency from the resonant frequency, see Eq. (2.30)], we also learn that the relationship R versus $\delta\omega$ is asymptotic to a power law at low $\delta\omega$ (i.e. close to resonance) with the exponent $-\tilde{p}$.

Visual inspection of figures such as Fig. 4.3 leads us to notice that an individual trajectory by itself differs qualitatively very little from a trajectory at $\mathcal{F} = 0$. In other words, we hypothesize that a cusp-like caustic happens as a result of a *nonuniform* perturbation of the zeroth-order trajectories seen in Fig. 4.1, even though the actual perturbation of *each* trajectory is small (meaning, for example that $\delta r(t) \ll r^{(0)}(t)$ at a given t). If so, we may be able to capture the phenomenon with a perturbation theory. Furthermore, it is reasonable to suggest that the perturbation takes place mostly around $r = 1$ where trajectories slow down because the symmetry-breaking component of velocity that appears at nonzero $\mathcal{F}$ is most important there,[11] as the nominal velocity is smallest at $r = 1$. To test this hypothesis we consider a different system by reversing the sign in the original hamiltonian g [see Eq. (2.40)]: instead of $\frac{1}{4}\left(Q^2 + P^2 - 1\right)^4$ we consider the system with $\frac{1}{4}\left(Q^2 + P^2 + 1\right)^4$. This eliminates the slowing-down effect. The escape trajectories for this system for $\eta = 0.2$ and $\mathcal{F} = 0.05$ are displayed in Fig. 4.7. We find that caustics can be found in this system as well. Hence, the slowing-down region around $r = 1$ is not crucial to the formation of caustics. In the next subsection we will demonstrate analytically that indeed,

[11] In the absence of damping, there is a circle of stagnation at $r = 1$, while the damping term introduces a gradient towards the origin.

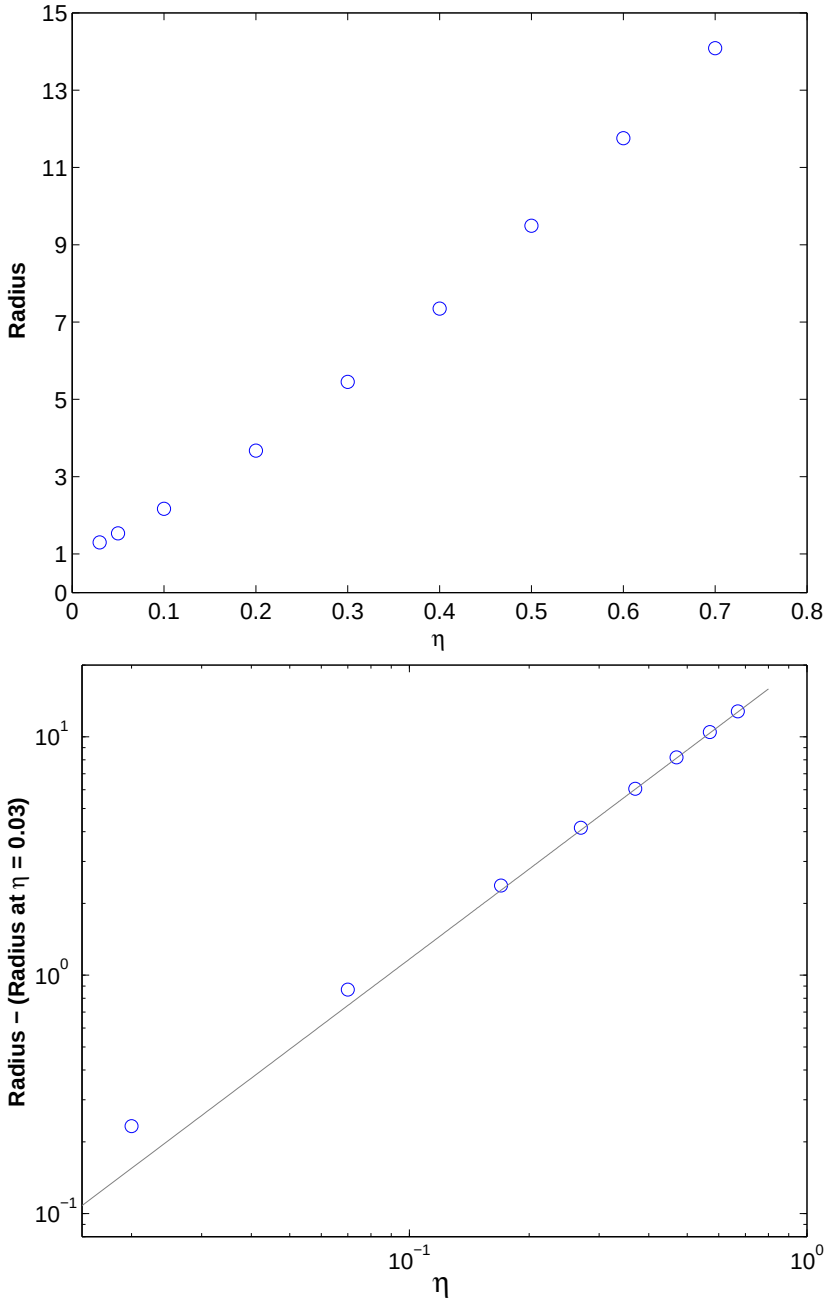


Figure 4.6: Radius of the cusp versus η for a fixed $\mathcal{F} \approx 0.0012$. The asymptotic exponent $\tilde{p}$ in $R = 1 + A\eta^{\tilde{p}}$ is ≈ 1.26.

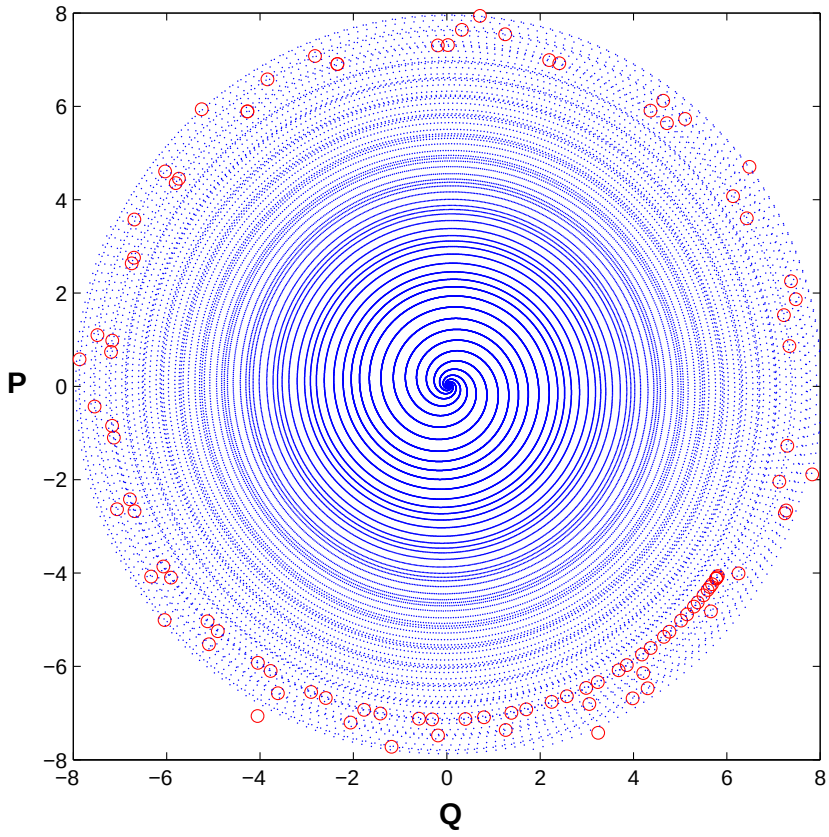


Figure 4.7: A portion of the caustic in the system based on $g(Q,P) = \frac{1}{4}\left(Q^2+P^2+1\right)^4$, indicated by open circles and a sample of 5 trajectories. Note the absence of the turn-around at $r = 1$.

the slowing-down region is not responsible, and caustics in fact, arrive from infinity. It appears that caustics form as a result of the non-uniform perturbation of each trajectory at large distances.

4.2.4 Analytical Predictions

One way to calculate trajectories is by a straightforward perturbation theory: let $Q(t) = Q^{(0)}(t) + \mathcal{F}Q^{(1)}(t)$, $P(t) = P^{(0)}(t) + \mathcal{F}P^{(1)}(t)$, and same for the momenta p_Q and p_P, and substitute into the auxiliary EOM, Eqs. (2.51)-(2.54). There will be four linear equations for first-order corrections $Q^{(1)}$, etc., but they will have time-dependent coefficients. We can partly ameliorate this inconvenience by eliminating the need for $\dot{p}_Q$ and $\dot{p}_P$ equations with the help of a separate perturbative analysis of

the action $S(Q_0, P_0)$.[12] This will hand us perturbed momenta as a function of Q and P, which can then be substituted in the remaining ODEs for $Q^{(1)}$ and $P^{(1)}$.

We would like to have an expression for the action at a point (Q_0, P_0) as a series in $\mathcal{F}$, i.e., $S(Q_0, P_0) = S^{(0)}(Q_0, P_0) + \mathcal{F}S^{(1)}(Q_0, P_0) + ...$. Consider a trajectory that leaves the attractor and arrives to a point (Q_0, P_0); for simplicity of notation denote this trajectory by $\xi(t)$. Then the action is given by

$$S(Q_0, P_0) = \int_{attr}^{(Q_0,P_0)} \mathcal{L}\left(\xi(t'), \dot{\xi}(t')\right) \, dt'. \tag{4.58}$$

Of course both $\mathcal{L}$ and the trajectory itself vary with $\mathcal{F}$, so we should expand both around $\mathcal{F} = 0$,

$$\begin{aligned} \mathcal{L}\left(\xi(t), \dot{\xi}(t)\right) &= \mathcal{L}^{(0)}\left(\xi^{(0)}, \dot{\xi}^{(0)}\right) + \mathcal{F}\left[\left(\frac{\partial \mathcal{L}^{(0)}}{\partial \xi}\right)_{(\xi^{(0)}, \dot{\xi}^{(0)})} \xi^{(1)} + \left(\frac{\partial \mathcal{L}^{(0)}}{\partial \dot{\xi}}\right)_{(\xi^{(0)}, \dot{\xi}^{(0)})} \dot{\xi}^{(1)}\right] \\ &+ \mathcal{L}^{(1)}\left(\xi^{(0)}, \dot{\xi}^{(0)}\right) + O\left(\mathcal{F}^2\right). \end{aligned} \tag{4.59}$$

When integrated, the first term will yield $S^{(0)}(Q_0, P_0)$ and the term in brackets will give zero by definition of $\mathcal{L}^{(0)}$: with integration by parts it can be seen as the integral of the Euler-Lagrange equations based on $\mathcal{L}^{(0)}$ along the trajectory which is a solution to these equations. Hence,

$$\begin{aligned} S(Q_0, P_0) &= \frac{\eta}{2}\left(Q_0^2 + P_0^2\right) + \mathcal{F}\int_{-\infty}^{t} \mathcal{L}^{(1)}\left(\xi^{(0)}(t'), \dot{\xi}^{(0)}(t')\right) \, dt' \\ &= S^{(0)}(Q_0, P_0) + \mathcal{F}S^{(1)}(Q_0, P_0). \end{aligned} \tag{4.60}$$

As long as the zeroth-order trajectories $\xi^{(0)}$ are unique, i.e., they do not reach the same point at the same time, as is the case in our problem, to calculate $S(Q_0, P_0)$ we choose that unique zeroth-order trajectory which reaches (Q_0, P_0), and substitute for t in the upper limit in Eq. (4.60) the time at which this point is reached. Given the scalar field $S(Q_0, P_0)$ we calculate momenta in the usual way, $p_Q = \frac{\partial S}{\partial Q_0}$, etc.,

[12] A better notation is $S(Q, P)$ emphasizing that S is a scalar field over the (Q, P) space, but this may be confused with $Q(t)$ and $P(t)$ of a trajectory, so we will stick to the notation Q_0 and P_0, with the subscript.

and we have produced for ourselves momenta to first order in $\mathcal{F}$ without the need to solve any equations! The method can be extended to higher order. We feed the obtained $\mathcal{F}$-series for momenta into the *full* equations for $\dot{Q}$ and $\dot{P}$ and proceed with the conventional perturbation theory, $Q(t) = Q^{(0)}(t) + \mathcal{F}Q^{(1)}(t)$, etc., making sure that all of the terms of a given order are included. At the end, we would obtain perturbed trajectories $(Q(t), P(t))$. It may be that nearby trajectories intersect. An envelope of these intersections is a caustic.

A few words need to be mentioned about the choice of the coordinate system in which to do the analysis. Polar system seems to be most useful only for obtaining zeroth-order solutions, because that is when the angular symmetry takes hold. At a nonzero $\mathcal{F}$, the fixed point moves away from zero and therefore $\theta(t)$ is no longer a small perturbation over its zeroth-order version for large negative times. After some consideration, it appears that at nonzero $\mathcal{F}$, it is best to perform the analysis in a Cartesian system in the time domain, i.e. to perform the action method for $Q(t)$ and $P(t)$ and the associated momenta, rather then $r(\theta)$ for example.

For convenience, we restate the zeroth-order solution:

$$\begin{aligned} r^{(0)}(t) &= r_0 e^{\eta t}, \\ \theta^{(0)}(t) &= \theta_0 + t - \frac{r_0^2}{2\eta}\left(e^{2\eta t} - 1\right). \end{aligned}$$

Or

$$\begin{aligned} Q^{(0)}(t) &= e^{\eta t}\left[Q_0 \cos\psi(t) - P_0 \sin\psi(t)\right], \\ P^{(0)}(t) &= e^{\eta t}\left[Q_0 \sin\psi(t) + P_0 \cos\psi(t)\right], \\ \text{where } \psi(t) &= t - \frac{Q_0^2 + P_0^2}{2\eta}\left[e^{2\eta t} - 1\right]. \end{aligned}$$

This solution is designed to arrive to point (r_0, θ_0) or (Q_0, P_0) at time $t = 0$. Remember that the superscript denotes the order of the solution versus time, while the subscript denotes the point in space; the subscripts were used to avoid confusion with

the full solution $Q(t)$ and $P(t)$. Applying the action method we have,

$$\begin{aligned} S(r_0,\theta_0) &= \int_{-\infty}^{0} \mathcal{L}^{(0)}\left(r^{(0)}(t),\theta^{(0)}(t)\right)\,dt + \mathcal{F}\int_{-\infty}^{0} \mathcal{L}^{(1)}\left(r^{(0)}(t),\theta^{(0)}(t)\right)\,dt, && (4.61)\\ \mathcal{L}^{(0)} &= \frac{1}{4}\left[(\dot{r}+\eta r)^2 + \left(r\dot{\theta} - r + r^3\right)^2\right], && (4.62)\\ \mathcal{L}^{(1)} &= -\frac{1}{2}\left[\left(r\dot{\theta} - r + r^3\right)\cos\theta + (\dot{r}+\eta r)\sin\theta\right]. && (4.63) \end{aligned}$$

In view of this, and Eqs. (4.44)-(4.45), see that

$$\begin{aligned} S^{(0)}(r_0,\theta_0) &= \frac{\eta}{2} r_0^2, && (4.64)\\ S^{(1)}(r_0,\theta_0) &= -\eta r_0 \int_{-\infty}^{0} e^{\eta t}\sin\left(\theta_0 + t - \frac{r_0^2}{2\eta}\left(e^{2\eta t}-1\right)\right)dt && (4.65)\\ &= -r_0\int_0^1 \sin\left(\theta_0 + \frac{1}{\eta}\left[\ln\xi - r_0^2\left(\frac{\xi^2-1}{2}\right)\right]\right)d\xi. && (4.66) \end{aligned}$$

Notice that compactly, the dependence of $S^{(1)}$ upon (r_0,θ_0) appears as parameters inside the integral, and $p_i^{(1)}$ can be effortlessly obtained by differentiation with respect to these parameters under the integral sign - we do not have to mess with differentiation of the limits of integrals! We now drop the subscripts such as 0 in r_0 or θ_0. The superscripts indicate the order in the perturbation theory. We have

$$S = S^{(0)} + \mathcal{F}S^{(1)}. \tag{4.67}$$

The Jacobian is

$$\begin{aligned} J &= \frac{\partial p_Q}{\partial Q}\frac{\partial p_P}{\partial P} - \frac{\partial p_Q}{\partial P}\frac{\partial p_P}{\partial Q} = \frac{\partial^2 S}{\partial Q^2}\frac{\partial^2 S}{\partial P^2} - \left(\frac{\partial^2 S}{\partial P \partial Q}\right)^2 && (4.68)\\ &= \frac{\partial^2 S^{(0)}}{\partial Q^2}\frac{\partial^2 S^{(0)}}{\partial P^2} + \mathcal{F}\left[\frac{\partial^2 S^{(0)}}{\partial Q^2}\frac{\partial^2 S^{(1)}}{\partial P^2} + \frac{\partial^2 S^{(0)}}{\partial P^2}\frac{\partial^2 S^{(1)}}{\partial Q^2}\right] + O(\beta). && (4.69) \end{aligned}$$

The mixed partials of $S^{(0)}$ disappear in view of Eq. (1). Also, in view of this we have

$$J = \eta^2 + \eta\mathcal{F}\nabla^2 S^{(1)} + O(\beta). \tag{4.70}$$

The appearance of the Laplacian is fortunate, since it is a geometric object and can therefore be taken in any coordinates. We will work in polar coordinates, so

$$\nabla^2 S^{(1)} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial S^{(1)}}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 S^{(1)}}{\partial \theta^2}. \tag{4.71}$$

We now compute the derivatives. First note that

$$r\frac{\partial S^{(1)}}{\partial r} = S^{(1)} + \frac{r^3}{\eta}\int_0^1 (\xi^2 - 1)\cos(\phi)\, d\xi. \tag{4.72}$$

So

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial S^{(1)}}{\partial r}\right) = \frac{1}{r}\frac{\partial S^{(1)}}{\partial r} + \frac{1}{r}\frac{\partial}{\partial r}\left[\frac{r^3}{\eta}\int_0^1 (\xi^2 - 1)\cos(\phi)\, d\xi\right]. \tag{4.73}$$

The first term on the rhs of Eq.(10) can be read off directly from two lines above:

$$\frac{1}{r}\frac{\partial S^{(1)}}{\partial r} = -\frac{1}{r}\int_0^1 \sin(\phi)\, d\xi + \frac{r}{\eta}\int_0^1 (\xi^2 - 1)\cos(\phi)\, d\xi. \tag{4.74}$$

The second term is

$$\frac{3r}{\eta}\int_0^1 (\xi^2 - 1)\cos(\phi)\, d\xi + \frac{r^3}{\eta^2}\int_0^1 (\xi^2 - 1)^2 \sin(\phi)\, d\xi. \tag{4.75}$$

Moreover,

$$\frac{1}{r^2}\frac{\partial^2 S^{(1)}}{\partial \theta^2} = \frac{1}{r}\int_0^1 \sin(\phi)\, d\xi. \tag{4.76}$$

All together,

$$\nabla^2 S^{(1)} = \frac{4r}{\eta}\int_0^1 (\xi^2 - 1)\cos(\phi)\, d\xi + \frac{r^3}{\eta^2}\int_0^1 (\xi^2 - 1)^2 \sin(\phi)\, d\xi. \tag{4.77}$$

So the meat of the calculation has to do with understanding the integrals.

$$I_1 = \int_0^1 (\xi^2 - 1)\cos(\phi)\, d\xi, \tag{4.78}$$

$$I_2 = \int_0^1 (\xi^2 - 1)^2 \sin(\phi)\, d\xi. \tag{4.79}$$

The analysis of the integrals I_1 and I_2, Eqs. (4.78)-(4.79), can be performed with the use of Γ-functions. Consider

$$\begin{aligned} I^{(p)} &= \int_0^1 \xi^p \exp\left(i\theta + \frac{i}{\eta}\left[\ln\xi - r^2\left(\frac{\xi^2-1}{2}\right)\right]\right) d\xi \\ &= e^{i\theta}e^{i\kappa}\int_0^1 \xi^p \xi^\alpha e^{-i\kappa\xi^2}\, d\xi, \end{aligned} \tag{4.80}$$

where $\kappa = \frac{r^2}{2\eta}$, $\alpha = \frac{i}{\eta}$ and p are either 0, 2 or 4. Next, change variables to $x = i\kappa\xi^2$. Then

$$I^{(p)} = \frac{e^{i\theta}e^{i\kappa}}{2(i\kappa)^z}\int_0^{i\kappa} x^{z-1}e^{-x}\, dx, \tag{4.81}$$

where $z = \frac{\alpha+p+1}{2} = \frac{i}{2\eta} + \frac{p+1}{2}$. Note that z lies in the right-half z-plane. When $r \gg \sqrt{\eta}$ the upper limit can be replaced by ∞ (notice that z is independent of r, so as r is increased but η is held fixed, only the upper limit changes, but the structure of the integrand does not, so its ok to make this replacement). The integral is taken in the x-plane along the imaginary axis, but the integrand is an analytic function of x, so we may rotate the contour to go along the real x-axis. Then, the integral coincides with the definition of the Γ-function in the variable z. Thus,

$$I^{(p)} = \frac{e^{i\theta}e^{i\kappa}}{2(i\kappa)^z}\Gamma(z). \tag{4.82}$$

Notice that in the limit of large $\eta \to \infty$, $z \to \frac{p+1}{2}$, i.e. $\alpha \to 0$. But $\alpha = 0$ is equivalent to starting out with the integral that does not contain the ln-term. This justifies dropping the ln-term in the case of large η.

Now,

$$\begin{aligned} I_1 &= \frac{I^{(2)}}{2} - \frac{I^{(0)}}{2} + c.c. \\ &= \frac{e^{i(\theta+\kappa)}}{4(i\kappa)^{\frac{i}{2\eta}}}\left[\frac{\Gamma\left(\frac{i}{2\eta}+\frac{3}{2}\right)}{(i\kappa)^{3/2}} - \frac{\Gamma\left(\frac{i}{2\eta}+\frac{1}{2}\right)}{(i\kappa)^{1/2}}\right] + c.c., \end{aligned} \tag{4.83}$$

$$\begin{aligned} I_2 &= \frac{I^{(4)}}{2i} - \frac{I^{(2)}}{i} + \frac{I^{(0)}}{2i} + c.c. \\ &= \frac{e^{i(\theta+\kappa)}}{4i(i\kappa)^{\frac{i}{2\eta}}}\left[\frac{\Gamma\left(\frac{i}{2\eta}+\frac{5}{2}\right)}{(i\kappa)^{5/2}} - 2\frac{\Gamma\left(\frac{i}{2\eta}+\frac{3}{2}\right)}{(i\kappa)^{3/2}} + \frac{\Gamma\left(\frac{i}{2\eta}+\frac{1}{2}\right)}{(i\kappa)^{1/2}}\right] + c.c.\ . \end{aligned} \tag{4.84}$$

These approximations match nicely with numerical evaluations of the full I_1 and I_2 from Eqs. (4.78)-(4.79). Each term in the expression for I_1 is of the form

$$\frac{e^{\frac{\pi}{4\eta}}e^{i\left(\theta+\kappa-\frac{\ln\kappa}{2\eta}-\frac{q\pi}{2}\right)}}{4\kappa^q}\Gamma\left(\frac{i}{2\eta}+q\right), \tag{4.85}$$

where $\kappa = \frac{r^2}{2\eta}$. The same is true for each term in the expression for I_2 with an extra phase shift by $-\pi/2$. Note that both integrals behave like $\sim 1/r$ for large r. With the help of Eq. (4.77), the r-dependence of the correction to the momentum (see Eq. (4.70)) is

$$\mathcal{F}\nabla^2 S^{(1)} \sim \mathcal{F}r^2, \tag{4.86}$$

and hence the comparison of this with the zeroth-order part suggests that $R \sim \mathcal{F}^{-1/2}$. The η-dependence is not as simple, but attains simpler forms in the asymptotic regimes of $\eta \gg 1$ and $\eta \ll 1$. In particular, for large η and r, both I_1 and I_2 behave like $\frac{\eta^{1/2}}{r}$, which implies, with the help of Eqs. (4.77) and (4.70) that

$$J \sim \eta^2 + \mathcal{F}\frac{r^2}{\eta^{1/2}}. \tag{4.87}$$

Equating the two terms in J suggests that

$$R \sim \frac{\eta^{5/4}}{\mathcal{F}^{1/2}}. \tag{4.88}$$

For small η and large r, both I_1 and I_2 behave like $\frac{\eta^{1/2}}{r} e^{\frac{\pi}{4\eta}} \Gamma\left(\frac{i}{2\eta} + \frac{1}{2}\right)$. We invoke Stirling's asymptotic approxmation of the Γ-function: for large z

$$\Gamma(z) \sim \sqrt{2\pi} z^{z-1/2} e^{-z}. \tag{4.89}$$

Then

$$\Gamma\left(\frac{i}{2\eta}\right) \sim \sqrt{2\pi} e^{i\left(-\frac{1}{2\eta} - \frac{1}{2} + \frac{1}{2\eta}\ln\frac{1}{2\eta}\right)} e^{-\frac{\pi}{4\eta}}. \tag{4.90}$$

Putting it all together, we see that both I_1 and I_2 behave like $\frac{\eta^{1/2}}{r}$; to our luck, the exponential factors $e^{-\frac{\pi}{4\eta}}$ and $e^{\frac{\pi}{4\eta}}$ cancel each other out and prevent essential singularity. Thus, we can again conclude that for large r,

$$R \sim \frac{\eta^{5/4}}{\mathcal{F}^{1/2}}. \tag{4.91}$$

4.3 Physics Without the Slowing-down Region

The only difference is that now the phase has a different sign:

$$\phi(\xi) = \theta + \frac{1}{\eta}\left[\ln\xi + r^2\left(\frac{\xi^2 - 1}{2}\right)\right], \tag{4.92}$$

and therefore

$$\begin{aligned} I^{(p)} &= \int_0^1 \xi^p \exp\left(i\theta + \frac{i}{\eta}\left[\ln\xi + r^2\left(\frac{\xi^2-1}{2}\right)\right]\right) d\xi \\ &= e^{i\theta} e^{i\kappa} \int_0^1 \xi^p \xi^\alpha e^{i\kappa\xi^2}\, d\xi, \end{aligned} \tag{4.93}$$

where $\kappa = \frac{r^2}{2\eta}$, $\alpha = \frac{i}{\eta}$ and p are either 0, 2 or 4. Next, change variables to $x = -i\kappa\xi^2$. Then

$$I^{(p)} = \frac{e^{i\theta}e^{-i\kappa}}{2(-i\kappa)^z}\int_0^{-i\kappa} x^{z-1}e^{-x}\,dx, \tag{4.94}$$

where $z = \frac{\alpha+p+1}{2} = \frac{i}{2\eta} + \frac{p+1}{2}$. Again, in the limit of $r \gg \sqrt{\eta}$ the upper limit can be replaced by $-\infty$ and the contour can be rotated to go to ∞ along the real axis. Thus,

$$I^{(p)} = \frac{e^{i\theta}e^{i\kappa}}{2(-i\kappa)^z}\Gamma(z), \tag{4.95}$$

and the scaling of R with $\mathcal{F}$ would remain unchanged.

4.4 Perturbation Theory for $\mathbf{r}(\mathbf{p})$

Our method above was to extract the physics by comparing the first-order correction to the zeroth order term, as this was a signature when the Jacobian J begins to blow up. There is no hope of actually observing any singularities of J with the use of the perturbation theory - for one thing, the first-order correction to the action $S^{(1)}$ is single-valued, so there is no hope to extract information about the folds directly. On the other hand, we know from numerical work that for small $\mathcal{F}$ the trajectories are indeed perturbed only a little. The infinity of J is only an artifact of the particular coordinate system that we use. In reality, the Lagrangian manifold obtains a small "wrinkle", which when viewed from just the right perspective has an infinite slope, but there is nothing special about this perspective. One would suspect then, that there must exist a perturbation theory which converges and which captures the real perturbed trajectories.

One such possibility is to express the manifold as $\mathbf{r(p)}$ rather then $\mathbf{p(r)}$, which means we have to express the action as a function of momenta, call it $\tilde{S}(\mathbf{p})$. Since $\mathbf{p} = \nabla S(\mathbf{r})$, $\tilde{S}(\mathbf{p})$ is just the Legendre transform of $S(\mathbf{r})$. In other words

$$\tilde{S}(p_P, p_Q) = p_P P(p_Q, p_P) + p_Q Q(p_Q, p_P) - S(Q(p_P, p_Q), P(p_Q, p_P)). \tag{4.96}$$

This is true, but we would like an expression for $\tilde{S}(p_P, p_Q)$ of the form of Eq. (4.65). We follow a procedure similar to the procedure employed in Landau and Lifshitz to perform the Legendre transformation on the Lagrangian and to derive the Hamiltonian. Illustrated on a 1-variable action we have:

$$\begin{aligned} dS &= \frac{\partial S}{\partial q} dq + \frac{\partial S}{\partial t} dt \\ &= p dq - \mathcal{H} dt. \end{aligned} \tag{4.97}$$

Add qdp to both sides of the equation. Then

$$dS + qdp = d(qp) - \mathcal{H} dt. \tag{4.98}$$

Rearranging terms we have

$$d(S - qp) = -(qdp + \mathcal{H}) dt, \tag{4.99}$$

and we recognize the lhs as the differential of $\tilde{S}$. Hence

$$\tilde{S}(p) = -\int q\, dp - \int \mathcal{H}(q(t), p(t))\, dt \tag{4.100}$$

$$= -\int q(t)\dot{p}(r) + \mathcal{H}(q(t), p(t))\, dt. \tag{4.101}$$

Now, let the perturbed trajectory be given by $\mathbf{r}(t) = \mathbf{r}^{(0)}(t) + \mathcal{F}\mathbf{r}^{(1)}(t)$ and similar for momenta $\mathbf{p}(t) = \mathbf{p}^{(0)}(t) + \mathcal{F}\mathbf{p}^{(1)}(t)$. They have the same initial and final points in the phase space as the unperturbed trajectories. Then

$$\begin{aligned} \tilde{S} = -\int \Big\{ &\left[q^{(0)} + \mathcal{F} q^{(1)}\right] \left[\dot{p}^{(0)} + \mathcal{F}\dot{p}^{(1)}\right] \, dt \\ &+ \mathcal{H}^{(0)}(q^{(0)}, p^{(0)}) + \mathcal{F} \left.\frac{\partial \mathcal{H}^{(0)}}{\partial q}\right|_{q^{(0)}, p^{(0)}} q^{(1)}(t) + \mathcal{F} \left.\frac{\partial \mathcal{H}^{(0)}}{\partial p}\right|_{q^{(0)}, p^{(0)}} p^{(1)}(t) + \mathcal{F}\mathcal{H}^{(1)}\left(q^{(0)}, p^{(0)}\right) \Big\} \, dt. \end{aligned} \tag{4.102}$$

With the help of the canonical equations we have

$$\tilde{S} = S^{(0)} - \mathcal{F} \int \left[q^{(1)} \dot{p}^{(0)} + q^{(0)} \dot{p}^{(1)} - q^{(1)} \dot{p}^{(0)} + \dot{q}^{(0)} p^{(1)} + \mathcal{H}^{(1)} \left(q^{(0)}, p^{(0)} \right) \right] \, dt. \quad (4.103)$$

The first and third terms obviously cancel out and we can perform integration by parts on the fourth term. The boundary terms are zero because $p^{(1)}$ is zero at the initial and final times and the other term cancels out the second term in the integrand above. We are this left with

$$\tilde{S} = S^{(0)} - \mathcal{F} \int \mathcal{H}^{(1)} \left(q^{(0)}, p^{(0)} \right) \, dt, \quad (4.104)$$

which now needs to be reexpressed in terms of momenta. Recall that at zeroth-order in $\mathcal{F}$ we have $p_Q = \eta Q$ and $p_P = \eta P$. So we can just use Eq. (4.64) to get

$$\tilde{S}^{(0)}(p_Q, p_P) = \frac{1}{2\eta} \left(p_Q^2 + p_P^2 \right). \quad (4.105)$$

Now,

$$\mathcal{H}^{(1)} = p_P. \quad (4.106)$$

Recall also that

$$P^{(0)} = r e^{\eta t} \sin \left(\theta + - \frac{r^2}{2\eta} \left(e^{2\eta t} - 1 \right) \right), \quad (4.107)$$

and as was just mentioned, $p_P^{(0)} = \eta P^{(0)}$. Therefore, the first-order term in Eq. (4.104) is *exactly* the same as that given by Eq. (4.65), except that the result must be expressed in terms of p_P and p_Q instead of P and Q, but that is trivial with the relations just given. Eq. (4.99) tells us that $\mathbf{q} = \nabla_{\mathbf{p}} \tilde{S}$ - this is the alternative expression for

the manifold $\mathbf{q}(\mathbf{p})$. The new Jacobian is

$$\begin{aligned}
\tilde{J} &= \frac{\partial^2 S^{(0)}}{\partial p_Q^2}\frac{\partial^2 S^{(0)}}{\partial p_P^2} + \mathcal{F}\left[\frac{\partial^2 S^{(0)}}{\partial p_Q^2}\frac{\partial^2 S^{(1)}}{\partial p_P^2} + \frac{\partial^2 S^{(1)}}{\partial p_Q^2}\frac{\partial^2 S^{(0)}}{\partial p_P^2}\right] \\
&= \frac{1}{\eta^2} + \frac{\mathcal{F}}{\eta^3}\left[\frac{\partial S^{(1)}}{\partial Q^2} + \frac{\partial S^{(2)}}{\partial P^2}\right] \\
&= \frac{J}{\eta^4}.
\end{aligned}$$

Thus, the new method does not predict anything new!

Part II

Renormalization Group Method for Predicting Frequency Clusters in a Chain of Nearest-Neighbor Phase Oscillators

Chapter 5

Synchronization: Background

More then three centuries have passed since Huygens's observation of entrainment of clocks [90]. In modern days, the study of synchronization in nature has been a subject of major revival, yet some of the most challenging questions remain unanswered.

A variety of models have been in use for these systems. A "toy model" for exploration of various issues in this field is the phase model, which represents oscillators that interact with each other through an odd function of their phase differences, usually taken to be sine:

$$\dot{\theta}_i = \omega_i + \sum_{i \neq j} K_{i,j} \sin(\theta_j - \theta_i). \tag{5.1}$$

Not unlike the Ising model of magnetism, this phase model of oscillators is sufficiently simple to allow for the development of intuition, yet it contains enough physics to capture some of the phenomenology exhibited by more realistic models. An analysis of a more general model that allows for a change of an amplitude of each oscillator reveals that under certain conditions the dynamics develops an attracting limit cycle with a fixed (or approximately fixed) amplitude, and can thus be reduced to the aforementioned phase model [47, 91]. An example of a procedure like this has been given recently in the theoretical study of nanomechanical oscillators [92]. In the phase mode the set of frequencies $\{\omega_i\}$ is random and is sampled from a distribution $g(\omega)$.

Analogously to systems in equilibrium Statistical Mechanics, questions about the transition to global synchronization arises in the context of synchronization as well; in

the limit $N \to \infty$ one may expect abrupt transitions to globally synchronized states. As in the equilibrium realm, such a transition is reflected in the abrupt change in an order parameter. For oscillators, one choice of an order parameter is

$$r = \frac{1}{N} \sum_{j=1}^{N} e^{i\theta_j}, \tag{5.2}$$

which may attain values $\gg 1/\sqrt{N}$ when some macroscopic fraction of oscillators shares a common period. The question of the existence of such a transition to macroscopic synchronization is subject to dimensionality, properties of coupling $K_{i,j}$ and properties of $g(\omega)$. For example, the well-known all-to-all coupling $K_{i,j} = K/N$ treated by Kuramoto [46],

$$\dot{\theta}_i = \omega_i + \frac{K}{N} \sum_{i \neq j} \sin(\theta_j - \theta_i), \tag{5.3}$$

allows for a threshold value of $K = K_c$ beyond which an order parameter attains values $\gg 1/\sqrt{N}$. Due to the mean-field nature of the all-to-all interaction, the value of K_c can be found analytically - it is $2/\pi g(0)$ [45, 47, 93]. Other features of this transition to the synchronized state, such as the growth of the fraction of locked oscillators as the coupling increases above the critical value, can also be calculated.

Many systems where synchronization might occur involve short-range coupling [48, 49], such as oscillators on a lattice with nearest-neighbor interactions. Although nanomechanical examples have not been constructed yet, it is plausible that interactions will be short range if carried by the support or a substrate between oscillating nanobeams. Systems with short range coupling are much harder to analyze, and much of our understanding of these systems rests on numerical simulations of the equations of motion as there are few theoretical tools that yield a quantitative description of the collective motion.

The focus here will be on a one-dimensional chain of phase oscillators interacting

with their nearest neighbors.

$$\dot{\theta}_i = \omega_i + K \sin(\theta_{i-1} - \theta_i) + K \sin(\theta_{i+1} - \theta_i). \tag{5.4}$$

In spite of its simplicity, this model presents a challenge as we will see below. It has been shown by Strogatz and Mirollo that a nearest-neighbor chain cannot exhibit extensive synchronization for any finite vale of K [51]. Strogatz and Mirollo also described the probability for a global synchronization at finite N and how this probability tends to zero as $N \to \infty$ [52]. Although global synchronization is not possible, interesting collective structures do occur. In particular, oscillators tend to organize themselves into clusters of common frequency, $\bar{\omega}$, defined as

$$\bar{\omega}_i \equiv \lim_{(t-t_0)\to\infty} \frac{\theta_i(t) - \theta_i(t_0)}{t - t_0}, \tag{5.5}$$

such that each oscillator in a given cluster shares its $\bar{\omega}$ with all the other oscillators in the same cluster [51], but the dynamics of different oscillators in a given cluster may be different. Hence, the structure of the unsynchronized phase in the nearest-neighbor chain is quite complex. The relevant questions are regarding the distribution of cluster lengths and distribution of cluster frequencies as a function of the distribution $g(\omega)$ of frequencies.[1] $\{\omega_i\}$ As the coupling strength K is increased, the average size of such clusters grows, Fig. 5.1, but it does not become infinite at a finite K [51].

In a sense, understanding the properties of such clusters is a more complicated problem then predicting the feasibility of global synchronization, because the latter question is equivalent to predicting the existence of an attracting fixed point of the dynamics (this by itself is far from trivial in the $N \to \infty$ limit), while there is no such clear strategy in the case of frequency clusters - they are a dynamic phenomenon. Twenty years ago, Strogatz and Mirollo wrote [51]:

> "The dynamical behavior of $\dot{\theta}_i = \omega_i + K \sum_{\langle i,j \rangle} \sin(\theta_j - \theta_i)$, $j \in \{1, ..., N\}^d$ is not well understood in the regime before phase-locking occurs. In par-

[1] also as a function of the distribution $G(K)$ of couplings $\{K_i\}$ if we allow disorder in these as well

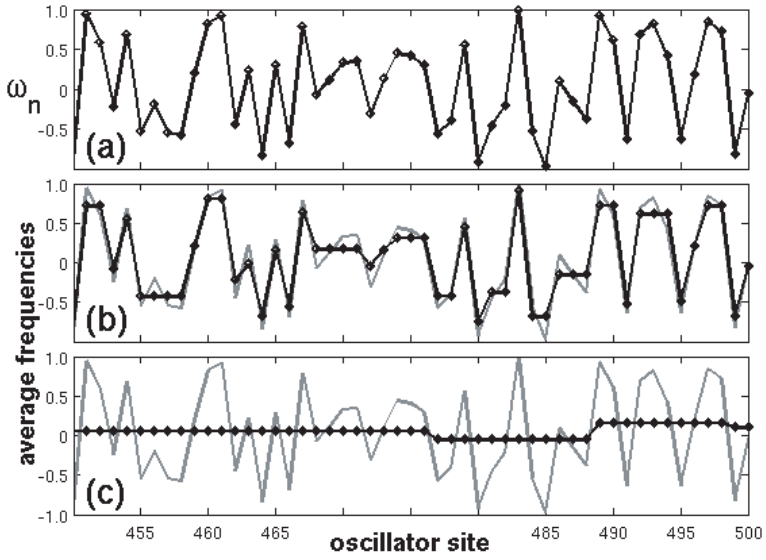


Figure 5.1: Development of frequency clusters as the value of coupling K is increased, shown on a small section of a 1000 oscillator chain. Solid diamonds: time-averaged running frequencies of oscillators, defined similarly to Eq. (6.40) with $t - t_0$ a large, but finite time interval; lines: intrinsic frequencies that appear in Eqn. (5.4).

> ticular, it is not known if or how the distribution of number and size of synchronized clusters scale with K, N, and d"

where d is the dimensionality of the system.

The understanding of these structures from the point of view of dynamical systems has been deepended by a series of papers by by Ermentrout and Kopell in early 1980s. For example, in [94] they considered a chain of oscillators with frequencies laid out linearly (not randomly) along the chain, proved the existence of limit cycles in the $\vec{\theta}$-space and related the first cluster break in the chain as the coupling strength is lowered to the appearance of a limit. Using this they were able to predict sizes of frequency differences between neighboring clusters. Around the same time, in a rather unrelated field, Dasgupta and Ma [95, 96], developed a renormalization group (RG) analysis of the ground state of 1-dimensional random quantum spin chains (see also more recent work by Daniel Fisher and co-workers [97, 98]). Their scheme was based on solving for the energy eigenvalues of the small parts of the chain (two spins), while treating the effect of the neighbors via a secon-order perturbation theory. Using this,

they were able to progressively thin out the spins interacting via largest couplings and in this way to set up an RG flow.

The main idea behind the strong disorder approach applied to quantum-mechanical problem is best demonstrated in the case of the spin-1/2 Heisenberg model:

$$\hat{H} = \sum_i J_i \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_{i+1}. \tag{5.6}$$

The ith term in the sum considered on its own, splits the four states of the spins i and $i+1$, into a singlet state, and three triplet states excited by the energy J_i. Thus a real-space decimation step can be proposed: freeze the spin pair with the strongest interaction, $J_n = \max\{J_i\}$, into a singlet state. Perturbative corrections to this state introduce an effective coupling between the neighbors of the spins n and $n+1$ which is:

$$H_{RG} = \frac{J_{n-1}J_{n+1}}{2J_n} \hat{\mathbf{S}}_n \cdot \hat{\mathbf{S}}_{n+1}, \tag{5.7}$$

an interaction identical in form to the operators in the bare Hamiltonian, but with a much suppressed strength: $J_{eff} = \frac{J_{n-1}J_{n+1}}{2J_n} < J_{n-1}, J_n, J_{n+1}$. By repeating this decimation step the number of free spins is gradually decreased; at the stage when all spins are bound into singlets, we obtain the ground state of the Hamiltonian. Note that this procedure is possible if the coupling to the rest of the chain is much weaker then between the two spins in question. Thus, their method was best applicable to chains that contained a large fraction of couplings which are significantly different from adjacent couplings, in other words, this technique applied to chains with broad coupling distributions.

More recently, the technique of strong randomness RG has been extended to explain the superfluid-insulator transition in random bosonic chains [99, 100] (a quantum reactive analog of the nearest-neighbor chain, where the $\ddot{\theta}$ replaces $\dot{\theta}$, and θ is a quantum operator). This latter system consists of superconducting grains of random sizes coupled to each other via Josephson couplings which are also random. Those superconducting grains with exceptionally large charging energies tend to introduce a

break in the chain, but perturbative calculation demonstrates that the effect of such a grain is to establish a weak effective coupling between its adjacent grains. Those couplings which are exceptionally large tend to lock together the phases of the order parameter of the superconducting grains which they couple, thus allowing these two grains to be treated as one grain with a new effective charging energy. An RG based on these decimation steps was built and its analysis led to the mapping out of a phase diagram for parameters under which the system behaves as a superconductor and those under which it behaves as a Mott insulator. This served as our motivation for an analogous task with the chain of oscillators.

The rest of the work serves a dual purpose - to develop an RG method as a quick and efficient tool for the analysis of the nearest-neighbor model and then use this tool to study the physics of frequency clusters exhibited by this model. The content of this part of the thesis follows the recent paper by Kogan, Rogers, Refael, and Cross [101], although the presentation is significantly expanded. The presentation in [101] is less intensive on the algebra, and therefore forms a good supplement to the more detailed derivations here, especially in Sections 6.2 and 6.3.

Chapter 6

Oscillator Chain RG

6.1 Idea

We will develop the strong-disorder real space RG method to analyze the random oscillator chain. The nearest-neighbor chain presents a classical problem with dissipative equations of motion, rather than a quantum problem with a conserved Hamiltonian. Nevertheless, we can apply the strong disorder RG method to a classical problem as well:

$$\dot{\theta}_i = \omega_i + K_{i-1} \sin(\theta_{i-1} - \theta_i) + K_i \sin(\theta_{i+1} - \theta_i), \tag{6.1}$$

where we generalized to introduce the randomness into the couplings. The distributions of ω_i is given by $g(\omega)$ and couplings K_i come from the distribution $G(K)$. These random frequencies and couplings are treated roughly speaking, as energies in a Hamiltonian. Indeed, a strong coupling between two oscillators will tend to force them to rotate at the same frequency, which allows for a renormalized description of two oscillators by one effective oscillator. Concomitantly, oscillators with frequencies which are anomalously different from each of their neighbors' frequencies will rotate essentially independently, or as free oscillators, with only small perturbative corrections from their neighbors; these neighbors are in turn perturbed only slightly by the anomalous oscillator. This allows us to get rid of such anomalous oscillators from the description of the chain and replace it with the description in which each of the neighbors' parameters are slightly changed.

As the RG proceeds the parameters K_i and ω_i are modified and the strength of their disorder changes.[1] As equations of motion are renormalized, they generally acquire new terms. We list these terms, but ignore them in this work, except for the appearance of the parameter m_i:

$$m_i\dot{\theta}_i = m_i\omega_i + K_{i-1}\sin(\theta_{i-1} - \theta_i) + K_i\sin(\theta_{i+1} - \theta_i). \tag{6.2}$$

It will follow from the strong-coupling decimation step that when two oscillators with m_i and m_{i+1} are described by one effective oscillator, its m is given simply by $m_i + m_{i+1}$. Hence, the value of m equals the number of original oscillators represented by this effective oscillator. The original chain, which is the starting point for RG will have all $m_i = 1$. It will become apparent that oscillators with larger m are less affected by perturbations from their neighbors, hence m is a measure of sensitivity of an oscillator to outside perturbations (similarly to inertia in Mechanics) and will be referred to as its "mass", although it is not a true mass in the sense of Newton's second law. We now proceed to develop each decimation step in detail.

6.2 Strong-Coupling Step

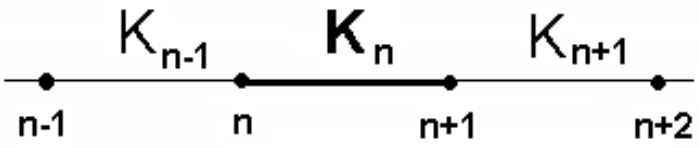


Figure 6.1: A small piece of the chain around the strongest coupling, labelled here as K_n. All the neighboring K_i and all the local ω_i are assumed to have much smaller magnitude then K_n.

Consider the largest coupling in the chain: suppose this coupling is K_n. Because of strong disorder, the frequencies ω_n and ω_{n+1} of the pair of oscillators that K_n couples as well as the couplings K_{n-1} and K_{n+1} to the neighbors of the pair will almost

[1] Even if the bare chain contains all identical K_i, the RG will create also $K_i = 0$. However, we do not expect the RG to work well on chains with all identical K_i.

always be small compared to K_n. Associated with this, there are small parameters $\omega_n/K_n, K_{n-1}/K_n$, etc. We will occasionally use a symbol ϵ to denote quantities of this order. Intuitively we know that oscillators n and $n+1$ rotate together as one oscillator, with a small phase difference between them that remains bounded. Notice that it is not the large magnitude of K_n by itself which causes this; the phase difference between θ_n and θ_{n+1} can not be assumed to be bounded if the coupling K_n between them is very large but the difference between their intrinsic frequencies or another neighboring K_n are comparably large. Therefore, it is the small value of ratios such as ω_n/K_n and K_{n-1}/K_n that determine whether an oscillator is subjected to the strong-coupling step.

Because we expect the phase difference to be small, it will be convenient to switch from variables θ_n and θ_{n+1} to the mass-weighted average phase Θ and the phase difference variable δ defined as

$$\Theta \equiv (m_n\theta_n + m_{n+1}\theta_{n+1})/M, \tag{6.3}$$

and

$$\delta \equiv \theta_{n+1} - \theta_n, \tag{6.4}$$

respectively, where $M = m_n + m_{n+1}$. For reference, the equations of motion for the strongly coupled oscillators and their neighbors read:

$$\begin{aligned}
m_{n-1}\dot{\theta}_{n-1} &= m_{n-1}\omega_{n-1} + K_{n-2}\sin(\theta_{n-2} - \theta_{n-1}) + K_{n-1}\sin(\theta_n - \theta_{n-1}), && (6.5)\\
m_n\dot{\theta}_n &= m_n\omega_n + K_{n-1}\sin(\theta_{n-1} - \theta_n) + K_n\sin(\theta_{n+1} - \theta_n), && (6.6)\\
m_{n+1}\dot{\theta}_{n+1} &= m_{n+1}\omega_{n+1} + K_n\sin(\theta_n - \theta_{n+1}) + K_{n+1}\sin(\theta_{n+2} - \theta_{n+1}), && (6.7)\\
m_{n+2}\dot{\theta}_{n+2} &= m_{n+2}\omega_{n+2} + K_{n+1}\sin(\theta_{n+1} - \theta_{n+2}) + K_{n+2}\sin(\theta_{n+3} - \theta_{n+2}). && (6.8)
\end{aligned}$$

Rewriting equations for θ_n and θ_{n+1} in terms of Θ and δ we get

$$M\dot{\Theta} = M\omega' + K_{n-1}\sin\left(\theta_{n-1} - \Theta + \frac{m_{n+1}}{M}\delta\right) + K_{n+1}\sin\left(\theta_{n+2} - \Theta - \frac{m_n}{M}\delta\right). \tag{6.9}$$

Similarly, oscillator θ_{n+2} is coupled to Θ and δ

$$m_{n+2}\dot{\theta}_{n+2} = m_{n+2}\omega_{n+2} + K_{n+1}\sin\left(\Theta - \theta_{n+2} + \frac{m_n}{M}\delta\right) + K_{n+2}\sin\left(\theta_{n+3} - \theta_{n+2}\right), \tag{6.10}$$

and analogously for oscillator θ_{n-1}. The phase difference δ satisfies

$$\dot{\delta} = (\omega_{n+1} - \omega_n) - \frac{K_n}{\mu}\sin(\delta) + \frac{K_{n+1}}{m_{n+1}}\sin\left(\theta_{n+2} - \Theta - \frac{m_n}{M}\delta\right) - \frac{K_{n-1}}{m_n}\sin\left(\theta_{n-1} - \Theta + \frac{m_{n+1}}{M}\delta\right). \tag{6.11}$$

By the strong-disorder assumption, the term $\sin\delta$ has a much larger prefactor $\frac{K_n}{\mu}$ then the other terms, so this term is responsible for quick relaxation of δ to the quasi-static equilibrium value given by

$$\delta(t) = \delta_0 + \frac{\mu}{K_n}\left[\frac{K_{n+1}}{m_{n+1}}\sin\left(\theta_{n+2} - \Theta\right) - \frac{K_{n-1}}{m_n}\sin\left(\theta_{n-1} - \Theta\right)\right] + O\left(\frac{\mu}{K_n}\right)^2, \tag{6.12}$$

where $\delta_0 = \mu\frac{\omega_{n+1}-\omega_n}{K_n}$. So δ consists of small oscillations of order ϵ centered around the time-independent mean δ_0 which is also of order ϵ and is proportional to $\omega_{n+1} - \omega_n$. The phase δ will mediate an interaction between oscillators $n-1$ and $n+2$, since each of these oscillators is coupled to δ. When this solution is substituted into $K_{n-1}\sin\left(\theta_{n-1} - \Theta + \frac{m_{n+1}}{M}\delta\right)$ we get

$$\begin{aligned} &K_{n-1}\sin\left(\theta_{n-1} - \Theta + \frac{m_{n+1}}{M}\delta_0\right) \\ &+ \frac{\mu}{M}\left[\frac{K_{n-1}K_{n+1}}{2K_n}\left(\sin\left(\theta_{n+2} + \theta_{n-1} - 2\Theta\right) + \sin\left(\theta_{n+2} - \theta_{n-1}\right)\right) - \frac{m_{n+1}}{m_n}\frac{K_{n-1}^2}{2K_n}\sin\left(2\theta_{n-1} - 2\Theta\right)\right] \end{aligned} \tag{6.13}$$

and when it is substituted into $K_{n+1}\sin\left(\theta_{n+2} - \Theta - \frac{m_n}{M}\delta\right)$ we get

$$\begin{aligned} &K_{n+1}\sin\left(\theta_{n+2} - \Theta - \frac{m_n}{M}\delta_0\right) \\ &+ \frac{\mu}{M}\left[\frac{K_{n-1}K_{n+1}}{2K_n}\left(\sin\left(\theta_{n+2} + \theta_{n-1} - 2\Theta\right) + \sin\left(\theta_{n-1} - \theta_{n+2}\right)\right) - \frac{m_n}{m_{n+1}}\frac{K_{n+1}^2}{2K_n}\sin\left(2\theta_{n+2} - 2\Theta\right)\right]. \end{aligned} \tag{6.14}$$

Therefore, Eq. (6.9) transforms into

$$M\dot{\Theta} = M\omega' + K_{n-1}\sin\left(\theta_{n-1} - \Theta + \frac{m_{n+1}}{M}\delta_0\right) + K_{n+1}\sin\left(\theta_{n+2} - \Theta - \frac{m_n}{M}\delta_0\right) + \frac{\mu}{M}\left[\frac{K_{n-1}K_{n+1}}{K_n}\sin\left(\theta_{n+2} - \theta_{n-1} - 2\Theta\right) - \frac{m_{n+1}}{m_n}\frac{K_{n-1}^2}{2K_n}\sin\left(2\theta_{n-1} - 2\Theta\right) - \frac{m_n}{m_{n+1}}\frac{K_{n+1}^2}{2K_n}\sin\left(2\theta_{n+2} - 2\Theta\right)\right], \quad (6.15)$$

where Ω is the mass-weighted frequency,

$$\omega' = \frac{\omega_n m_n + \omega_{n+1} m_{n+1}}{M}, \quad (6.16)$$

and

$$M = m_1 + m_2, \quad (6.17)$$

while Eq. (6.10) transforms into

$$m_{n+2}\dot{\theta}_{n+2} = m_{n+2}\omega_{n+2} + K_{n+1}\sin\left(\Theta - \theta_{n+2} + \frac{m_n}{M}\delta_0\right) + K_{n+2}\sin\left(\theta_{n+3} - \theta_{n+2}\right) + \frac{\mu}{M}\left[\frac{K_{n-1}K_{n+1}}{2K_n}\left(\sin\left(2\Theta - \theta_{n+2} - \theta_{n-1}\right) - \sin\left(\theta_{n-1} - \theta_{n+2}\right)\right) - \frac{m_n}{m_{n+1}}\frac{K_{n+1}^2}{2K_n}\sin\left(2\Theta - 2\theta_{n+2}\right)\right]. \quad (6.18)$$

An analogous expression exists for θ_{n-1}. To leading order it turns out that the oscillating corrections to δ give small extra coupling terms of forms not included in the equation of motion Eq. (6.2): a second-nearest-neighbor interaction between oscillators, a three body interaction involving the new effective oscillator and the neighboring pair, and a second harmonic correction to the form of the pairwise interaction. As is conventional in real space renormalization group approaches, we will neglect these more complicated interaction terms. Then, aside from the factor δ_0 resulting equations take the original form. Fortunately, in one dimension the phase δ_0 can be unwound by redefining the phases θ_{n-1} and all the phases prior to it as

$$\theta_j' = \theta_j + \frac{m_{n+1}}{M}\delta_0, \;\; j \leq n-1, \quad (6.19)$$

as well as the phase θ_{n+2} and all the phases following it as

$$\theta'_j = \theta_j - \frac{m_n}{M}\delta_0, \;\; j \geq n+2. \tag{6.20}$$

Finally, we arrive at

$$\begin{aligned}
m_{n-1}\dot{\theta}_{n-1} &= m_{n-1}\omega_{n-1} + K_{n-2}\sin(\theta_{n-2}-\theta_{n-1}) + K_{n-1}\sin(\Theta-\theta_{n-1}), & (6.21)\\
M\dot{\Theta}_n &= M\omega' + K_{n-1}\sin(\theta_{n-1}-\Theta) + K_n\sin(\theta_{n+2}-\Theta), & (6.22)\\
m_{n+2}\dot{\theta}_{n+2} &= m_{n+2}\omega_{n+2} + K_{n+1}\sin(\Theta-\theta_{n+2}) + K_{n+2}\sin(\theta_{n+3}-\theta_{n+2}). & (6.23)
\end{aligned}$$

In summary, the effect of the strong-coupling decimation step is to replace the pair of oscillators with phases θ_n and θ_{n+1} by a single oscillator with phase Θ, frequency ω' and mass M respectively given by Eqs. (6.3), (6.16), and (6.17).

6.3 Crazy Oscillator Step

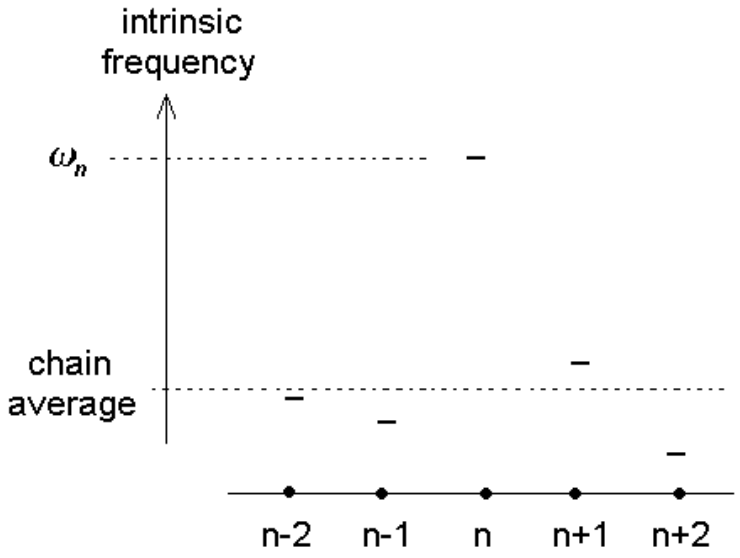


Figure 6.2: A small piece of the chain around the oscillator with the largest frequency, labelled here as Ω. All the neighboring K_i and all the neighboring ω_i are assumed to have much smaller magnitude then ω_n.

Consider the oscillator with the largest frequency in the entire chain: suppose this

is oscillator n. We use Ω for its frequency ω_n to emphasize this is a large quantity. Because of strong disorder the coupling strengths K_{n-1} and K_n to the neighboring oscillators, and the frequencies $\omega_{n\pm1}$ of these oscillators will almost always be small compared to Ω. Associated with this, there are small parameters $K_{n-1}/\Omega, \omega_{n-1}/\Omega$, etc. In the present section we will use the symbol ϵ to denote quantities of this order. To zeroth order in ϵ the nth oscillator runs freely at the frequency Ω. We call such oscillators "crazy oscillators", Fig. 6.2. However, as in the case of the strong coupling, it is not the absolute magnitude of Ω that determines whether an oscillator is subjected to this step, but the $O(\epsilon)$ ratios such as K_{n-1}/Ω and ω_{n-1}/Ω: the phase of an oscillator with a large Ω, but strongly coupled to a slow neighbor or positioned next to another fast neighbor can not be assumed to advance freely (see also Section 6.3.1 for elaboration).

Deviations from the free-running solution can be calculated perturbatively in ϵ. The most important effect we might look for is an induced effective coupling via oscillator n between the neighbors $n-1$ and $n+1$ (similar to what we have seen in the strong-coupling step), since this might lead to their synchronization. However, at least to order ϵ^2 as we describe below, we find that there is no induced interaction of the type that tends to lock the neighboring oscillators. [2] Thus the main effect of decimating the largest frequency oscillator in the chain is to eliminate this oscillator from further consideration (it forms a cluster in the final tally at frequency close to Ω and with a size given by the mass parameter m) and *to cut the chain into two pieces.* This is a key feature of the strongly random chain that limits the size of synchronized clusters. At order ϵ^2 we do find a renormalization of the frequencies of both the eliminated oscillator and the two neighbors remaining in the chain. We now analyze the elimination of the crazy oscillator (oscillator n in Fig. 6.2) perturbatively. The

[2] This should be compared with the strongly disordered spin chain, where an effective interaction between the neighbors is given by eliminating the high energy spins.

EOM for oscillator n and the neighboring oscillators read:

$$\frac{d\theta_{n-1}}{dt} = \omega_{n-1} + \frac{K_{n-1}}{m_{n-1}} \sin(\theta_n - \theta_{n-1}) + \frac{K_{n-2}}{m_{n-1}} \sin(\theta_{n-2} - \theta_{n-1}) \quad (6.24)$$

$$\frac{d\theta_n}{dt} = \Omega + \frac{K_{n-1}}{m_n} \sin(\theta_{n-1} - \theta_n) + \frac{K_n}{m_n} \sin(\theta_{n+1} - \theta_n) \quad (6.25)$$

$$\frac{d\theta_{n+1}}{dt} = \omega_{n+1} + \frac{K_{n+1}}{m_{n+1}} \sin(\theta_n - \theta_{n+1}) + \frac{K_{n+2}}{m_{n+1}} \sin(\theta_{n+2} - \theta_{n+1}) \quad (6.26)$$

As Ω/K_n and Ω/K_{n-1} become very large, the dynamic of a crazy oscillator is given approximately by $\theta_n = \Omega t$. The corrections are caused by the interaction with the neighbors. There will be corrections to the average frequency at which θ_n grows, but there will also be new periodic terms that arise due to the periodic slow down and speeding up as the phase θ_n advances in the "washboard potential" of the slow neighbors. There will be other effects, for example, the neighbors pick up a fast component to their dynamics imparted by the oscillator n and this change will be reflected back in the dynamics of the nth oscillator itself. Regardless of the effect, we must make the ansatz that allows for both the change in the average growth rate of θ_n *and* the appearance of small new (periodic) terms. Thus, we let

$$\begin{aligned} \theta_n &= \Omega \left(1 + \frac{\alpha}{\Omega} + \frac{\beta}{\Omega^2} + ...\right) t + \frac{\psi}{\Omega} + \frac{\phi}{\Omega^2} + ... \quad (6.27) \\ &= \Omega' t + \frac{\psi}{\Omega} + \frac{\phi}{\Omega^2} + ... \ . \end{aligned}$$

This is the perturbation ansatz for θ_n. Occasionally we will write Ω' to save space rather then the series for which it stands. As for the dynamics of each of the neighbors, $\theta_{n\pm1}$, let $l(t)$ be the solution to $n-1$st oscillator when it is completely disconnected

from the crazy oscillator, i.e., this $l(t)$ satisfies:

$$\begin{aligned} \frac{dl}{dt} &= \omega_{n-1} + \frac{K_{n-2}}{m_{n-1}} \sin\left(\theta_{n-2} - l\right) && (6.28) \\ \frac{d\theta_{n-2}}{dt} &= \omega_{n-2} + \frac{K_{n-3}}{m_{n-2}} \sin\left(\theta_{n-3} - \theta_{n-2}\right) + \frac{K_{n-2}}{m_{n-2}} \sin\left((l - \theta_{n-2}\right) && (6.29) \\ &... \end{aligned}$$

(an analogue of $l(t)$ for $n+1$st oscillator will be called $r(t)$). Then, at finite K_{n-1}/Ω, the dynamic of oscillator $n-1$ is given by

$$\theta_{n-1} = l(t) - \frac{K_{n-1}}{m_{n-1}\Omega} \cos\left(\Omega t - l(t)\right) + O\left(\frac{1}{\Omega^2}\right), \tag{6.30}$$

as can be verified by direct differentiation and substitution of the crazy oscillator ansatz, Eq. (6.27) and the definition of $l(t)$, Eq. (6.30). This gives the fast correction imparted onto the slow motion of a nearest-neighbor of a crazy oscillator, and we will calculate the effect of this fast component of the motion of the neighbor back on CO.

We will now take the self-consistent, iterative approach to extract the first few terms in the perturbation series for θ_n. We have

$$\begin{aligned} \dot{\theta}_n &= \Omega\left(1 + \frac{\alpha}{\Omega} + ...\right) + \frac{\dot{\psi}}{\Omega} + ... \\ &= \Omega + \frac{K_{n-1}}{m_n} \sin\left(l(t) - \Omega\left(1 + \frac{\alpha}{\Omega} + ...\right)t\right) + \frac{K_n}{m_n} \sin\left(r(t) - \Omega\left(1 + \frac{\alpha}{\Omega} + ...\right)t\right). \end{aligned} \tag{6.31}$$

We made an assumption that ψ/Ω, etc. terms are much smaller then the terms we keep in the arguments of sines (we can not assume $\left(\frac{\alpha}{\Omega} + ...\right)t$ to be small because t can be large). We will calculate ψ and verify that the "smallness" assumption was correct, hence making the method self-consistent. Then integrating the equation while *treating* $l(t)$ *and* $r(t)$ *as constants* we have

$$\left(\frac{\alpha}{\Omega} + \frac{\beta}{\Omega^2} + ...\right)t + \frac{\psi}{\Omega} = \frac{K_{n-1}}{m_n\Omega} \cos\left(l(t) - \Omega' t\right) + \frac{K_n}{m_n\Omega} \cos\left(r(t) - \Omega' t\right) + O\left(\frac{1}{\Omega^2}\right). \tag{6.32}$$

Comparing terms we see that α must be set to zero, since it was defined to be a

constant, not a function of time. Then we have found the first order corrections ψ:

$$\psi = \frac{K_{n-1}}{m_n}\cos\left(l(t) - \Omega\left(1 + \frac{\beta}{\Omega^2} + ...\right)t\right) + \frac{K_n}{m_n}\cos\left(r(t) - \Omega\left(1 + \frac{\beta}{\Omega^2} + ...\right)t\right). \tag{6.33}$$

We now repeat the procedure:

$$\begin{aligned}
&\Omega\left(1 + \frac{\beta}{\Omega^2} + ...\right) + \frac{\dot{\psi}}{\Omega} + \frac{\dot{\phi}}{\Omega^2}... = \\
&\Omega + \frac{K_{n-1}}{m_n}\sin\left(l(t) - \frac{K_{n-1}}{m_{n-1}\Omega}\cos\left(l(t) - \Omega' t\right) - \Omega' t - \frac{K_{n-1}}{m_n\Omega}\cos\left(l(t) - \Omega' t\right) - \frac{K_n}{m_n\Omega}\cos\left(r(t) - \Omega' t\right)\right) \\
&+ \frac{K_n}{m_n\Omega}\sin\left(r(t) - \frac{K_n}{m_{n+1}\Omega}\cos\left(r(t) - \Omega' t\right) - \Omega' t - \frac{K_{n-1}}{m_n\Omega}\cos\left(l(t) - \Omega' t\right) - \frac{K_n}{m_n\Omega}\cos\left(r(t) - \Omega' t\right)\right) \\
&+ ... \,.
\end{aligned} \tag{6.34}$$

Again, in the spirit of the self-consistent approach, we neglect ϕ-terms from the arguments of the sines. We now expand the sines in $1/\Omega$ and cancel out the zeroth-order terms on the right-hand side with the $\dot{\psi}$ term on the left-hand side, leaving

$$\begin{aligned}
&\frac{\beta}{\Omega} + ... + \frac{\dot{\phi}}{\Omega^2} + ... \\
&= -\frac{K_{n-1}^2}{2m_n\mu_{n,n-1}\Omega}\left(1 + \cos\left(2l(t) - 2\Omega' t\right)\right) - \frac{K_{n-1}K_n}{m_n^2\Omega}\cos\left(l(t) - \Omega' t\right)\cos\left(r(t) - \Omega' t\right) \\
&- \frac{K_n^2}{2m_n\mu_{n,n+1}\Omega}\left(1 + \cos\left(2r(t) - 2\Omega' t\right)\right) - \frac{K_nK_{n-1}}{m_n^2\Omega}\cos\left(r(t) - \Omega' t\right)\cos\left(l(t) - \Omega' t\right) + ... \\
&= -\frac{K_{n-1}^2}{2m_n\mu_{n,n-1}\Omega}\left(1 + \cos\left(2l(t) - 2\Omega' t\right)\right) - \frac{K_n^2}{2m_n\mu_{n,n+1}\Omega}\left(1 + \cos\left(2r(t) - 2\Omega' t\right)\right) \\
&- \frac{K_nK_{n-1}}{m_n^2\Omega}\left[\cos\left(l(t) + r(t) - 2\Omega' t\right) + \cos\left(r(t) - l(t)\right)\right] + ...,
\end{aligned} \tag{6.35}$$

with $\mu_{i,j}^{-1} = m_i^{-1} + m_j^{-1}$. From this we learn two things. First, we read off β:

$$\beta = -\frac{K_{n-1}^2}{2m_n\mu_{n,n-1}} - \frac{K_n^2}{2m_n\mu_{n,n+1}} \tag{6.36}$$

Second, we read off an equation for ϕ:

$$\begin{aligned}\frac{\dot{\phi}}{\Omega} &= -\frac{K_{n-1}^2}{2m_n\mu_{n,n-1}}\cos\left(2l(t)-2\Omega' t\right)) - \frac{K_n^2}{2m_n\mu_{n,n+1}}\cos\left(2r(t)-2\Omega' t\right) \\ &- \frac{K_nK_{n-1}}{m_n^2}\cos\left(l(t)+r(t)-2\Omega' t\right) - \frac{K_nK_{n-1}}{m_n^2}\cos\left(r(t)-l(t)\right) \qquad (6.37)\end{aligned}$$

We will not proceed with integrating this; what we learned at this stage is that an equation for the crazy oscillator can be rewritten without explicitly containing the variable θ_n, at the expense of introducing extra terms:

$$\begin{aligned}\dot{\theta}_n = \Omega&\left[1 - \frac{K_{n-1}^2}{2m_n\mu_{n,n-1}\Omega^2} - \frac{K_n^2}{2m_n\mu_{n,n+1}\Omega^2} - \frac{K_nK_{n-1}}{m_n^2\Omega^2}\cos\left(r(t)-l(t)\right)\right] \\ &+ \frac{K_{n-1}}{m_n}\sin\left(l(t)-\Omega' t\right) + \frac{K_n}{m_n}\sin\left(r(t)-\Omega')t\right) \\ &- \frac{K_{n-1}^2}{2m_n\mu_{n,n-1}\Omega}\cos\left(2l(t)-2\Omega' t\right)) - \frac{K_n^2}{2m_n\mu_{n,n+1}\Omega}\cos\left(2r(t)-2\Omega' t\right) \\ &- \frac{K_nK_{n-1}}{m_n^2\Omega}\cos\left(l(t)+r(t)-2\Omega' t\right) + ..., \qquad (6.38)\end{aligned}$$

where

$$\Omega' = \Omega\left(1 + \frac{\beta}{\Omega^2} + ...\right), \qquad (6.39)$$

and β is given by Eq. (6.36).

We can define a long-time average rate at which θ_n advances as

$$\bar{\omega}_n \equiv \langle\dot{\theta}_n\rangle_t = \frac{1}{T}\int_{t_0}^{T}\dot{\theta}_n\,dt = \frac{\theta_i(t_0+T)-\theta_i(t_0)}{T}. \qquad (6.40)$$

This is very similar to the definition in Eq. (6.40), with one difference that T is some characteristic time of the slow dynamics (for example of $r(t)$ and $l(t)$). Over this time period, the parts that depend explicitly on $\Omega' t$ are much suppressed and we have

$$\bar{\omega}_n \approx \Omega\left[1 - \frac{K_{n-1}^2}{2m_n\mu_{n,n-1}\Omega^2} - \frac{K_n^2}{2m_n\mu_{n,n+1}\Omega^2} - \frac{K_nK_{n-1}}{m_n^2\Omega^2}\cos\left(\theta_{n+1}-\theta_{n-1}\right)\right], \qquad (6.41)$$

where we have also substituted $l(t) \to \theta_{n-1}$ and $r(t) \to \theta_{n+1}$ without introducing errors at the order considered. This is our final result for the crazy oscillator: we

started with the unperturbed motion given by $\dot{\theta}_n = \Omega$ and we calculated the long-time average after taking the effect of the neighbors into account to second order in $1/\Omega$, resulting in Eq. (6.41). Stepping back we see that there are two $O(\epsilon^2)$ effects that slightly perturb the average rate of the uniform advancement of the phase θ_n. The first is the fact that crazy oscillator rotates in the tilted washboard potential created by its neighbors. Since typically $\omega_{n-1}, \omega_{n+1} \ll \Omega$ for strong disorder, this potential can be treated as effectively static. The second effect is a small component of fast dynamics added to the motion of the neighboring oscillators, which then acts back on the crazy oscillator. Both of these effects change the average frequency of the crazy oscillator by a shift of order K^2/Ω. The term $\cos(r(t) - l(t))$ comes only from the first effect - the positions of each of the neighbors affects washboard potential, and hence, the slowing down and speeding up of the crazy oscillator.[3]

Corresponding corrections happen to the motion of the neighboring oscillators as well. Indeed, a similar procedure can be performed for either of CO's neighbors and we can obtain a renormalized equation for both $\theta_{n\pm1}$ which will not be explicitly coupled to θ_n. Thus, we will remove θ_n from the description of the chain! For example, when we substitute perturbation anzatz for θ_n, Eq. (6.27) and for θ_{n-1}, Eq. (6.30) (and analogously θ_{n+1}), then Eq. (6.24) for the dynamics of θ_{n-1} becomes, after all the algebra

$$\begin{aligned}
\dot{\theta}_{n-1} &= \omega_{n-1} + \frac{K_{n-1}^2}{2m_n\mu_{n,n-1}\Omega} + \frac{K_{n-1}K_n}{2m_n m_{n-1}\Omega}\cos(l(t) - r(t)) \\
&+ \frac{K_{n-1}}{m_{n-1}}\sin(\Omega' t - l(t)) - \frac{K_{n-1}^2}{2m_{n-1}\mu_{n,n-1}\Omega}\cos(2l(t) - 2\Omega' t) \\
&+ \frac{K_n K_{n-1}}{2m_{n-1}m_n\Omega}\cos(l(t) + r(t) - 2\Omega' t) + \frac{K_{n-2}}{m_{n-1}}\sin(\theta_{n-2} - \theta_{n-1}) + \dots \,.
\end{aligned} \tag{6.42}$$

As in the case of the crazy oscillator n, we drop the terms containing $\Omega' t$ because the effect of these terms is to introduce small oscillations to the dynamics θ_{n-1} on top of

[3]The somewhat different derivation which is based directly on calculating these two aforementioned effects, rather then a more formal manipulation of the equations of motion is presented in [101].

whatever dynamics this oscillator has without these terms (of course oscillator θ_{n-1} is coupled to other oscillators, for example to θ_{n-2}, so this nominal dynamics may be complicated). We then again make the substitution $l(t) \to \theta_{n-1}$ and $r(t) \to \theta_{n+1}$ to end up with

$$\dot{\theta}_{n-1} = \omega_{n-1} + \frac{K_{n-1}^2}{2m_n\mu_{n,n-1}\Omega} + \frac{K_{n-1}K_n}{2m_n m_{n-1}\Omega}\cos(\theta_{n-1} - \theta_{n+1}) + \frac{K_{n-2}}{m_{n-1}}\sin(\theta_{n-2} - \theta_{n-1}) + \dots . \tag{6.43}$$

Finally, we argue that the cosine term that couples oscillators $n-1$ and $n+1$ can be dropped. It turns out that the corresponding term in the equation for $\dot{\theta}_{n+1}$ has same sign! In other words, this new type of interaction does not have action-reaction symmetry. This sign peculiarity arises because the decimated oscillator has a particular direction of rotation. Due of the lack of action-reaction symmetry, the cosine term does not tend to pull neighboring oscillators together, and so does not directly influence the clustering. We therefore omit this term in the renormalization procedure. We will also ignore this term in the motion of the crazy oscillator Eq. (6.41) since it averages to zero on the long timescale of the period of oscillators $n \pm 1$ [longer timescale then the characteristic time T over which the average was performed to arrive at Eq. (6.41)].

So far, the crazy oscillator decimation step can be summarized as follows: (i) remove the crazy oscillator from the chain and set the interaction between its neighbors to zero; (ii) adjust the frequencies of each of the neighbors by

$$\delta\omega_{n-1} = \frac{K_{n-1}^2}{2m_{n-1}\mu_{n,n-1}\omega_n}, \ \delta\omega_{n+1} = \frac{K_n^2}{2m_{n+1}\mu_{n,n+1}\omega_n}, \tag{6.44}$$

with μ the reduced mass of the pair; (iii) adjust the frequency of the crazy oscillator

itself by the corresponding correction from each of the neighbors[4]

$$\delta\bar{\omega}_n = -\frac{K_{n-1}^2}{2m\mu_{n,n-1}\Omega} - \frac{K_n^2}{2m_n\mu_{n,n+1}\Omega}. \tag{6.45}$$

Two observations are in order. First, notice that the constant frequency renormalizations of the crazy oscillator is reciprocal to the sum of its neighbors. This is because at the pairwise level interactions between the oscillators conserve the mass-weighted frequency average: by adding all equations of motion we can see that $\sum_{i=1}^{N} m_i\delta\omega_i = 0$, which is just a consequence of the fact that interactions are odd. Moreover, at the level of approximation that we consider, $\delta\omega_i = 0$ for all but the crazy oscillator n and its nearest neighbors, $n-1$ and $n+2$. Hence $m_{n-1}\delta\omega_{n-1} + m_n\delta\omega_n + m_{n+1}\delta\omega_{n+1} = 0$.

Second, the fact that frequency correction terms (the first two terms in Eq. (6.41) are additive suggests that they can be obtained from the analysis of two 2-oscillator systems, each of which can be solved exactly. The phase difference $\phi = \theta_n - \theta_{n-1}$ of two coupled oscillators satisfies

$$\dot{\phi} = (\omega_n - \omega_{n-1}) - \frac{K_{n-1}}{\mu_{n,n-1}}\sin\phi. \tag{6.46}$$

When $\omega_n - \omega_{n-1} \leq \frac{K_{n-1}}{\mu_{n,n-1}}$ the phase difference ϕ is attracted to one of the stable fixed points and locking of two oscillators occurs. Otherwise, a solution with an ever-increasing ϕ takes place. When $(\omega_n - \omega_{n-1})$ is slightly greater then $\frac{K_{n-1}}{\mu_{n,n-1}}$, the dynamics of $\phi(t)$ consists of a series of long plateaus (the bottleneck part of the dynamics of ϕ) which periodically undergo a rapid slip: a "staircase" solution with some average growth rate. A period τ over which this phase difference grows by 2π

[4]The appearance of the overbar in $\delta\bar{\omega}_n$ and its lack in $\delta\omega_{n\pm1}$ is on purpose: while $\bar{\omega}_n$ represents the RG solution to the actual running frequency of the crazy oscillator, the frequency correction to oscillators $n-1$ and $n+1$ is a renormalization of the parameters that appears in the equations of motion of these oscillators, but does not yet represent the final value of the running frequency of those oscillators.

can be defined by

$$\begin{aligned}\tau &= \int_0^{2\pi} \frac{d\phi}{(\omega_n - \omega_{n-1}) - \frac{K_{n-1}}{\mu_{n,n-1}} \sin\phi} \\ &= \frac{2\pi}{\sqrt{(\omega_n - \omega_{n-1})^3 - \left(\frac{K_{n-1}}{\mu_{n,n-1}}\right)^2}}. \end{aligned} \tag{6.47}$$

So

$$\phi = \theta_n - \theta_{n-1} = (\omega_n - \omega_{n-1})\sqrt{1 - \left(\frac{K_{n-1}}{\mu_{n,n-1}(\omega_n - \omega_{n-1})}\right)^2}\, t + \text{oscillating terms.} \tag{6.48}$$

Also notice that that $m_{n-1}\dot{\theta}_{n-1} + m_n\dot{\theta}_n = m_{n-1}\omega_{n-1} + m_n\omega_n$, so

$$\Theta \equiv m_{n-1}\theta_{n-1} + m_n\theta_n = (m_{n-1}\omega_{n-1} + m_n\omega_n)t. \tag{6.49}$$

From Eqs. (6.48) and (6.49) we see that $m_i\dot{\theta}_i = [m_i(\omega_i + \delta\omega_i) + \text{oscillating terms}]$ where

$$\delta\omega_i = \frac{\mu_{i,j}}{m_i}(\omega_j - \omega_i)\left[1 - \sqrt{1 - \left(\frac{K_{i,j}}{\mu_{i,j}(\omega_j - \omega_i)}\right)^2}\right], \tag{6.50}$$

and i and j are the two oscillators in question (n and $n-1$). Note the reciprocity: $m_i\delta\omega_i = -m_j\delta\omega_j$.

This was a 2-body system: not surprisingly a fully solvable case. Going back to the many-body system of a crazy oscillator and the rest of the chain, we will neglect corrections due to second nearest neighbors and beyond. Then the frequency corrections are simply the superposition of two 2-body effects. In other words, each of the nearest neighbors of the crazy oscillator receives a frequency correction given by Eq. (6.50), and its own frequency correction is given as

$$\delta\bar{\omega}_n = -\frac{m_{n-1}}{m_n}\delta\omega_{n-1} - \frac{m_{n+1}}{m_n}\delta\omega_{n+1}. \tag{6.51}$$

If the square roots are expanded to second power in $1/\Omega$, we recover exactly the same

frequency corrections summarized in Eqs. (6.44)-(6.45). Square root formulas of this type include some of the higher-order effects (such as $\Omega - \omega_i$ in the denominator, rather then Ω). There may be other terms of this order, for example $\omega_{n-1}\omega_{n+1}/\Omega^2$, which can not be obtained from a pair of independent 2-oscillator analysis. Therefore, if used in the full oscillator chain, these expressions form an uncontrolled approximation at orders higher then $O(1/\Omega)$, but empirically we found them to produce a slightly better match with the simulation data (below) in some cases; we explore this in the following subsection. Hence, in the numerical renormalization of the chain, we employ the square root formulas of the type of Eq. (6.50).

6.3.1 Comments on the Crazy Oscillator Step

Now is the appropriate time to reflect on what it means for an oscillator to be crazy. Notice that $\dot{\theta}_n/\omega_n \equiv 1 + (\text{K-terms})$, where the "K-terms" are terms which become zero when the couplings to the neighbors, K_{n-1} and K_n are set to zero. Based on what we just learned, an oscillator is crazy when these K-terms are $\ll 1$. This is just a definition; we still need to specify what causes an oscillator to possess this property. Based on this definition, for an oscillator which is crazy, setting the couplings to the neighbors to zero results in a small correction to $\dot{\theta}_n/\omega_n$, while for an oscillator which is not crazy, the correction can be substantial. From the reference frame of the oscillator in question, when the neighbors have significantly different frequencies relative to the couplings, their pushing and pulling on θ_n averages out to approximately zero. This is seen directly from the 2-oscillator analysis, Eqs. (6.46)-(6.50). Hence, setting the coupling to zero will have a small impact on $\dot{\theta}_n$. Thus, it is the frequency difference relative to the couplings that determines whether an oscillator is crazy. Indeed, the physics that dictates whether the K-terms are small must clearly be independent of the reference frame; we can always switch to the frame in which all ω values are large relative to couplings, so all oscillators appear fast - but not all are crazy! Therefore, only the frequency differences must enter into the criterion. We define a parameter r_{CO} as a measure of craziness of an oscillator, or the measure of validity of the

corresponding decimation step:

$$r_{CO} \equiv \text{Max}\left[\frac{K_n}{\mu_{n,n+1}|\omega_n - \omega_{n+1}|}, \frac{K_{n-1}}{\mu_{n,n-1}|\omega_n - \omega_{n-1}|}\right] \tag{6.52}$$

Our discussion of the 2-oscillator problem shows that when $r_{CO} > 1$, the frequencies of the two oscillators entrain, and for $r_{CO} < 1$, the corrections due to the interaction quickly decays, due to the rapid variation of the square root for the small value of its argument. It is not obvious what the craziness criterion should be in the case of a crazy oscillator embedded in a many-oscillator case. Numerical experiments, such as the ones we now present suggest that $r_{CO} < 1$ is still a reasonable criterion, so we use it in our implementation of the RG.

Demonstrations

First, consider a system of three oscillators, labelled as oscillators 3, 4, and 5 in Fig. 6.3 (appears at the end of this chapter). The frequencies of oscillators 3 and 5, as well as the couplings $K_{3,4}$ and $K_{4,5}$ were fixed, while a series of numerical simulations of Eqs. (6.2) were performed, one for each different value of $\omega_4 \equiv \Omega$. The resultant frequencies of each of the oscillators, defined by Eq. (6.40) are measured. This is the content of Fig. 6.3 (a), along with comparisons with the square-root predictions, such as Eqs. (6.50),(6.51) and the simpler formula, such as Eq. (6.45).

We see that as Ω becomes large, the average running frequency of oscillator 4, $\bar{\omega}_4$ approaches its intrinsic frequency Ω and the average running frequencies of the other oscillators $\bar{\omega}_i$ also approach their own intrinsic frequencies ω_i. We also see that this approach is rather rapid and closely follows the square-root predictions.

The detail of the low-Ω region is shown in Fig. 6.3 (b) along with the average frequencies of all three oscillators, and of the pairs of oscillators 3-4 and 4-5. As Ω is varied from negative to positive values, first oscillator 3 and 4 form a cluster (point A), then at even lower Ω all three oscillators entrain to the average frequency (point B),[5] followed by oscillator 3 breaking off from the three-oscillator cluster (point C)

[5] as is expected, since this is the phase-locking situation.

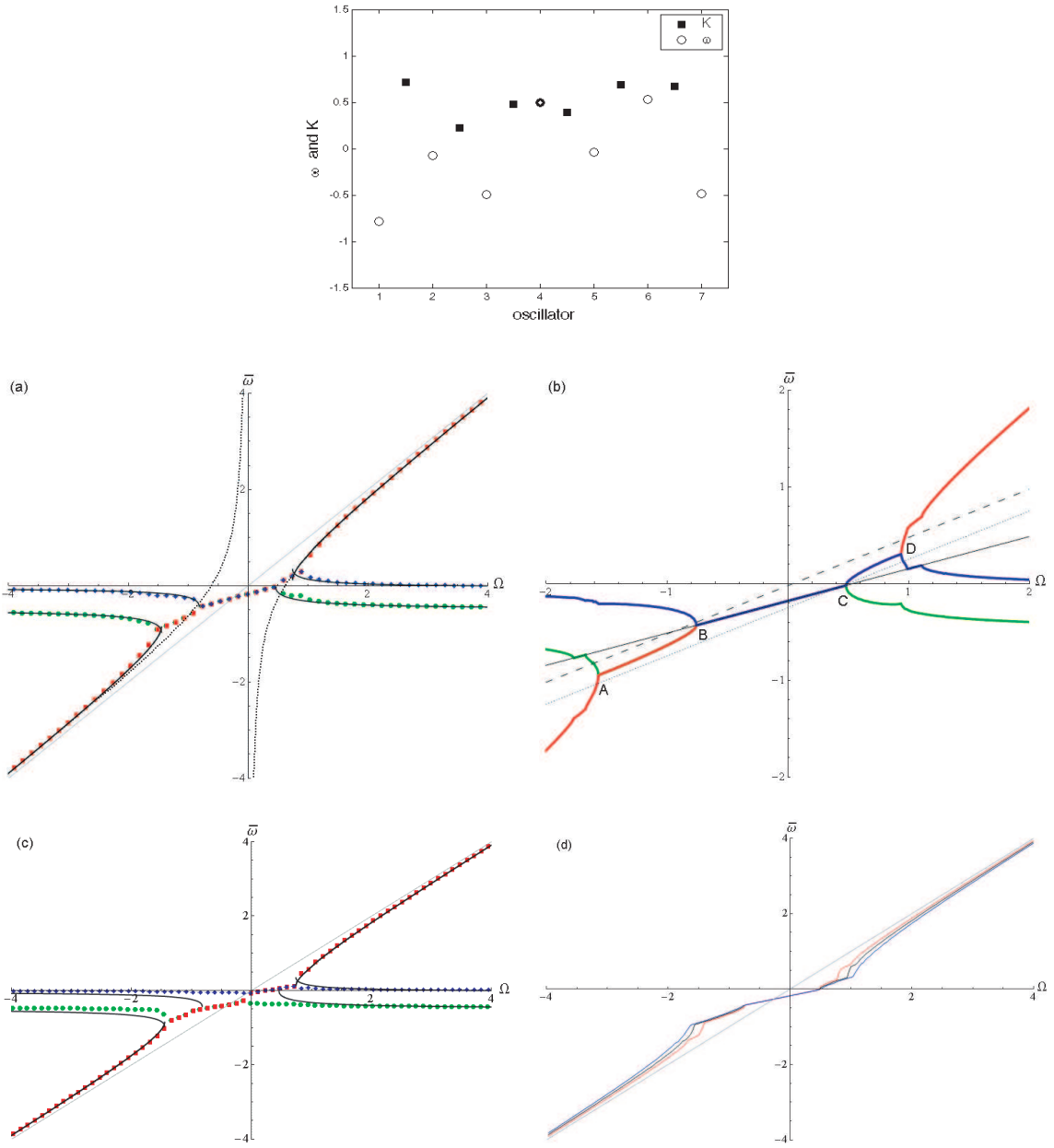


Figure 6.3: Top: the set of intrinsic frequencies and couplings. The frequency of oscillator 4 has been varied. (a) $\omega_i(\Omega)$ for oscillator 3 (green circles), oscillator 4 (red squares) and oscillator 5 (blue diamonds). The square-root approximations, Eqs. (6.50),(6.51 are shown by the thick curves, the simple formula such as Eq. (6.45) appears as the dotted curve and the function $\bar{\omega} = \Omega$ appears in light grey. (b) Close-up of the low-Ω region. Also shown are the average frequencies of all three oscillators (thin solid line), of oscillators 3 and 4 (dotted line) and of oscillators 4 and 5 (dashed line). (c) Same as (a) when oscillators 3, 4, 5 are embedded into a 7-oscillator system with frequencies and couplings specified in the top plot. (d) $\bar{\omega}_4(\Omega)$ for three different functional forms of the interaction. The curve closest to $\bar{\omega} = \Omega$ grey line is the saw-tooth interaction, the one furthest is the 4-th order polynomial interaction and the middle curve is the sine interaction, i.e. same as in (a).

and finally the 4-5 oscillator cluster breaking up (point D). Moreover, the point of breakout of oscillator 4 from both of its neighbors (points A and D) in Fig. 6.3 (b) is close to the point when r_{CO} becomes 1, as the comparison with Fig. 6.3 (a) indicates.

We also embedded the three oscillator system into a larger system with two additional oscillators on each side. The $\bar{\omega}_i(\Omega)$ curves for this situation appears in Fig. 6.3 (c). While the detailed features change somewhat in the low-Ω regions, the rest of the findings qualitatively remain the same. We note that not all realizations of frequencies and couplings would result in a qualitatively equivalent $\bar{\omega}_i(\Omega)$ plot. For example, it is possible to come up with such values of parameters for which all three oscillators are not entrained for any value of Ω. But it appears that it always remains true that $\bar{\omega}$ of a crazy oscillator approaches Ω for a relatively small values of Ω.

We hypothesized that one of the reasons for this quick approach to the crazy oscillator regime is the steep decay of the square root correction. The square root comes from the slowing down as the oscillator passes through the bottleneck of the washboard potential, see Eqs. (6.46) and (6.47). If this is so, changing the waveform to increase the duration of this bottleneck should result in a slower approach to the crazy regime, and changing the waveform to decrease the duration of the bottleneck should have the opposite effect. To check this hypothesis made numerical simulations of the 3-oscillator case for three different types of waveforms: sine, the results for which are already shown, as well as the sawtooth which should decrease the bottleneck slowing down and the periodic quartic polynomial which should enhance the slowing down. To avoid unnecessary clutter, we define these wave forms graphically in Fig. 6.4.

The comparison of the results appears in Fig. 6.3 (c). Interestingly, our expectations appear to be confirmed: the sawtooth wave leads to the quickest approach to the crazy oscillator regime, while the quartic polynomial interaction leads to the slowest approach. In fact, the analytical form for the period of ϕ in the case of the

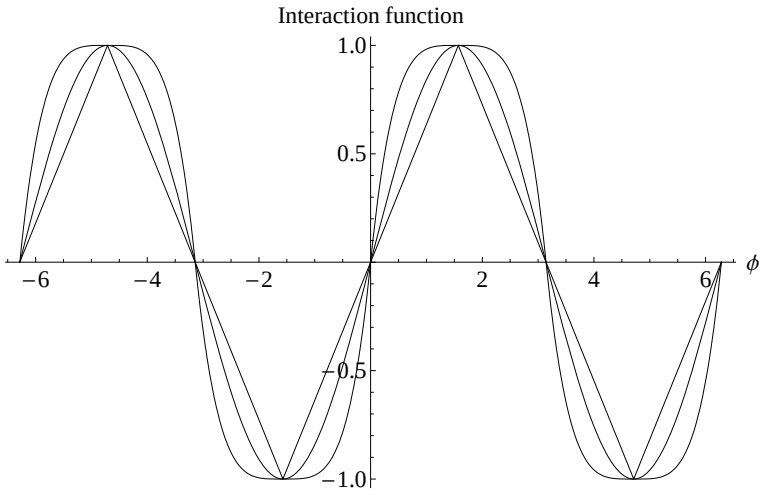


Figure 6.4: Triangle wave, sine (middle curve) and quartic polynomial interaction (outer curve).

sawtooth interaction $Saw(\phi)$ is logarithmically diverging:

$$\begin{aligned} \tau_{Saw} &= \frac{1}{\Delta\omega}\int_0^{2\pi}\frac{1}{1-\kappa Saw(\phi)}\,d\phi \\ &= \frac{\pi}{\kappa\Delta\omega}\ln\left(\frac{1+\kappa}{1-\kappa}\right), \end{aligned} \tag{6.53}$$

where $\kappa = \frac{K}{\mu\Delta\omega}$. For any value of κ this τ is smaller then $\tau_{Sin} = \frac{2\pi}{\Delta\omega\sqrt{1-\kappa^2}}$, the functional form in the case of the sine-wave interaction, so the slowing down of ϕ and decrease of its $\bar{\omega}$ relative to Ω is indeed smaller.

6.4 Renormalization Group Implementation

The chain of oscillators is renormalized by successive application of the two decimation steps developed above. These decimation steps are executed numerically on a list of parameters (m_i, ω_i, K_i) representing the chain of oscillators, beginning from the largest values of K and ω and working down to smaller ones during successive steps. Initial values of masses in the RG are all set to 1. As already mentioned, for the purpose of maintaining strong disorder which forms the assumption for the decimation steps, intrinsic frequencies and couplings in Eqs. (6.2) will be taken from

"wide" distributions, i.e., distributions with long tails. The numerical procedure identifies the oscillators with ω and K that lie in a narrow band of magnitudes at the top of the spectrum of remaining values, and then decimates those oscillators according to the steps defined in Sections 6.2 and 6.3. A band is used, rather than just selecting the largest values, to improve the efficiency of the code. The width of the band is chosen to be narrow (1% of the chain) to maintain the descending order in ω and K for the decimation of nearby oscillators. Oscillators decimated in the crazy oscillator step form a cluster in the final tally, with the effective mass describing the size of the cluster and the frequency giving the time-averaged rate at which the phase of this cluster advances. In decimating these oscillators, we use the squareroot approximations of the type of Eq. (6.50). As described in Section 6.3, these oscillators are removed from the chain. After the two decimation steps on the band of oscillators, the remainder of the chain consists of fewer oscillators, which have lower coupling strengths and frequencies within a narrowed spectrum. The procedure is then repeated for successively lower bands of K and ω.

During the execution of the RG, we keep track of the values of r_{SC} and r_{CO}. For most cases in the beginning of the RG procedure, these are not close to 1 because we work with distributions that exhibit strong disorder. In other words, given a potential crazy oscillator, in most of the cases the nearby ω and K are small in comparison: this assumption defines the physical content of the strong disorder. However, it is possible to have neighbors with parameters which are somewhat lower, yet the difference is not large enough to guarantee that effects of these neighbors are small. This is especially important as the spectrum of the remaining K and ω values shrinks. This is where the importance of r_{SC} and r_{CO} comes in. As we explained in Sections 6.3 and 6.2, we set set the criteria to $r_{CO} <$ and $r_{SC} > 1$.

Aside from some r_{CO} and r_{SC} being close to 1, we occsionally encounter another special case related to violations of the assumptions based on strong disorder. Note that single oscillators with $m = 1$ or oscillators representing clusters with $m > 1$ are normally subjected to the crazy oscillator decimation step at some point in the RG procedure: most of the bare oscillators either join other oscillators in building

clusters, which are eventually decimated via the crazy oscillator step, or they are decimated as crazy oscillators directly without undergoing a strong-coupling step. However, there exist rare combinations of frequencies and couplings that will prevent the decimation from occurring at all. For example, while an oscillator's frequency may lie in the executable band, this oscillator may be coupled via a K which is outside this band (it has a somewhat lower magnitude), but may nevertheless cause r_{CO} to exceed 1. In general, when the RG reaches the scale of that coupling, it will perform a strong-coupling decimation step. However, in rare cases this may not happen due to intervening steps that have modified the neighboring parameters, rendering that coupling no longer strong according to the r_{SC} criterion. As a result, the site in question has not been subjected to either decimation step. To address such loopholes, the algorithm sweeps repeatedly through the entire spectrum of ω and K values until nothing is left to decimate.

6.4.1 Strain Check

This subsection introduces an additional piece of the RG algorithm which we call "strain check". Both of the decimation steps just presented are local: the decision as to which step to implement and the outcome of each step are based on frequencies and couplings in the immediate neighborhood of one oscillator or one bond. As clusters are built up, only two pieces of collective information about the constituents of the cluster are retained: the average frequency of the constituents' ω' and the total number of oscillators in that cluster m; all other internal information is forgotten. However, it is not obvious that only ω' and m is enough to accurately predict the properties of the clusters. We propose an additional measure, the phase strain of a cluster, not unlike the strain of a material under stress.

It was mentioned in the work of Strogatz and Mirollo [51] that a chain of size N will phase synchronize, i.e., will allow a solution for θ_1 through θ_N such that

$\theta_{i+1} - \theta_i = const$ for all $i \in [1, N]$, only when the quantity

$$X_n = \sum_{i=1}^{n} (\omega_i - \langle\omega\rangle) < K_n, \tag{6.54}$$

for all $1 \leq n \leq N$. Here $\langle\omega\rangle$ is the average of all bare frequencies of the constituent oscillators. This conclusion was reached by adding up equations of motion for oscillators 1 through n and making use of the fact that all but the last interaction terms, $K_n \sin(\theta_{n+1} - \theta_n)$ cancel each other out. In order for phase synchronization to happen, the magnitude of the coupling K_n must be greater then or equal to X_n for every $1 \leq n \leq N$. Notice that X_n is a sum of random numbers, so for larger chains, it is increasingly likely that X_n will exceed K_n at some oscillator n (see [52] for calculation of probability that $X_n < K_n$ for all $1 < n < N$ in the case of all identical $K_n = K$); as the size of the cluster increases, the range of values of $|X_n|$ typically grows, and increasing phase strains $|\theta_{n+1} - \theta_n|$ are needed to keep the various parts of the cluster synchronized. This X_n, which we call "accumulated randomness" serves as a measure of a "strain" of a potential phase-synchronized cluster. When $X_n < K_n$ for all n, there will be a solution such that all local phase differences lie in $[-\pi/2, +\pi/2]$. If the strain becomes so large that some $|\theta_{n+1} - \theta_n|$ becomes $\pi/2$, any further increase of the strain will cause the cluster to break.

Also note that if two clusters are combined through a strong-coupling decimation step to form a putative larger cluster, the mean frequency $\langle\omega\rangle$ to be used in the sum in Eq. (6.54) changes, and the condition $|X_n| < K_n$ may become violated for a bond in the interior of one or other of the component clusters. This means that the combined set of oscillators are not, in fact, synchronized, and should be broken into smaller synchronized clusters in the RG procedure. We proceed by making the assumption that the oscillators form just two synchronized clusters, and identify the break between the two clusters as where the condition $|X_n| < K_n$ is violated in the presumed cluster.[6] Thus if X_n exceeds K_n at some n, we do not allow the cluster to

[6]Note, it is easy to see that if X_n exceeds K_n only at one place in the cluster, it does not matter from which end the quantity X_n is computed. If X_n exceeds K_n at more then one place, this only tells us that the hypothetical cluster would not phase synchronize, but it does not tell where it would

form. Instead, we reform the bare oscillators into two different subclusters, such that the first subcluster is formed out of the bare oscillators 1 through n, and the second is formed form the remaining bare oscillators. The RG then proceeds as normal. The new subclusters may be subjected to either of the two decimation steps later in the sweep, or in the next sweep if more are required.

The strain may appear as incompatible with the idea of the RG because its determination requires the complete information about the constituent oscillators, rather then some renormalized degree of freedom which hides microscopic details of the original system. However, we propose the strain as a way to check whether additional information beyond ω' and m is necessary. If it turned out that a strain check improves the correspondence of RG and simulation results, it would suggest that we are on the right track and motivate us to come up with some other method that measures strain but does not require the knowledge of the oscillators in the original chain. In our numerical implementation of the RG, the use of the strain check was optional in order to compare predictions with and without it. We will see in Section 7.2.4 that the use of the strain check does not seem to provide an obvious improvement in the correspondence of the RG and simulation data.

---

break up. In our algorithm, the break was made at the first place where X_n exceeded K_n.

Chapter 7

Results

7.1 Testing the validity of the numerical RG

We now switch the discussion from the general RG algorithm applicable to a wide class of distributions and parameters to specifics of the RG and simulations as they were used in this work: these include a particular choice of frequency distributions $g(\omega)$ and coupling distributions $G(K)$ as well as other details such as system size, the set of parameters defining the distributions that are investigated, and cutoff values in ω and K. Intrinsic frequencies will be assumed to have a symmetric Lorentzian distribution because they exhibit long tails:

$$g(\omega) = \frac{C_1}{1+\omega^2}, \tag{7.1}$$

with $-\omega_c \leq \omega \leq \omega_c$. The couplings will be positive and taken from the half Lorentzian for the same reason:

$$G(K) = \frac{2\mu C_2}{\mu^2 + K^2}, \tag{7.2}$$

with $0 \leq K \leq K_c$. The C_1 and C_2 are normalization constants that become $1/\pi$ when the cutoff values ω_c and K_c are infinite (i.e., no cutoff). Notice that due to the structure of Eqs. (6.2), it is always possible to divide both equations by a constant such that the width of one of these distributions will be unity (assuming the width of that distribution is not zero to begin with) - this only changes the timescale of all the processes in the chain. So, without loss of generality we chose to define $g(\omega)$ to have

the width of unity and explore the physics by varying the other width parameter μ.

7.1.1 RG Analysis of Lorentzian Chains

The RG procedure is applied to a chain of 10^6 oscillators. Twelve values of coupling width μ are studied: $\mu = 0.025$, 0.05, 0.25, 0.5, 0.625, 1.25, 2.5, 3.75, 4.5, 5, 6.25, and 7.5. The set of random numbers defining the particular realizations of the chain were different in the RG and in the simulations (below) for the statistical comparisons, but were identical for the real-space comparison described in Section 7.1.3. In order to facilitate a better comparison with simulations, we chose the same cutoff values for $\{\omega\}$ and $\{K\}$ as in the simulations: ω values were chosen to lie in $[-100, 100]$ and K values were chosen to lie in $[0, 100]$.

7.1.2 Simulations

To better understand the system characteristics and the reliability of the renormalization group, Eqs. (6.1) are also integrated numerically. The numerical method is a variable stepping Runge-Kutta algorithm. Systems of $N = 10000$ oscillators are solved with N intrinsic frequencies chosen randomly from the symmetric Lorentzian given by Eq. (7.1). Similarly, $N - 1$ coupling constants are randomly selected from the half Lorentzian given by Eq. (7.2). For both distributions the same cutoff value is used: $\omega_c = 100$ and $K_c = 100$; increasing cutoff values increases the time needed for simulations. The same twelve values of coupling width μ that are used in the RG are studied in the simulations: $\mu = 0.025$, 0.05, 0.25, 0.5, 0.625, 1.25, 2.5, 3.75, 4.5, 5, 6.25, and 7.5. Results for each value of μ are averaged over 100 realizations of the intrinsic frequencies and coupling constants. To help facilitate comparisons between different distribution widths the same 100 random number seeds are used for each coupling width.

The simulations are integrated for a relatively long time of $t = 10000$. Running phases at each site are recorded at regular intervals and used to calculate average

frequencies over a time T according to

$$\bar{\omega}_i(T) = \frac{\theta_i(t_0 + T) - \theta_i(t_0)}{T}; \tag{7.3}$$

with $T = t - t_0$ [compare with the theoretical definition of $\bar{\omega}$ given by Eq. (6.40)]. The time t_0 is chosen to eliminate transients: increasing values are tried until average frequencies remain unchanged. In these simulations, the value of $2\pi/T$ sets the resolution limit in $\bar{\omega}_i$ to distinguish two neighboring frequency clusters. A group of oscillators is determined to be members of a synchronized cluster only if all the members ($i = n, n+1, \ldots, n+m+1$) have the same value of $\bar{\omega}_i$ as the neighboring sites within some tolerance $|\bar{\omega}_\ell - \bar{\omega}_i| \leq \eta$ ($\ell = i - 1$ and $i + 1$). Since T^{-1} is $O(10^{-4})$ the tolerance η is set to 10^{-3}.

7.1.3 Comparison of RG and Numerics

Insight into the validity of this RG can be found by careful comparison of the frequency cluster structure that it predicts with results from numerical integration of Eqs. (6.2). If the renormalization method is wrong there is expected to be little agreement between the two. Figure 7.1 plots these predictions over a representative sample of 150 oscillators from chains of $N = 10000$ at 3 different coupling distribution widths. For all 3 values of μ, the RG predicts clusters with large frequency oscillators interspersed. As expected, the characteristic size of these clusters grows with μ. Also shown are time-averaged frequencies, Eq. (7.3), calculated from numerical solutions of Eqs. (6.2). There is excellent agreement of the cluster sizes, locations, and frequencies between the two approaches. This provides some confidence that the essential mechanisms are captured by the RG procedure.

Analysis of these and similar plots, shows that the agreement of the cluster frequencies becomes more accurate with increasing cluster size. We do in fact expect the dynamics of most of the oscillators in large cluster to be very similar to the dynamics if that cluster was isolated from the rest of the chain. The reason for this is the smaller boundary-to-bulk ratio of larger clusters, which is more apparent if Eq. (6.2)

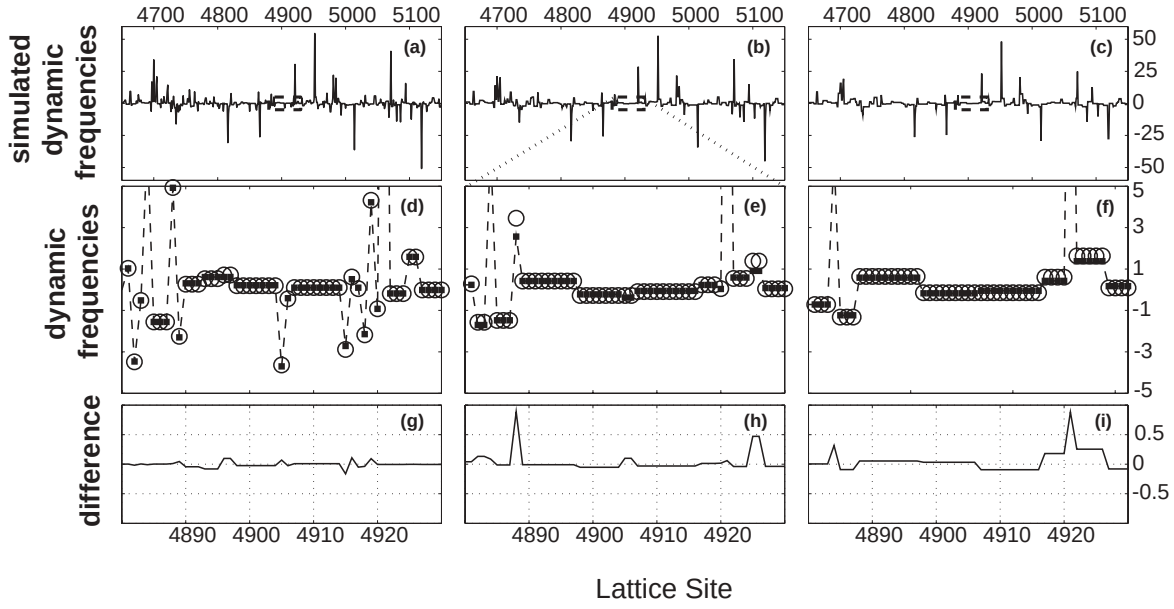


Figure 7.1: Cluster structure of a small section of a 10000-oscillator chain. The three columns correspond to $\mu = 1.25$, $\mu = 3.75$ and $\mu = 7.5$ respectively. The top row displays 500 time-averaged running frequencies $\bar{\omega}$ from simulations of Eq. (6.2). The middle row zooms in on a smaller section of 50 oscillators and compares $\bar{\omega}$ from simulations (solid squares) and from the RG (open circles). The bottom row displays the difference between the frequencies $\bar{\omega}$ of the RG and the simulations for the same 50 oscillators as in the middle row.

is rewritten as

$$\dot{\theta}_i = \omega_i + \frac{K_{i-1}}{m_i} \sin(\theta_{i-1} - \theta_i) + \frac{K_i}{m_i} \sin(\theta_{i+1} - \theta_i). \tag{7.4}$$

As clusters become larger, the importance of the coupling to the rest of the chain is weighted down by a $1/m$ factor. In the crazy oscillator decimation, the boundary-to-bulk ratio is manifested by the appearance of the $1/m$ factor in the frequency correction. Notice that in time-averaging Eq. (6.2), the influence of $1/m$ suggests that

$$\bar{\omega}_i \equiv \langle \dot{\theta}_i \rangle_t \approx \omega_i, \tag{7.5}$$

but ω_i is given by the mass-weighted average of the constituent bare ω values. Also, when a cluster is built up by a strong-coupling decimation step, any frequency renor-

malizations of the component oscillators by prior fast oscillation decimation steps cancel when the weighted sum of the frequencies is calculated to give the frequency of the new cluster. Thus the cluster frequencies $\bar{\omega}_i$ are given approximately by the average of the constituent bare ω values, plus the small correction due to the crazy oscillator decimation step that yields the final cluster. This is what we see in comparison of RG with simulations: the $\bar{\omega}$ of clusters in simulations is accurately predicted by the RG predictions, which are mostly given as by the mass-weighted average of the constituent bare ω values (plus the small correction due to the crazy oscillator decimation step).

The agreement with exact numerics in a small chain gave us reassurance that the RG method works, at least in some parameter regimes. Investigating the frequency distribution of clusters of a given size provides further insight into the validity of this RG procedure, but this discussion is postponed until Section 7.2.4.

7.2 Physics of the Unsynchronized Phase

We are now ready to turn to the prime goal of this work - understanding the behavior of the random oscillator chain, and its universal aspects. We first use the results of the RG and the simulations to discuss the statistical properties of the frequency clustering in the one-dimensional oscillator chain with strong randomness.

7.2.1 Cluster Size Distribution

The central quantity for understanding the behavior of the random oscillator chain is the distribution of cluster sizes. In Fig. 7.2 we show the distribution of cluster sizes for several widths μ of the coupling distribution from the RG and simulations. The RG was performed on one single realization of 10^6 oscillators, while the simulations were performed on 100 realizations of 10^4 oscillator chains. Since the characteristic cluster size is much smaller then 10^4, we combine results from all 100 realizations and analyze them as a single chain of size 10^6. As can be seen in Fig. 7.2, good agreement exists between the RG and the simulations. The apparent discrepancies at the largest

cluster sizes are due to statistical fluctuations resulting from the small number of such clusters present in the 10^6 oscillator ensemble.

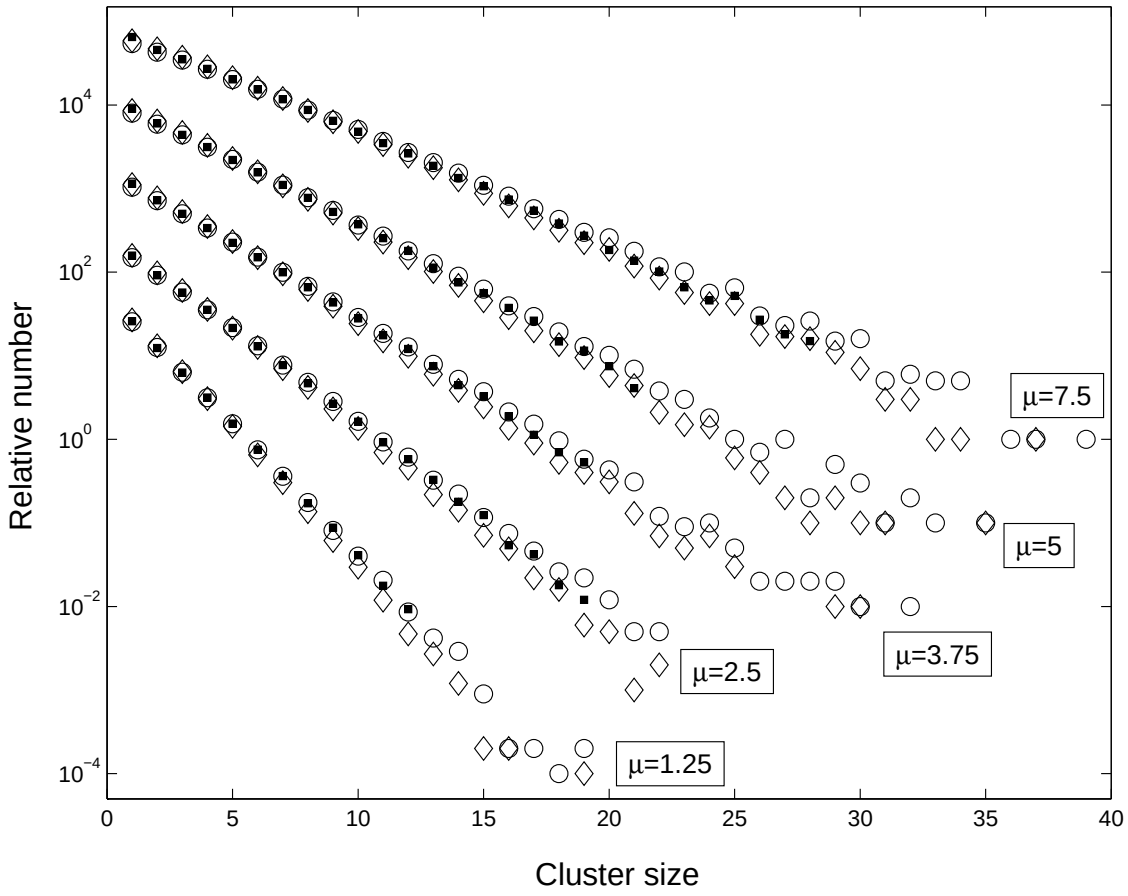


Figure 7.2: Number of clusters of a given size versus the size. Comparison between simulations of Eq. (6.2) (solid squares), the RG without strain check (open circles) and the RG with the strain check (diamonds). The curves are spread out for clarity by multiplying each successive data set as μ decreases by 0.1.

For a given value of the coupling width parameter μ, the probability of finding a cluster is predicted to fall off rapidly with increasing cluster size. The linear scaling in this semilog plot suggests that distribution of cluster sizes has the form of an exponential: $P(n) \propto \exp(-n/\xi)$, where $P(n)$ is the number of clusters of a given size, for example. The characteristic length ξ was extracted by making a linear fit to the data (not shown in Fig. 7.2) between twice the estimated characteristic size ξ (so it is an iterative process), and the cluster size at which there are 20 occurrences

of clusters of that given size.[1] The following section focuses on the dependence of ξ with μ.

We also investigated the effect of the larger system size. In Fig. 7.3 we show comparison of the RG for a system of a one million and ten million oscillators. Aside from continuing the exponential trend for larger cluster sizes, there does not seem to be any significant extra information conveyed by the larger system size.

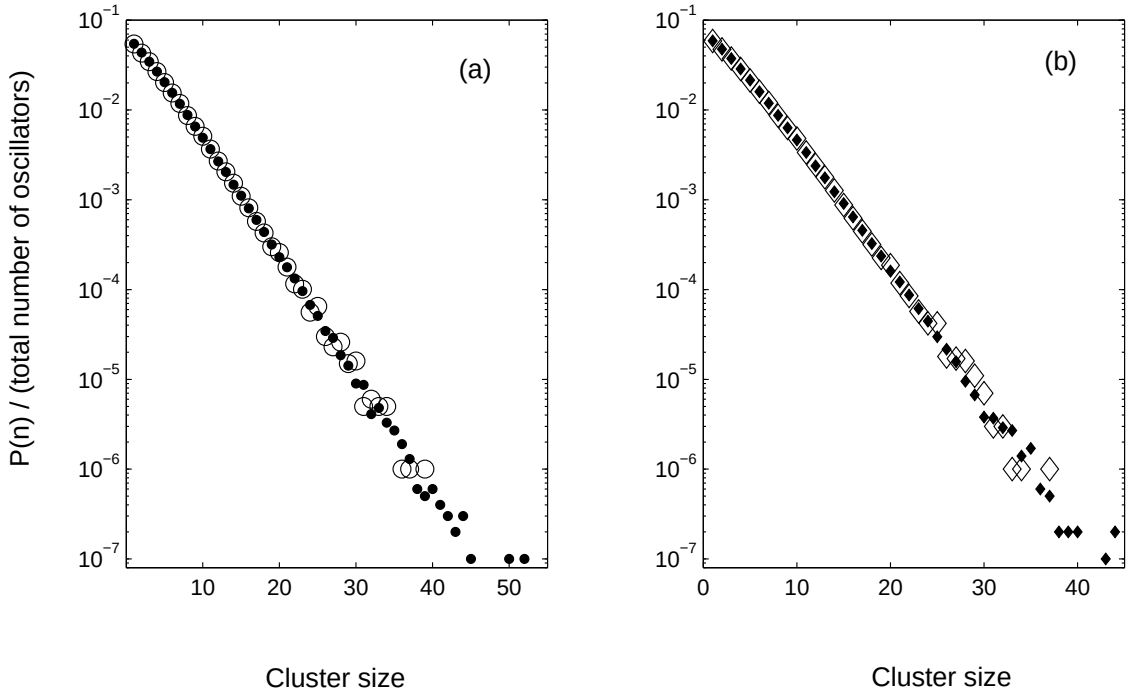


Figure 7.3: Number of clusters of a given size $P(n)$, normalized by the system size for $\mu = 7.5$ at two system sizes: 10^6 oscillators (open symbols) and 10^7 oscillators (smaller filled symbols). (a) RG without strack check and (b) RG with the strain check.

The values of ξ were obtained as described in the previous paragraph. We find that in the system with 10^6 oscillators, $\xi = 3.28$ and $\xi = 3.05$ without and with the strain check respectively, while in the system with 10^7 oscillators, $\xi = 3.26$ and $\xi = 2.97$ without and with the strain check respectively.

The distributions $P(n)$ are not strictly exponential, but they appear to be asymp-

[1] since the standard deviation is expected to be of order of the square root of the number of occurrences.

totically so, as seen from the plot of $P(n)$ for a system of 10^7 oscillators and the fit which was made between twice the characteristic length (7) and the cluster size 35 at which there are 27 clusters of that size: see the low-n part in Fig. 7.4 (a). Also shown in Fig. 7.4 (b) is the plot of slopes of partial fits, i.e., fits to $P(n)$ between cluster size m and cluster size 35. Aside from the low-statistics scatter at the end, we see that the variation in this data is much smaller then the average, and the data appears to be constant, indicating asymptotically exponential $P(n)$.

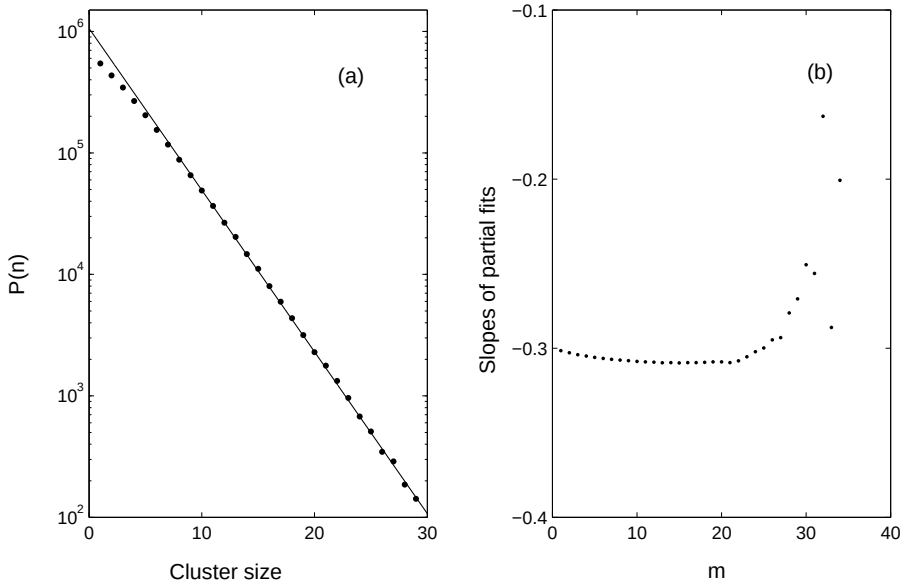


Figure 7.4: (a) P(n) along with the fit between cluster size 7 and 35 as described in the text. (b) The slopes of partial fits between cluster size m and cluster size 35.

7.2.2 Dependence of ξ on Disorder Strength

The width of the coupling distribution, μ is the only control parameter that characterizes the chain. Also, as we showed in the previous section, distributions of cluster sizes are asymptotically exponential, and hence they are characterized by a single parameter ξ. Hence, the function $\xi(\mu)$ is a comprehensive quantity that characterizes the physics of the chain in the unsynchronized regime. We plot this function in Fig. 7.5. We find good agreement between simulation (solid dots) and RG (open circles) until $\mu \approx 2.5$, after which the two results begin to deviate from each other. The small,

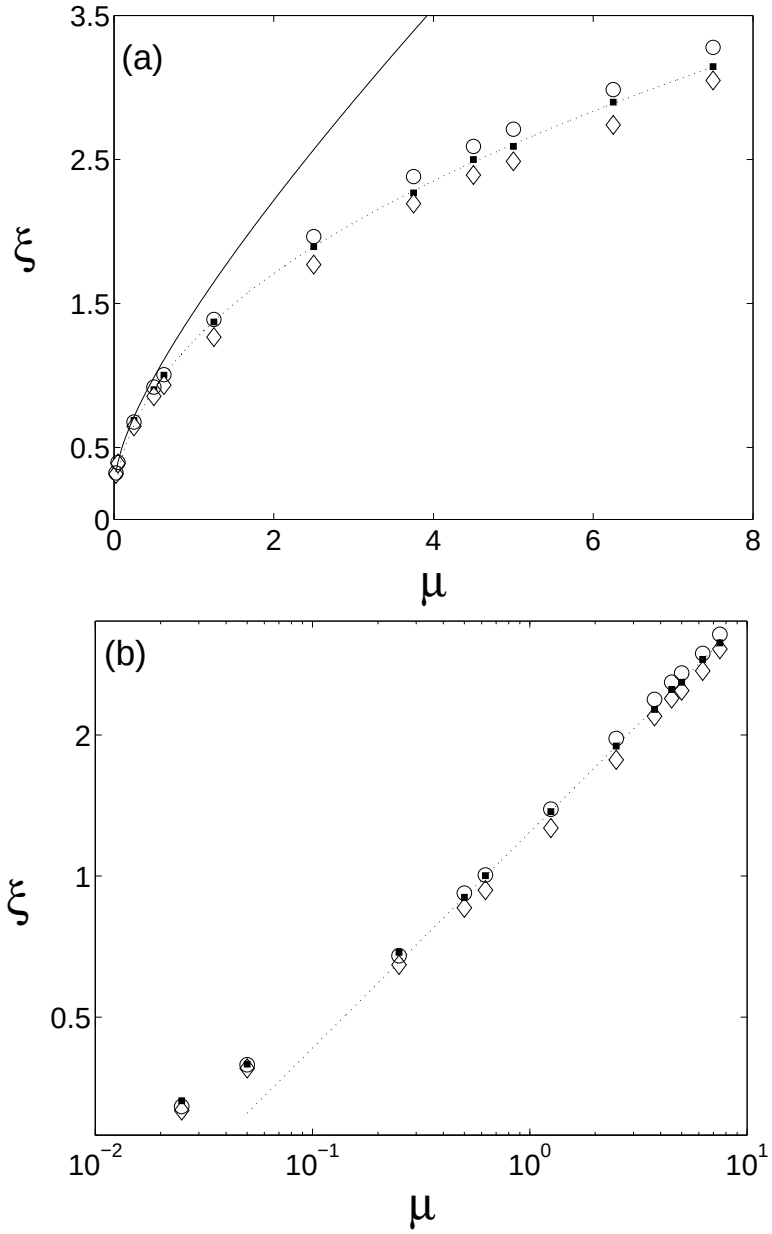


Figure 7.5: Characteristic cluster length ξ given by the inverse of the fits to data in Fig. 7.2. (a): Results of simulation (filled squares), the numerical RG of Section 6.1 (open circles) and an enhanced numerical RG with strain check (diamonds) described in Section 7.2.2. The solid line is a simple prediction $\xi(\mu) = -1/\log(1-\wp(\mu))$ where $\wp(\mu)$ comes from Eq. (7.8). The dashed line is a linear fit to the simulation data (solid squares) for $\mu > 0.625$ and has a slope of 0.461; similar fits to the circles give the slope 0.477 and to the diamonds the slope of 0.478. (b): Log-log plot of the same data excluding the analytic prediction.

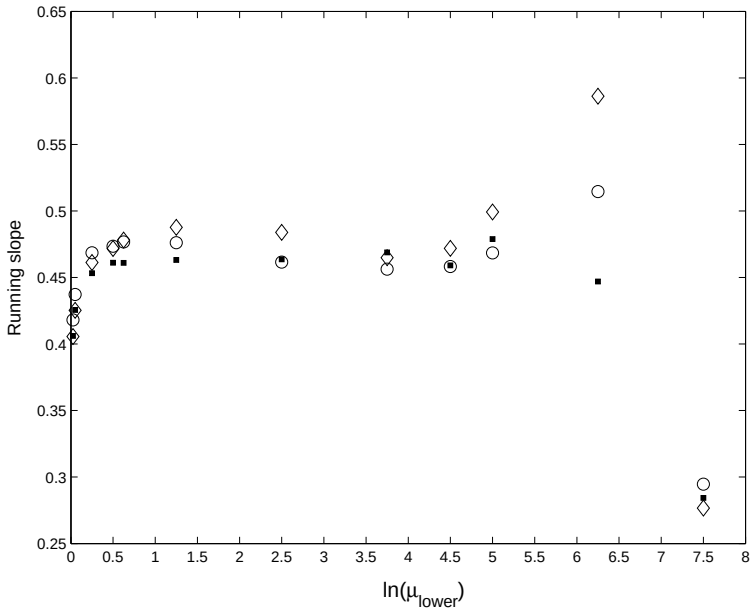


Figure 7.6: Slopes of partial fits of ξ versus $\ln \mu$ between $\ln \mu_{lower}$ and $\ln 7.5$. Solid dots: simulation; open circles: RG without strain check; diamonds: RG with the strain check.

but growing discrepancy beyond that μ suggests two possibilities: (i) existence of an additional mechanism in the physics of cluster formation that becomes important at larger μ and is not taken into account by the decimation steps of the RG procedure; (ii) corrections are needed to the steps already included in the RG procedure. The $\xi(\mu)$ obtained with the version of RG which included the strain check is shown in diamonds in Fig. 7.5. All three sets of data, when plotted on a log-log scale follow approximately straight lines for $\mu > 0.5$, so $\xi(\mu)$ appears to become a power law in this range. The straight-line fits to the data were made for μ between 7.5 and 0.625. The error bars in the values of these exponents can be estimated from the scatter in the partial fits in Fig. 7.6 (we chose the range $0.625 < \mu < 7.5$ for doing this). This leads to the following estimate of ξ: 0.47 ± 0.01 for simulations, 0.48 ± 0.02 for RG with and without the strain check (when rounded to two digits).

7.2.3 Approximate Analytical Argument

A simple estimate of characteristic cluster sizes that reproduces the exponential form of $P(n)$ and also puts an upper bound on the characteristic cluster size can be made by noting the role of weak couplings in limiting cluster sizes.[2] Indeed, that small clusters are likely to form in between weak couplings (also between the crazy oscillators, but these are padded by a weak coupling on each side, so are automatically taken into account), and since the distribution of separations between such weak couplings follow Poisson statistics the exponential distribution should be expected. Using this hypothesis the probability of finding a cluster of size n (i.e., a cluster delimited by two weak couplings n units apart) is $P(n, \wp) = C(1 - \wp)^n = Ce^{-n/\xi}$ where $\xi = -1/\ln(1 - \wp)$ and $\wp$ is a probability that any randomly chosen coupling is weak. We make an estimate of this $\wp$ based on bare distributions of ω and K. Similarly to the RG, we define a coupling to be weak if $2K < |\omega_l - \omega_r|$. For each such bond, the probability that the required ratio is > 1 is given by

$$\wp = 2\int_0^\infty \left(\int_{-\infty}^\infty g(\omega)g(\omega+\delta)\,d\omega\right) \times \left(\int_0^{\delta/2} G(K)\,dK\right)\,d\delta. \tag{7.6}$$

The integral of $g(\omega)g(\omega + \delta)$ is the probability that for each of the two oscillators connected by this bond, $\omega_l - \omega_r = \delta$. For each such frequency difference, all the bonds with $K < |\delta|/2$ are considered weak. Finally, all possible values of δ are integrated in the outer integral. After performing the inner integrals over ω and K we obtain

$$\wp(\mu) = \frac{4}{\pi^3}\int_0^\infty \frac{2\pi}{4+\delta^2}\arctan\left(\frac{\delta}{2\mu}\right)\,d\delta. \tag{7.7}$$

To handle this integral we differentiate $\wp$ with respect to μ under the integral sign, eliminating *arctan* and permitting the resulting expression to be integrated in δ in

[2]The exponential dependence of the correlation function, which is closely related to distribution of cluster sizes, is actually rather familiar from equilibrium statistical mechanics. Examples include the Ising model and XY model that display exponentially decaying correlations [102]. This can be shown using standard techniques, such as transfer matrices, or by performing a high-temperature expansion [50]. However, underpinning these explanations of the exponentially decaying correlations is a partition function, which is inapplicable in our non-Hamiltonian model, so we are forced to seek a different explanation

closed form. Finally, by reintegrating with respect to μ and redefining variables once more we find

$$\wp(\mu) = \int_{2\mu}^{\infty} \frac{4}{\pi^2} \frac{\ln(x^2/4)}{(x^2-4)}\, dx. \tag{7.8}$$

This result is compared with both the RG predictions and numerical solutions in Fig. 7.5. The expression gives a good description of the data for $\mu \to 0$, but overestimates the characteristic cluster size for larger values of μ. Of course such simple estimate for ξ is but an upper bound; it does not take into account any information about the constituent ω or K. The renormalization group method, which does not treat all scales of frequencies and couplings at once, clearly provides better predictions then an estimation based on bare distributions. The noteworthy feature of the RG procedure is that the process of building up clusters by successive applications of very simple decimation steps leads to good predictions of structures which are far from trivial!

7.2.4 Cluster Frequency Distributions

Investigating the frequency distribution of clusters of a given size provides further insight into the physics and the validity of this RG procedure. Distributions in Fig. 7.7, serve as representative examples.[3] The RG predictions are compared with the results from simulations of Eqs. (6.1) for several clusters sizes and parameters μ. Notice the excellent agreement for the larger values of m and also for all m at small values of μ. The comparisons in Fig. 7.7 do show some underestimation of the number of smallest sized clusters ($m = 1$ and $m = 2$) near zero frequency. The region of the greatest discrepancy, small m and ω, is in fact where the crazy oscillator decimation is expected to be least accurate, since by the time the RG procedure reaches small frequencies, the distributions of remaining frequencies and couplings are no longer expected to be wide, so the approximations based on strong randomness are not as accurate. However the disagreement decreases rapidly with increasing m, because as

[3] A reader is invited to see [101] for a different, but complimentary way of plotting these distributions.

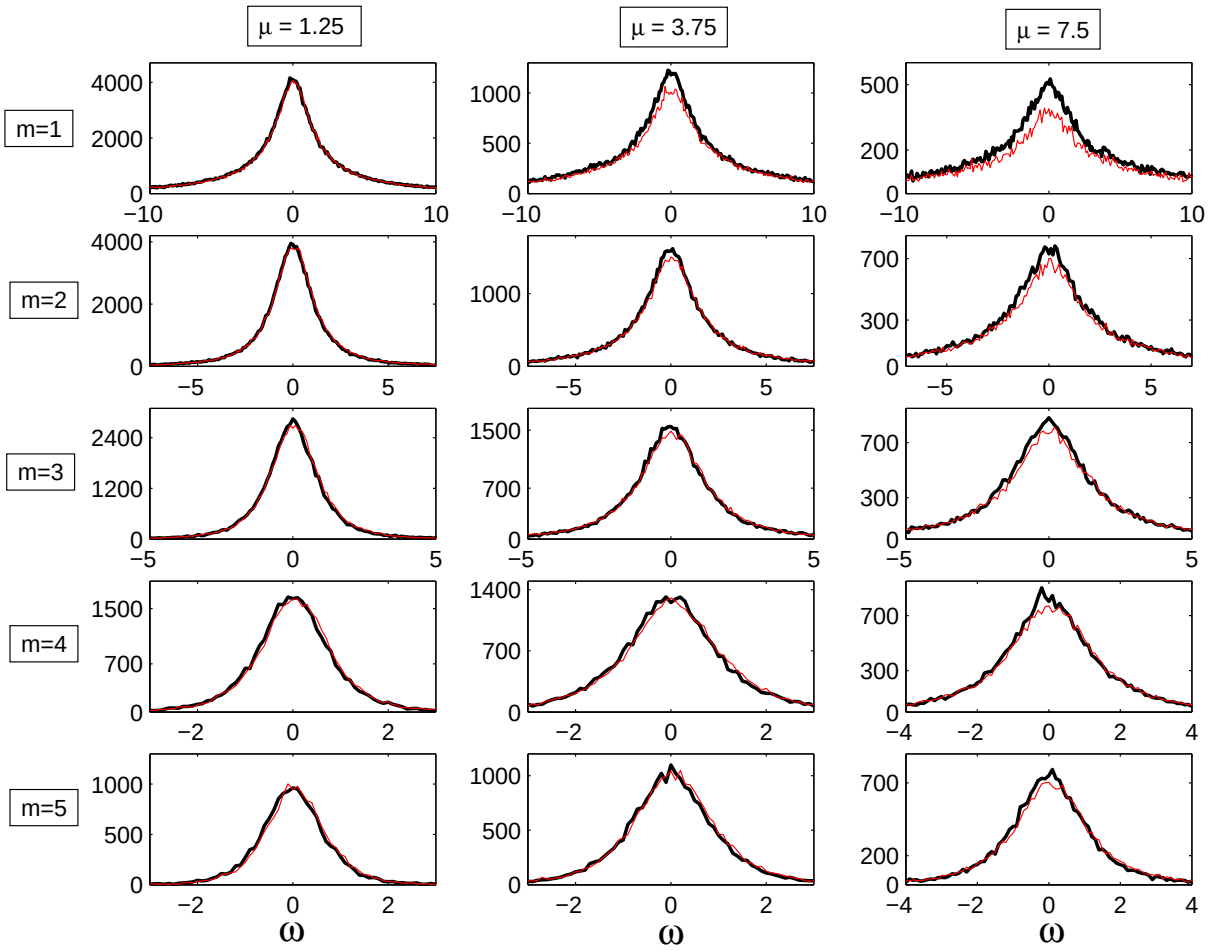


Figure 7.7: Frequency distributions. The y-axis displays the number of clusters of a given size m in a frequency bin of width $\Delta\omega = 0.1$ at frequency specified on the x-axis.We display distributions for $m = 1, 2, 3, 4$ at $\mu = 1.25, 3.75$ and 7.5. Thick curve: numerical simulation; thin red curve: numerical RG *without* the use of the strain check.

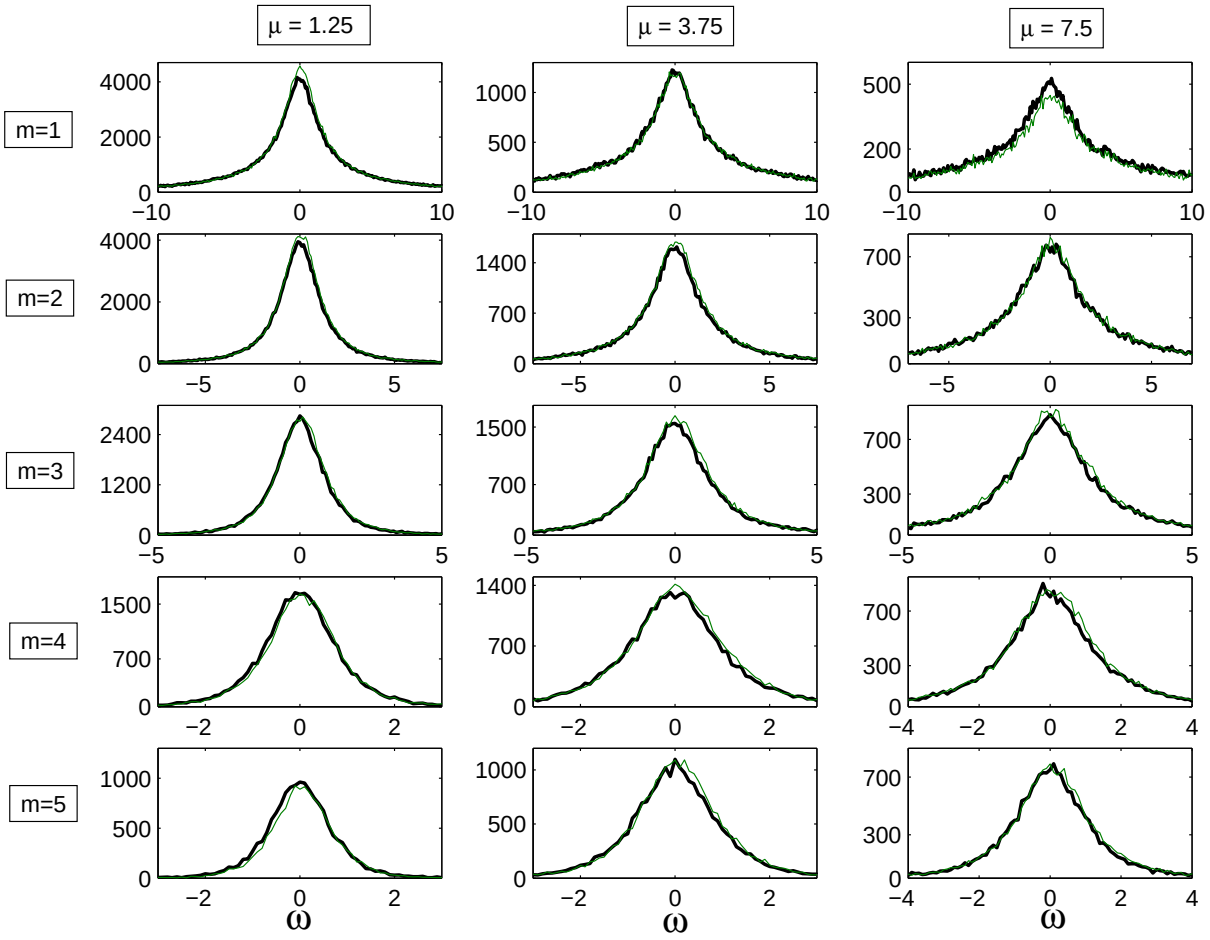


Figure 7.8: Frequency distributions. The y-axis displays the number of clusters of a given size m in a frequency bin of width $\Delta\omega = 0.1$ at frequency specified on the x-axis.We display distributions for $m = 1, 2, 3, 4$ at $\mu = 1.25, 3.75$ and 7.5. Thick curve: numerical simulation; thin green curve: numerical RG *with* the use of the strain check.

described above, the frequencies of larger clusters are insensitive to the corrections from neighbors: for large m the running frequency is mostly given by the average of intrinsic values of ω. We also note in passing that the disagreement decreases with decreasing μ, but refrain from proposing an explanation for this.

Another remark is regarding the shape of distributions for $m = 1$ oscillators: they do not fit pure Lorentizian at all values of ω, however at very large ω the tails of RG and simulation results match the tails of the bare frequency distribution. This is expected because these tails are comprised of very high frequency oscillators which are typically decimated out as crazy oscillators to form one-oscillator "clusters." Because of their high bare ω, the frequency correction from the neighbors is typically negligible, so their renormalized frequencies equal bare frequencies. An interesting observation however (not shown in Fig. 7.7), is that the frequency interval over which there is a good match between RG and simulation for these $m = 1$ oscillators extends to lower frequencies then where the RG and simulation data match with the bare Lorentzian, providing yet another indication for the validity of the frequency adjustments included in the fast oscillator decimation steps.

The use of strain check in the RG appears to increase the number of low-frequency clusters for m not too large and for all sampled μ, as can be seen from Fig. 7.8. Although not obvious from these figures, closer inspection reveals that the tails are not significantly affected - the increase is mostly at low frequencies. This net increase in the number of small-m clusters is compensated at larger m - this is not shown in Fig. 7.8 but can be seen from closer inspection of large-m distributions and can also be seen from statistics of cluster sizes, Fig. 7.2 (notice the diamonds slightly above the circles for small m and below the circles for larger m). Indeed, the strain check acts to mostly break up large clusters, therefore decreasing their number and increasing the number of smaller ones, but it is not trivial to see why this increase affects mostly low frequencies.

We leave this section with a reminder that a more refined measure of a cluster strain *or* other improvement in the RG strategy, which would reduce differences between all RG results and simulation results, is due to be invented in the future.

Chapter 8

Concluding Remarks

In this work we investigated the collective behavior of a one-dimensional chain of coupled nonlinear oscillators with random frequencies and nearest-neighbor couplings. For this purpose we developed a real space renormalization group approach, which is expected to be reliable in analyzing chains with large disorder. As we report above, our RG approach did an excellent job in capturing both the dynamics of individual oscillators, and the large scale behavior of the chain.

The disordered oscillator chain has the potential of establishing macroscopic order, which is exhibited by a finite fraction of all oscillators in the system moving in accord with the same frequency. Indeed in systems with higher connectivities, such global synchronization may occur, but in the limit of short range interactions, the macroscopic order is stymied by fluctuations. This dynamical behavior is reminiscent of critical phenomena in equilibrium statistical mechanics, where for each system there exists a lower critical dimension, and only dimensionality higher than that allows macroscopic ordering. In our analysis, we concentrated on the quantities that most directly capture the collective aspects of the chain's behavior, and therefore also the competition between the interactions which seek to establish a synchronized phase, and the disorder-induced fluctuations which destroy it. This physics is fully contained in the cluster-size distribution, $P(n) \sim e^{-n/\xi(\mu)}$, which we find as a function of the interaction-strength tuning parameter, μ.

The behavior of the average cluster size, $\xi(\mu)$, acts as an effective correlation length for the oscillator chain, where we once more think of our system in terms of

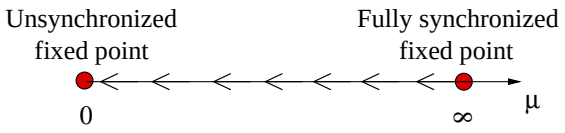


Figure 8.1: Fixed points and flow diagram of the oscillator chain. The short-range oscillator chain has only two fixed points: the unstable fully synchronized fixed point at infinite interactions (or zero disorder) $\mu \to \infty$, and the stable unsynchronized fixed point. The cross over flow between the two points is associated with a correlation length, captured by the average cluster size, $\xi(\mu)$.

the theory of critical phenomena. Since our oscillator chain is below its lower critical dimension, its phase diagram resembles that of the one-dimensional Ising model, with μ serving the role of inverse temperature $1/T$. The chain has an ordered phase, where all oscillators are synchronized, which is reached in the limit $\mu \to \infty$ (analogous to the Ferromagnetic phase at $T \to 0$). This phase, however, is described by an *unstable* fixed point. If an RG flow can be extended to the entire parameter space, the full-synchronization fixed point flows under RG to the stable fixed point found at the noninteracting point, $\mu = 0$, which describes the unsynchronized phase (analogous to the paramagnetic fixed point at $T \to \infty$). The putative flow diagram of the chain is shown in Fig. 8.1. The flow from the full-synchronization point to the unsynchronized $\mu = 0$ point is indeed associated with a length scale, whose meaning is the coarse-graining length at which the chain seems fully unsynchronized, i.e., consisting of single free oscillators. This length is the size of synchronized clusters in the original chain, $\xi(\mu)$. Since $\xi(\mu)$ serves as an effective correlation length, it may diverge in a universal fashion in the limit $\mu \to \infty$. While our RG method can only access moderate values of μ, we cautiously argue that we can access the critical region, as can be seen in Fig. 7.2. From our results we obtain the critical exponent ν of the synchronization transition:

$$\xi(\mu) \sim \mu^{\nu} \qquad \nu \simeq 0.48 \pm 0.02 \tag{8.1}$$

Note that this behavior is not analogous to the divergence of the Ising chain correlation length $\xi \sim \ln \frac{1}{T}$.

The method by which we obtained the above results constitutes the main achieve-

ment of this work. We developed a real space renormalization group scheme that successfully predicts the detailed behavior of a nonlinear oscillator chain with short-range interactions. We have implemented the RG scheme numerically on a chain of 10^6 oscillators and compared its results with exact simulations of the model equations, Eqs. (6.1), in chapter 7. The agreement is very good, and for some characteristics, such as the cluster frequency distributions at low frequencies, it improves with larger cluster sizes. This indicates particularly that our local RG scheme successfully captures the nontrivial collective features of the chain expressed in large synchronized clusters.

One practical implication of our analysis may be found when considering coupled laser arrays, where only limited connectivity can be achieved. Diode lasers, for example, are manufactured in one-dimensional bars that allow for evanescent coupling through overlapping electric fields. Due to the exponential falloff in the electric field envelope with distance, the coupling between well-behaved diodes can only be short ranged. The tools developed here could be used to study the extent of coherence that such systems exhibit. A potentially promising area of current research is constructing grids of diodes coupled in 2-dimensions by stacking 1-dimensional bars. Our RG technique could possibly also be extended to such a system, and help clarify both its properties, and whether it can establish macroscopic coherence.

Our work suggests several other directions for future research. First, we may ask how can one improve the applicability of our RG scheme to larger values of the interaction parameter μ. This may be pursued by focusing on the concept of intra-cluster strain, which is roughly the variance of phase differences between neighboring oscillators within a cluster. As interaction increases, cluster sizes grow, and bigger strain is needed to accommodate the spread of frequencies within a cluster (cf. Sec. 7.2.2). In our current scheme we made the first attempt to take this into account with the strain check based on the argument of [51], a condition expressed by Eq. (6.54). This led to some improvement - for example, $\xi(\mu)$ was modified in the right direction, but this modification was overestimated. The strain check may be augmented in the future.

A second natural question is the degree of universality in the oscillator chain. Our result for the correlation length critical exponent ν, Eq. (8.1), was obtained using Lorentzian distributions. It is possible that the exponent depends on the nature of the tails of the distributions used; Lorentzians, for instance, do not have a variance. We intend to investigate the universality of ν alongside investigating the applicability of our RG scheme to distributions with narrower tails. This work is already underway [103].

A more general question concerns the development of an RG scheme and the analysis of synchronization in short-range networks at higher dimensions, and especially at two dimensions, which is very close to the lower critical dimension: Refs. [104, 105] suggest that in two dimensions there may be a transition to a macroscopically synchronized phase, or that $d = 2$ is the lower critical dimension giving rise to interesting anomalous behavior (see also [45]). The development of a decimation procedure at dimensions higher than one is quite challenging because of the higher connectivity of the system, and the possibility that real-space local decimation steps change the geometry of the system.

Appendix A

Relating Beam Theory to Duffing Equation

In this appendix, the parameters of the Duffing equation (except damping) are extracted from the underlying equation of motion of a doublyclamped beam. This is useful to obtain an estimate of the realistic values of the parameters that go into the Duffing equation and also to see explicit dependence of these parameters in the two regimes: the tension-dominated regime and the stiffness-dominated regime. In the process of preparation of this appendix, I learned that the review article [5] contains a similar discussion. I chose to include this appendix in any case, because it was part of the work I have done while collaborating with experimentalists. Also, figures are included here which do not appear in [5]. A different method (Galerkin approximation) has recently been used for similar calculations in [4].

The beam will be described by a 1-dimensional displacement function $Y(x,t)$. We start with the following equation describing the dynamics of the beam:

$$YI\frac{\partial^4 y}{\partial x^4} - \tau_0\left[1 + \frac{YA}{2l}\int_0^l \left(\frac{\partial y}{\partial x}\right)^2 dx\right]\frac{\partial^2 y}{\partial x^2} = -\rho A\frac{\partial^2 y}{\partial t^2} \tag{A.1}$$

The integral term corresponds to increasing tension when the beam is deformed, i.e.,

the tension depending on the shape of the beam. Let us first nondimentionalize:

$$y = ry' \tag{A.2}$$

$$x = l_0 x' \tag{A.3}$$

$$t = t'\sqrt{\frac{\rho A l_0^2}{\tau_0}} \tag{A.4}$$

The resulting dimensionless equation (dropping $'$) is:

$$\xi\frac{\partial^4 y}{\partial x^4} - \left[1 + \nu\int_0^{l/l_0}\left(\frac{\partial y}{\partial x}\right)^2 dx\right]\frac{\partial^2 y}{\partial x^2} = -\frac{\partial^2 y}{\partial t^2}, \tag{A.5}$$

where

$$\xi = \frac{YI}{l_0^2\tau_0}, \tag{A.6}$$

$$\nu = \frac{YA}{2\tau_0}\frac{r^2}{ll_0}. \tag{A.7}$$

For a double-clamped beam, l remains strictly equal to l_0 (and to a good accuracy for a cantilever as well, although technically it can stretch). Also, for a doublyclamped beam the boundary conditions are

$$y(0,t) = y(1,t) = \left.\frac{\partial y}{\partial x}\right|_{(0,t)} = \left.\frac{\partial y}{\partial x}\right|_{(1,t)} = 0. \tag{A.8}$$

For $\nu = 0$ the equation is linear with the general solution

$$y_0(x,t) = \sum_{n=1}^{\infty} Y_n(x)e^{-i\omega_n t}, \tag{A.9}$$

where Y_n are eigenfunctions of the operator $\hat{L}$:

$$\hat{L}Y_n(x) = \omega_n^2 Y_n(x), \tag{A.10}$$

$$\hat{L} = \xi\frac{d^4}{dx^4} - \frac{d^2}{dx^2}, \tag{A.11}$$

with boundary conditions

$$Y(0) = Y(1) = Y'(0) = Y'(1) = 0. \tag{A.12}$$

It is straightforward to show that with these boundary conditions, $\hat{L}$ is Hermitian, i.e.,

$$\int_0^1 Y_n(x)\hat{L}Y_m(x)_{,}^{,}x = \int_0^1 Y_m(x)\hat{L}Y_n(x)_{,}^{,}x. \tag{A.13}$$

Therefore, its eigenmodes form a complete set, so a nonlinear solution can be expanded in terms of them

$$y(x,t) = \sum_{n=1}^{\infty} a(t)Y_n(x). \tag{A.14}$$

Analogously to the time-dependent perturbation theory in Quantum Mechanics the nonlinearity will spread solution among the linear modes, even if initially the initial condition is fully in one single mode. Nevertheless, we will make an approximation based on the physical grounds that when the amplitude is sufficiently small, the system is approximately in the ground state, i.e., $y(x,t) \approx a(t)Y_{0,1}(x)$, and we would like to know the effect of nonlinearity on the timedependence of y, i.e., on a_1. In this spirit, we substitute $a(t)Y_{0,1}(x)$ into Eq. (A.5), multiply by $Y_{0,1}(x)$ and integrate out the x-dependence. The resulting equation for $a(t)$ is

$$\ddot{a} = -\omega^2 a(t) - \gamma a^3(t), \tag{A.15}$$

where

$$\omega^2 = \frac{\int_0^1 \left(\xi \frac{\partial^4 Y_1}{\partial x^4} - \frac{\partial^2 Y_1}{\partial x^2}\right) Y_1(x)\, dx}{\int_0^1 Y_1^2\, dx}, \tag{A.16}$$

and

$$\gamma = -\mu \frac{\left(\int_0^1 \frac{\partial^2 Y_1}{\partial x^2} Y_1(x)\, dx\right)\left(\int_0^1 \left(\frac{\partial Y_1}{\partial x}\right)^2 dx\right)}{\int_0^1 Y_1^2\, dx}. \tag{A.17}$$

Actually, Eq. (A.16) is identical to Eq. (A.10) with $n = 1$, so in the approximation used, the coefficient ω does not depend on the presence of nonlinearity. Eq. (A.16) can also be seen as a variational solution to ω. However, we do not need to employ

the variational method (which by its nature gives an approximate values) because it is best to solve the eigenproblem defined by Eqs. (A.10)-(A.12) directly. As an ODE, Eq. (A.10) has the following general solution

$$Y_n(x) = a_n e^{\lambda_n x} + b_n e^{-\lambda_n x} + c_n \cos(\lambda'_n x) + d_n \sin(\lambda'_n x), \tag{A.18}$$

where

$$\lambda_n = \sqrt{\frac{1 + (1 + 4\xi\omega_n^2)^{1/2}}{2\xi}}, \tag{A.19}$$

$$\lambda'_n = \sqrt{\frac{-1 + (1 + 4\xi\omega_n^2)^{1/2}}{2\xi}}. \tag{A.20}$$

However, Eq. (A.10) is a boundary-value problem which needs to satisfy the boundary conditions given by Eq. (A.12). These conditions will determine the eignevalues $\{\lambda_n, \lambda'_n\}$, as well as the coefficients $\{a_n, b_n, c_n, d_n\}$. The doublyclamped boundary conditions lead to the following equation (dropping the subscripts n for a moment):

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ e^{\lambda} & e^{-\lambda} & \cos\lambda' & cos\lambda' \\ \lambda & -\lambda & 0 & \lambda' \\ \lambda e^{\lambda} & -\lambda e^{-\lambda} & -\lambda' \sin\lambda_2 & \lambda_2 \cos\lambda_2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{A.21}$$

For a nontrivial solution we demand that the matrix on the left hand side is noninvertible by setting its determinant to 0. This gives an implicit equation for the set $\{\omega_n(\xi)\}$:

$$f(\omega, \xi) = 4\sqrt{\frac{\omega_n^2}{\xi}}\left[1 - \cos\lambda'_n \cosh\lambda_n\right] + \frac{2}{\xi}\sin\lambda'_n \sinh\lambda_n = 0, \tag{A.22}$$

where λ'_n and λ_n are given by Eqs. (A.19)-(A.20). For a given value of ξ, there are an infinite number of zeros of the function f at a discrete set of ω_n. Using this, we can substitute the resulting $Y_1(x)$ into Eq. (A.17) to obtain γ. See Fig. A.1.

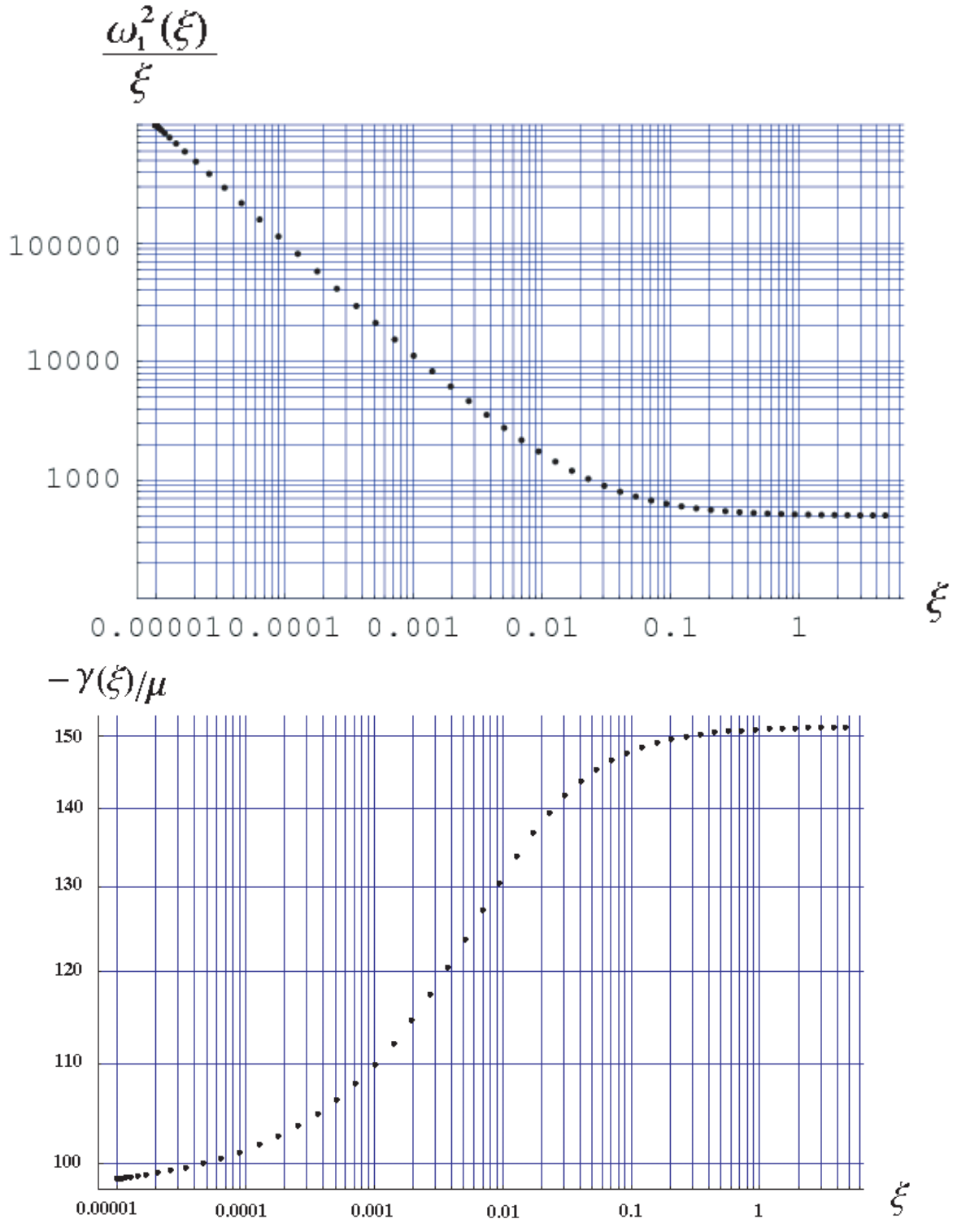


Figure A.1: Top: $\omega_1^2(\xi)/\xi$. Notice the crossover from the the string-dominated regime at lower ξ to the stiffness-dominated regime at larger ξ. Bottom, $-\gamma/\mu$.

Appendix B

Properties of Noise in the Amplitude Equations

Referring back to definitions in 2.1.1, we note that

$$\mathcal{N}_Q(T) = \frac{3^{1/2}}{2}\left(\frac{\eta}{\epsilon}\right)^{3/2}\frac{\Omega}{\pi}\int_{\tau}^{\tau+2\pi/\Omega}\tilde{f}(\xi)\sin(\Omega\xi)\,d\xi \tag{B.1}$$

$$= \frac{3^{1/2}}{2}\left(\frac{\eta}{\epsilon}\right)^{3/2}\frac{\Omega}{\pi}\int_{2\eta T/\epsilon}^{2\eta T/\epsilon+2\pi/\Omega}\tilde{f}(\xi)\sin(\Omega\xi)\,d\xi. \tag{B.2}$$

Similarly

$$\mathcal{N}_P(T) = \frac{3^{1/2}}{2}\left(\frac{\eta}{\epsilon}\right)^{3/2}\frac{\Omega}{\pi}\int_{2\eta T/\epsilon}^{2\eta T/\epsilon+2\pi/\Omega}\tilde{f}(\xi)\cos(\Omega\xi)\,d\xi. \tag{B.3}$$

We compute correlation functions,

$$\langle\mathcal{N}_Q(T_1)\mathcal{N}_Q(T_2)\rangle = \frac{3}{4}\left(\frac{\eta}{\epsilon}\right)^3\left(\frac{\Omega}{\pi}\right)^2\int_{2\eta T_1/\epsilon}^{2\eta T_1/\epsilon+2\pi/\Omega}d\xi'\int_{2\eta T_2/\epsilon}^{2\eta T_2/\epsilon+2\pi/\Omega}d\xi''\langle\tilde{f}(\xi')\tilde{f}(\xi'')\rangle\sin(\Omega\xi')\sin(\Omega\xi''). \tag{B.4}$$

Using the correlation property of $\tilde{f}$, Eq. (2.10) we have

$$\langle \mathcal{N}_Q(T_1)\mathcal{N}_Q(T_2)\rangle = \\ \frac{3}{2}B'\left(\frac{\eta}{\epsilon}\right)^3\left(\frac{\Omega}{\pi}\right)^2\int_{2\eta T_1/\epsilon}^{2\eta T_1/\epsilon+2\pi/\Omega} d\xi'\int_{2\eta T_2/\epsilon}^{2\eta T_2/\epsilon+2\pi/\Omega} d\xi''\delta(\xi'-\xi'')\sin(\Omega\xi')\sin(\Omega\xi''). \tag{B.5}$$

We make a simple change of variables $y = \Omega\xi$ and also the property of Delta functions $\delta(ay) = \frac{1}{a}\delta(y)$. Then

$$\langle \mathcal{N}_Q(T_1)\mathcal{N}_Q(T_2)\rangle = \frac{3}{2}\frac{B'}{\Omega}\left(\frac{\eta}{\epsilon}\right)^3\left(\frac{\Omega}{\pi}\right)^2\int_{L_1}^{L_1+2\pi} dy'\int_{L_2}^{L_2+2\pi} dy''\delta(y'-y'')\sin(y')\sin(y''), \tag{B.6}$$

where $L_i = 2\eta T_i\Omega/\epsilon$. Similarly,

$$\langle \mathcal{N}_P(T_1)\mathcal{N}_P(T_2)\rangle = \frac{3}{2}\frac{B'}{\Omega}\left(\frac{\eta}{\epsilon}\right)^3\left(\frac{\Omega}{\pi}\right)^2\int_{L_1}^{L_1+2\pi} dy'\int_{L_2}^{L_2+2\pi} dy''\delta(y'-y''),\cos(y')\cos(y'') \tag{B.7}$$

$$\langle \mathcal{N}_P(T_1)\mathcal{N}_Q(T_2)\rangle = \frac{3}{2}\frac{B'}{\Omega}\left(\frac{\eta}{\epsilon}\right)^3\left(\frac{\Omega}{\pi}\right)^2\int_{L_1}^{L_1+2\pi} dy'\int_{L_2}^{L_2+2\pi} dy''\delta(y'-y'').\cos(y')\sin(y'') \tag{B.8}$$

Let us investigate the properties of

$$I_s = \int_{L_1}^{L_1+2\pi} dy'\int_{L_2}^{L_2+2\pi} dy''\delta(y'-y'')\sin(y')\sin(y''). \tag{B.9}$$

The value of this integral is zero unless the square of integration includes the $y' = y''$ diagonal line, in which case it is equal to

$$I_s = \pi - \frac{1}{2}\left[|\Delta| + \frac{1}{2}\sin(2L_1) - \frac{1}{2}\sin(2(L_1+|\Delta|))\right], \tag{B.10}$$

where $\Delta = L_2 - L_1$. In other words,

$$\langle \mathcal{N}_Q(T_1)\mathcal{N}_Q(T_2)\rangle = \frac{3}{2}\frac{B'}{\Omega}\left(\frac{\eta}{\epsilon}\right)^3\left(\frac{\Omega}{\pi}\right)^2 \times \begin{cases} \pi - \frac{1}{2}\left[\frac{2\eta\Omega}{\epsilon}\Delta T + \frac{1}{2}\sin\left(\frac{4\eta\Omega}{\epsilon}T_1\right) - \frac{1}{2}\sin\left(4\left(\frac{\eta\Omega}{\epsilon}T_1 + \frac{\eta\Omega}{\epsilon}\Delta T\right)\right)\right] & |\Delta| \le \frac{\pi\epsilon}{\eta\Omega}, \\ 0 & |\Delta| > \frac{\pi\epsilon}{\eta\Omega}, \end{cases} \tag{B.11}$$

where $\Delta T = T_2 - T_1$. This finding is very peculiar. First, the correlation function is not a delta function in ΔT, so the noise terms $\mathcal{N}$ are not strictly white, and second it is dependent not only on the time difference, but also on the time itself! The integral of this autocorrelation function over $-\frac{\pi\epsilon}{\eta\Omega} < \Delta T < \frac{\pi\epsilon}{\eta\Omega}$ is

$$\left[\frac{3}{2}\frac{B'}{\Omega}\left(\frac{\eta}{\epsilon}\right)^3\left(\frac{\Omega}{\pi}\right)^2\right] \times \left[\frac{\pi^2\epsilon}{\eta\Omega} - \frac{\pi\epsilon}{2\eta\Omega}\sin\left(\frac{4\eta\Omega T_1}{\epsilon}\right)\right]. \tag{B.12}$$

However, the support of the autocorrelation function of $\mathcal{N}_Q$ is very narrow (compared to the height) and the fast time dependence of the area averages out to zero on the long timescale, so we will make an approximation

$$\begin{aligned} \langle \mathcal{N}_Q(T_1)\mathcal{N}_Q(T_2)\rangle &\approx \left[\frac{3}{2}\frac{B'}{\Omega}\left(\frac{\eta}{\epsilon}\right)^3\left(\frac{\Omega}{\pi}\right)^2\right] \times \frac{\pi^2\epsilon}{\eta\Omega}\delta(\Delta T) \\ &= \frac{3}{2}\frac{\gamma}{\omega_0^5}\left(\frac{\eta}{\epsilon}\right)^2 B\delta(\Delta T) \\ &\approx \left(\frac{3\gamma\eta^2}{8\omega_0^3\Gamma^2}B\right)\delta(\Delta T), \end{aligned} \tag{B.13}$$

with the identical result for $\langle \mathcal{N}_P(T_1)\mathcal{N}_P(T_2)\rangle$. As for the cross-correlation functions, it is odd in ΔT and so its integral over $-\frac{\pi\epsilon}{\eta\Omega} < \Delta T < \frac{\pi\epsilon}{\eta\Omega}$ zero. Thus we consider $\langle \mathcal{N}_Q(T_1)\mathcal{N}_P(T_2)\rangle = \langle \mathcal{N}_P(T_1)\mathcal{N}_Q(T_2)\rangle = 0$.

Appendix C

Features in the Pattern of Classical Trajectories

The goal here is simply to provide examples of some of the interesting features observed in the bistable regime which were not mentioned in the main body of the thesis.

C.1 Shadow regions

Shadows can be seen in the pattern of escape trajectories terminated at the separatrix. Because there is nothing fundamentally special about the separatrix, there is nothing special about shadows; they can of course be reached from the attracting fixed point by following a trajectory which leaves the basin and then re-enters it and eventually wanders into the shadow region. They are important however, when one is concerned about the distribution of exit locations along the separatrix: probability of exiting the basin through a given point on the separatrix, which has been discussed in Maier and Stein's long SIAM paper [74].

Here, a specific example will be considered. Using the notation of my thesis, we will choose $\eta = 0.18$ and $\mathcal{F} \approx 0.164$ which corresponds to the 10% of the hysteresis and will look at transitions from the lower amplitude branch (the down $\rightarrow$ up transitions). Consider first a set of 50 trajectories that escape from the attracting fixed point depicted in Fig. C.1. All these trajectories satisfy the Hamilton's equation with energy 0, but they differ in trajectory parameters (see Section 3.1 of the Thesis). The range

of trajectory parameters in Fig. C.1 is $[0, 2\pi]$. There is something interesting going on near the saddle. In Fig. C.2 we consider a narrower range of trajectory parameters, and also zoom in closer to the saddle. A pair of caustics: envelopes of intersections of nearby trajectories and the cusp at which they meet is clearly seen. Trajectory pattern for yet a smaller range of trajectory parameters is shown in Fig. C.3. Incidentally, an obvious but dramatic fact that trajectory parameters that differ very little can lead to drastically different end points is illustrated in Fig. C.4; I believe that the actions will also differ significantly. As stated above, the shadow region can be reached by first exiting the basin, but the action of a trajectory that reaches the separatrix on the shadow side, should generally be larger then the action of a trajectory on the other side, which reaches the separatrix directly, without having to leave the basin first. Let us make another zoom in the range of trajectory parameters: this is shown in Fig. C.5. We see that the projection unto the (Q, P) space of the most repelling eigenvector in fact forms the cut, or the boundary of the shadow region. It must be noted that if the plot was made over a larger range, the fan of trajectories actually intersects the projection of the eigenvector. This may need a more thorough investigation: to look at narrower and narrower ranges of trajectory parameters and observing if they intersect the projection of the eigenvector closer and closer to the saddle. In any case, it appears that the eigenvector defines the boundary of the shadow region only infinitesimally close to the saddle.

There is another technicality. One of the checks that is to repeat calculations at ever decreasing step sizes until the results do not change significantly (the other check of course is the comparison with available analytics). We made the calculations shown in Fig. C.5 at 1/10 the step size and aside from a slight shift of the trajectory parameter of the most likely escape trajectory, nothing much changes qualitatively.

C.2 Caustics emanating from saddles

There are situations when something interesting appears as we continue to zoom in to narrower and narrower range of trajectory parameters around the one that hits

the saddle. This is illustrated in Fig. C.6, where the parameters were $\eta = 0.09$ and $\mathcal{F} \approx 0.343$ is here). A new type of caustic is apparent and does not seem to be a direct continuation of the cusp pair formed by the trajectories in the larger parameter range. For some other values of parameters, it was observed that as the parameter range is varied from the relatively wide one to a narrower one, the position at which the trajectories appear to intersect moves in closer to the saddle. Evidently there is a caustic emanating from the saddle. In fact, caustics emanating from saddles have been predicted in [31]. That paper argues that whenever the ratio of the larger repelling eigenvalue to the smaller repelling eigenvalue $\lambda = \lambda_{lg}/\lambda_{sm}$ is less then 3/2, there is a caustic that emanates from the saddle point tangentially to the cut. In the above example, $\lambda = 1.37$, our observation is consistent with this prediction. So far we have not seen violations of this prediction. For example, when $\eta = 0.13$ and $\mathcal{F} \approx 0.19$ the behavior just described does not happen, and indeed, $\lambda \approx 1.92$ there. However, a more systematic numerical study is needed to verify this claim of [31], and we intend to do this in the near future.

C.3 Rotation of eigenvectors

Shadows appear to exist also in the case of up $\rightarrow$ down transitions. In some cases this phenomenon has been responsible for the seeming change in the slope of the effective barrier height versus the driving strength. Upon the examination of trajectories it was noticed that before the change in slope (say for $\mathcal{F}_1$) the most likely exit trajectory approaches the saddle from one side and for after the change in slope (say for $\mathcal{F}_2 > \mathcal{F}_1$) it approaches the saddle from the other side. This lead me to suppose that there is something akin to the first order transition, where there are actually several (locally minimizing) paths that lead to the saddle, and the values of their actions switch. Another possibility was simply a rotation of the eigenvector that defines the cut. This turned out to be the case, and moreover, the calculations show that in general this rotation is smooth, thus eliminating other exotic possibilities.

Fig. C.7 shows the snapshots of the fan of trajectories for three different values of

$\mathcal{F}$ for up $\rightarrow$ down transitions. All of these cases are actually very far from the up $\rightarrow$ down bifurcation point. The probable reason that the slope of the effective barrier height versus the driving strength appears to change is that it becomes harder to hit the saddle, as is evident from the middle plot in Fig. C.7. Indeed, upon a very painstaking bracketing, the slope change disappears, and nothing special is seen in the function of the effective barrier height versus the driving strength.

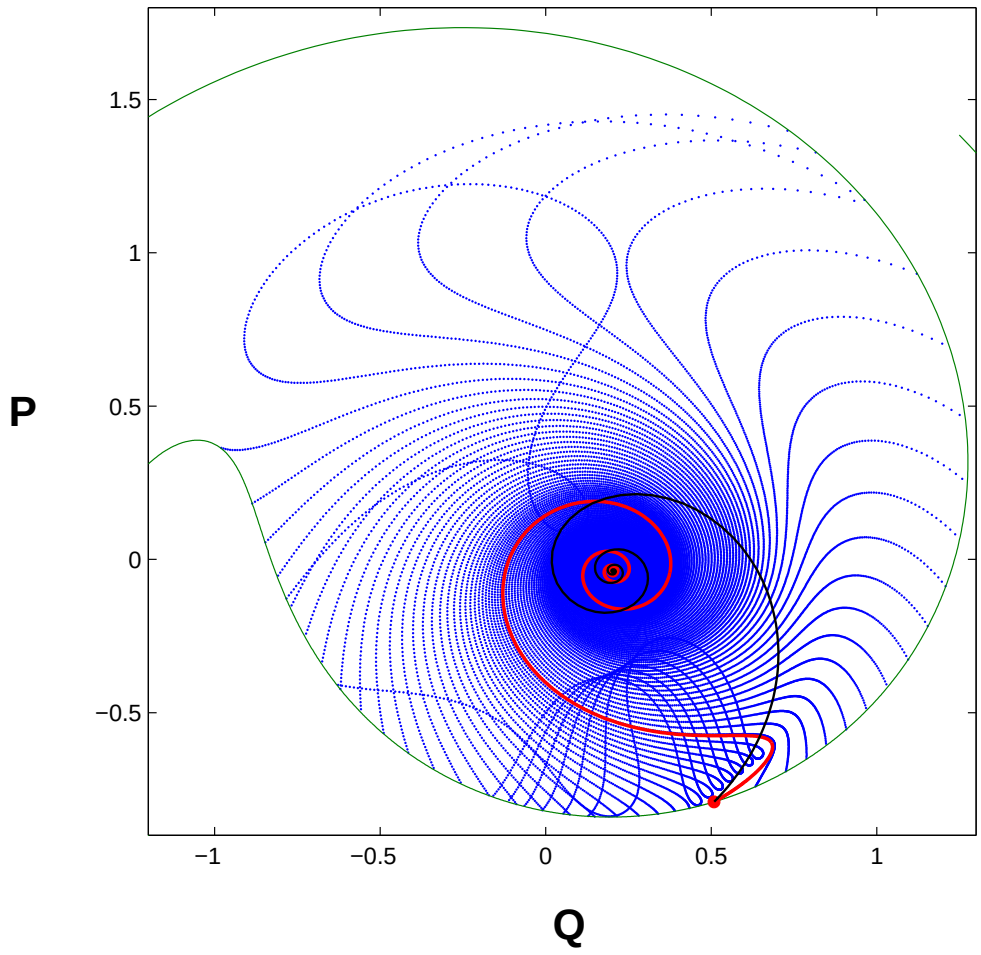


Figure C.1: A set of 50 trajectories that escape the attracting fixed point are shown in blue, the most likely escape trajectory (which we know must pass through the saddle) is in red and the relaxational trajectory from the saddle to the attracting fixed point appears in black. Note, that these two are not just related by reversing the sign of the friction, as would be in the case of detailed balance. The separatrix of the lower-amplitude basin is shown in green.

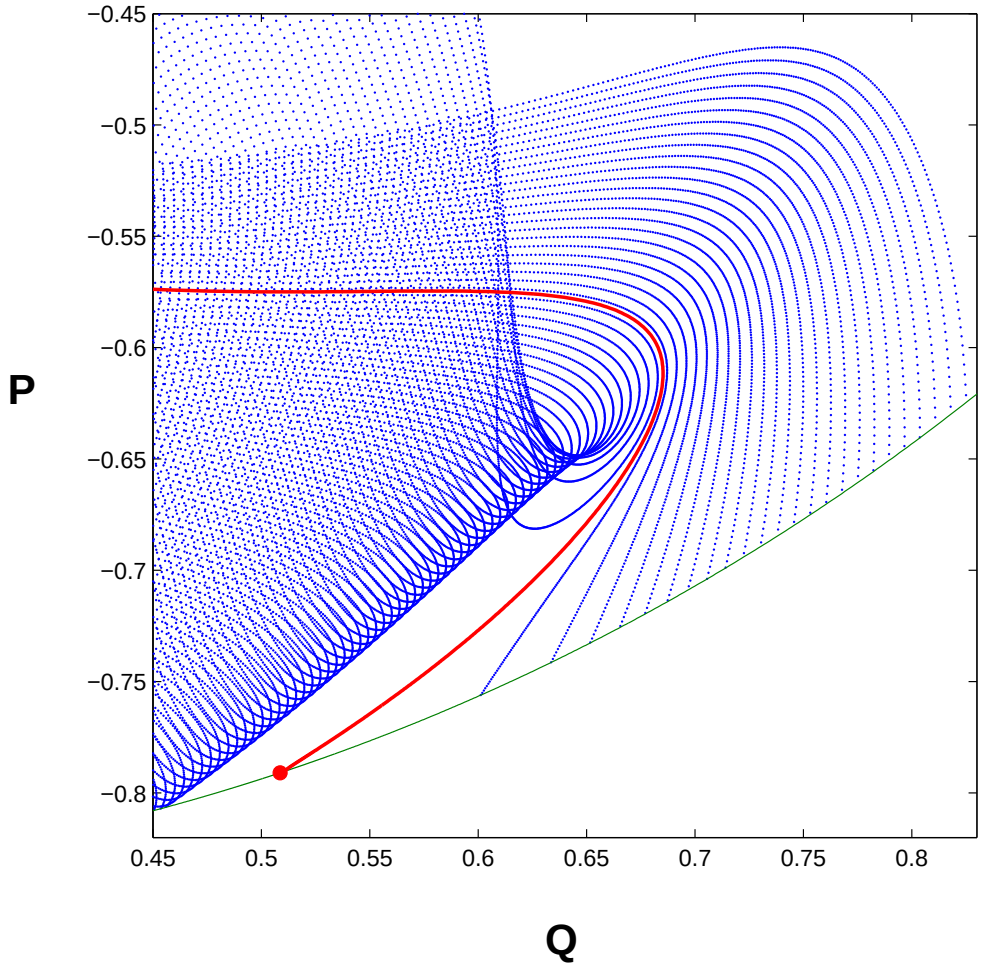


Figure C.2: Same features as in Fig. C.1, but for a narrower range of parameters, $[4.5, 5.985]$ and closer to the saddle. The cusp and the pair of caustics are clearly seen.

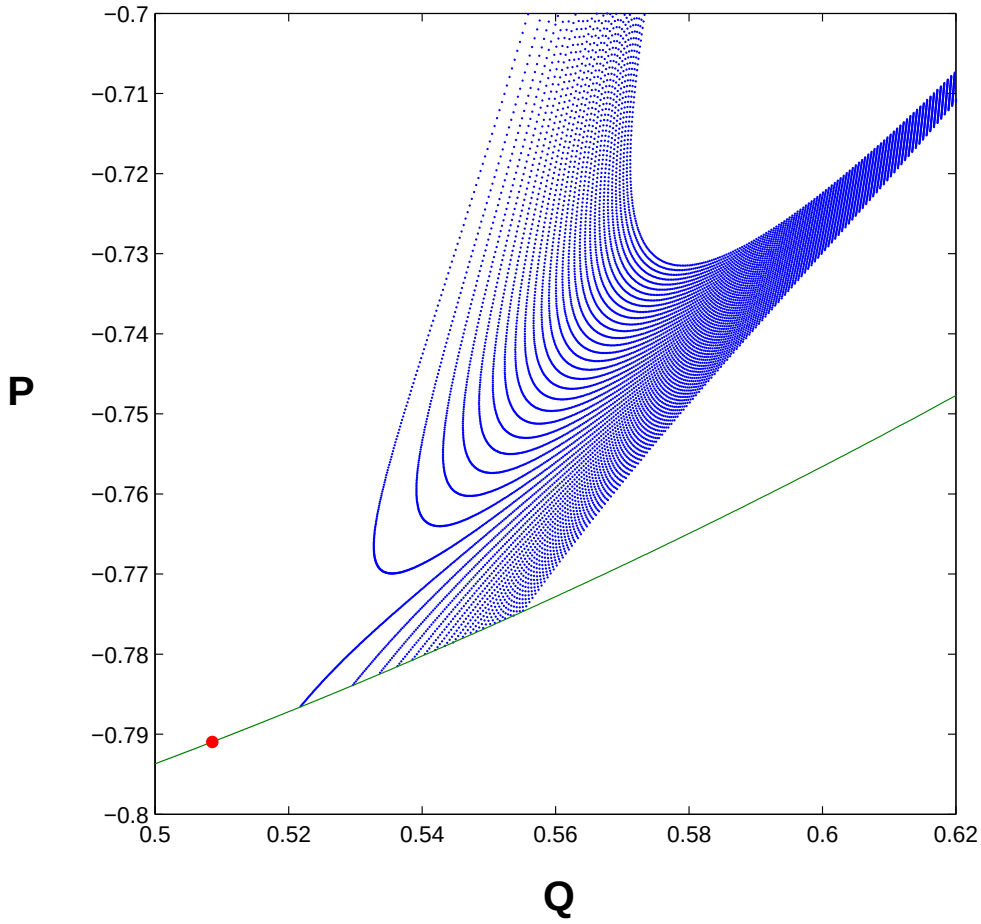


Figure C.3: A set of 50 trajectories in the range of parameters $[5.6328, 5.63398]$. The the region shown is even smaller then in Fig. C.2.

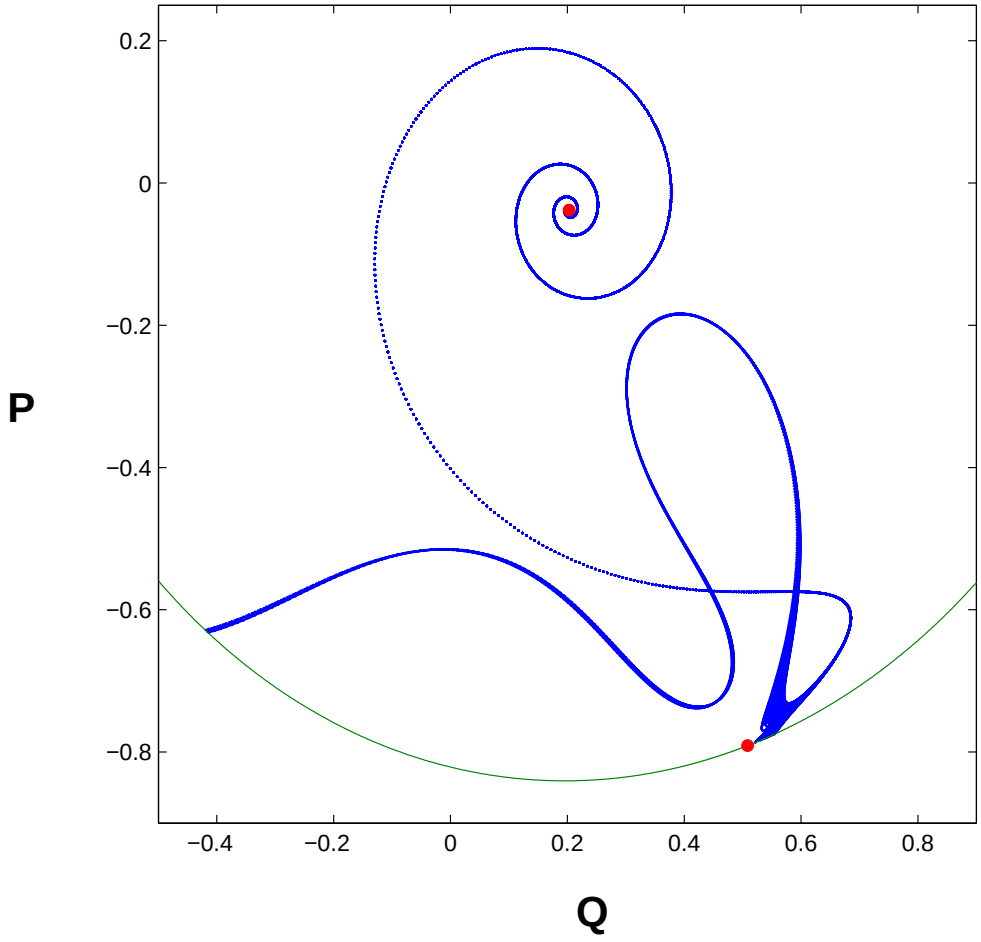


Figure C.4: Same features as in Fig. C.3, but showing the larger region of the (Q, P) space. Some trajectories hit the separatrix very close to the saddle and some, for a slightly different trajectory parameter, lead to a completely different region of the (Q, P) space.

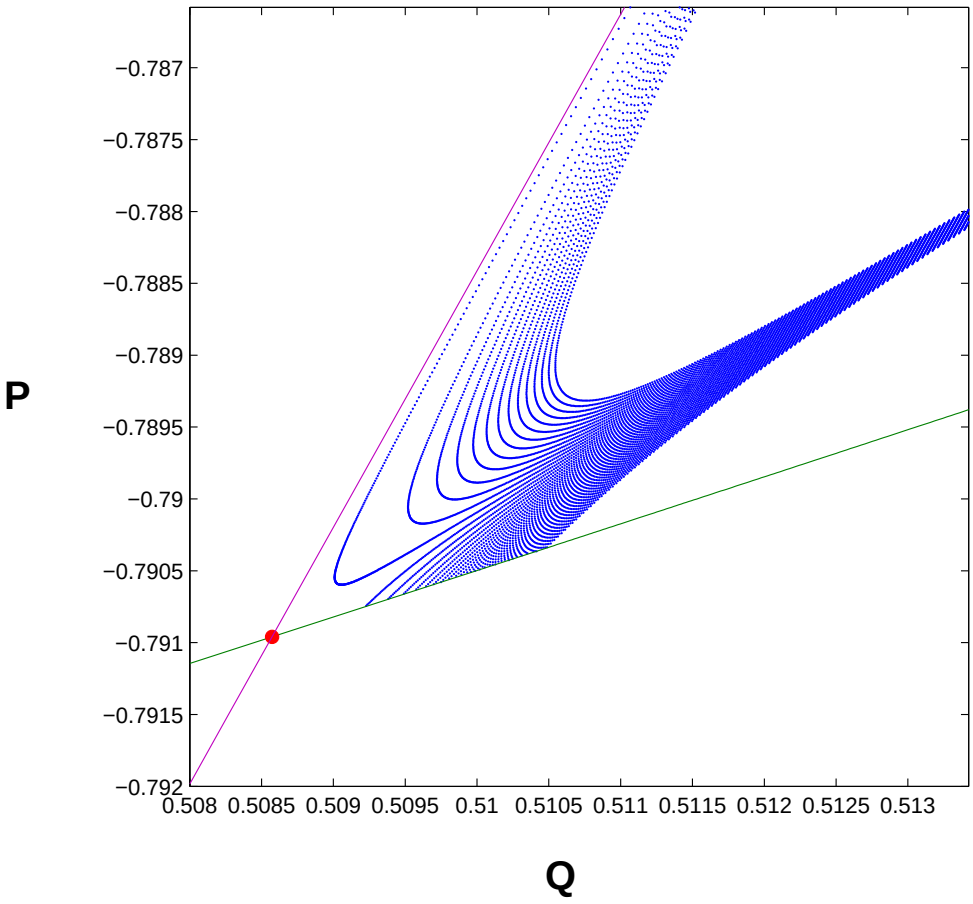


Figure C.5: A set of 50 trajectories in a very narrow window of trajectory parameters, $[5.633489536, 5.63348954776]$. The projection of the eigen-direction of the most-repelling eigenvector is also shown in magenta.

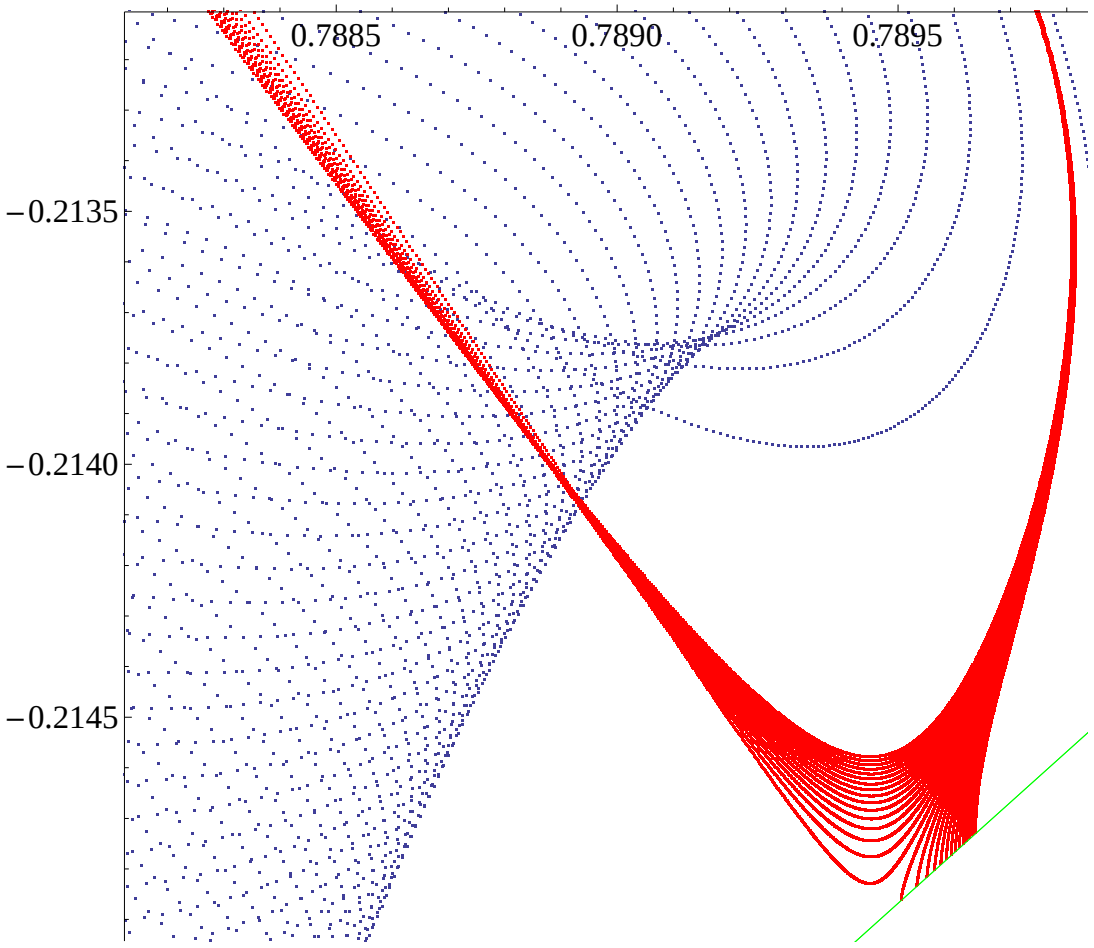


Figure C.6: Two sets of trajectories, one for a wide parameter region and one for a much narrower one. The saddle point is not shown.

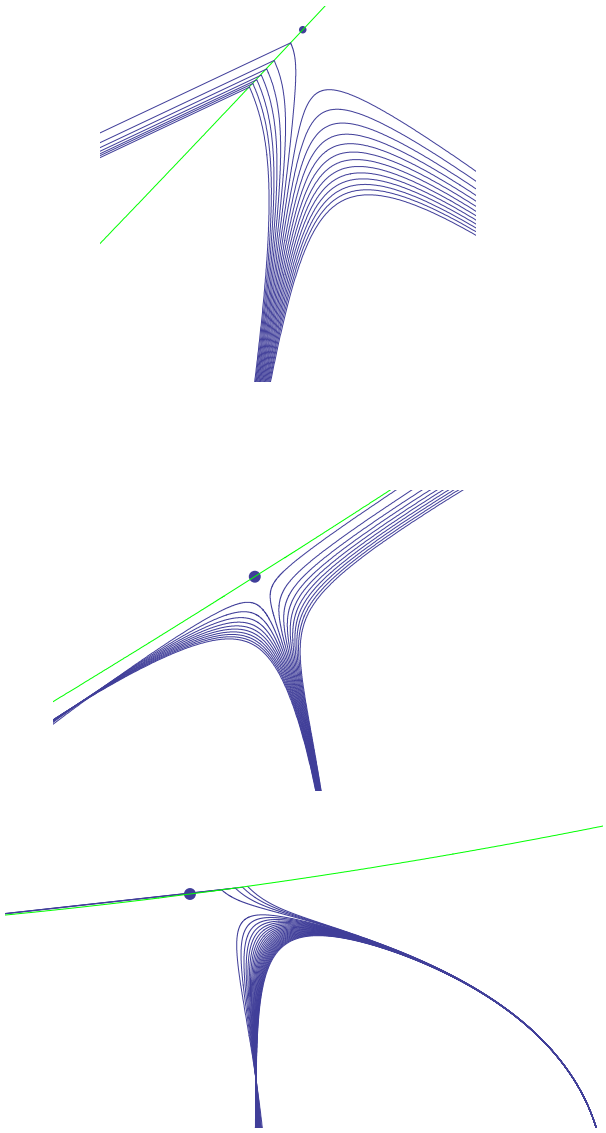


Figure C.7: Pattern of escape trajectories for up → down transitions at $\eta = 0.15$. Top to bottom: patterns of escape trajectories at $\mathcal{F}$ corresponding to 0.8%,0.9% and 0.989998% of the hysteresis respectively. The separatrix (green) and the saddle point (dark blue dot) are also shown.

Appendix D

Full Locking Criterion

Looking for fixed points of a set of m oscillators is an interesting problem by itself, so what follows is a brief summary of this problem.

The dynamical equations for phase differences, $\phi_i = \theta_i - \theta_{i-1}$ are:

$$\begin{aligned}
\dot{\phi}_{-1} &= \Delta\omega_{-1} + K_0 \cos\phi_0 - 2K_{-1}\sin\phi_{-1} + K_{-2}\sin\phi_{-2} \\
\dot{\phi}_0 &= \Delta\omega_0 + K_1 \sin\phi_1 + K_{-1}\sin\phi_{-1} \\
\\
\dot{\phi}_1 &= \Delta\omega_1 + K_2 \sin\phi_2 - 2K_1\sin\phi_1 - K_0\cos\phi_0 \\
\dot{\phi}_2 &= \Delta\omega_2 + K_3 \sin\phi_3 - 2K_2\sin\phi_2 + K_1\sin\phi_1 \\
... & \\
\dot{\phi}_{m-1} &= \Delta\omega_{m-1} + K_m \sin\phi_m - 2K_{m-1}\sin\phi_{m-1} + K_{m-2}\sin\phi_{m-2} \\
\dot{\phi}_m &= \Delta\omega_m + K_{m+1}\cos\phi_{m+1} - 2K_m\sin\phi_m + K_{m-1}\sin\phi_{m-1} \\
\\
\dot{\phi}_{m+1} &= \Delta\omega_{m+1} + K_{m+2}\sin\phi_{m+2} + K_m\sin\phi_m \\
\dot{\phi}_{m+2} &= \Delta\omega_{m+2} + K_{m+3}\sin\phi_{m+3} - 2K_{m+2}\sin\phi_{m+2} + K_{m+1}\cos\phi_{m+1}
\end{aligned}$$

Now isolate oscillators θ_0 through θ_m from the rest of the chain, by setting $K_0 = K_{m+1} = 0$. The EOM are:

$$\begin{pmatrix} \dot{\phi}_1 \\ \dot{\phi}_2 \\ \dot{\phi}_3 \\ ... \\ \dot{\phi}_m \end{pmatrix} = \begin{pmatrix} \Delta\omega_1 \\ \Delta\omega_2 \\ \Delta\omega_3 \\ ... \\ \Delta\omega_m \end{pmatrix} + \begin{pmatrix} -2K_1 & K_2 & 0 & 0 & ... & ... \\ K_1 & -2K_2 & K_3 & 0 & ... & ... \\ 0 & K_2 & -2K_3 & K_4 & 0 & ... \\ ... & ... & ... & ... & ... & ... \\ ... & ... & 0 & 0 & K_{m-1} & -2K_m \end{pmatrix} \begin{pmatrix} \sin\phi_1 \\ \sin\phi_2 \\ \sin\phi_3 \\ ... \\ \sin\phi_m \end{pmatrix}$$

The FP is:

$$-\begin{pmatrix} -2K_1 & K_2 & 0 & 0 & ... & ... \\ K_1 & -2K_2 & K_3 & 0 & ... & ... \\ 0 & K_2 & -2K_3 & K_4 & 0 & ... \\ ... & ... & ... & ... & ... & ... \\ ... & ... & 0 & 0 & K_{m-1} & -2K_m \end{pmatrix}^{-1} \begin{pmatrix} \Delta\omega_1 \\ \Delta\omega_2 \\ \Delta\omega_3 \\ ... \\ \Delta\omega_m \end{pmatrix} = \begin{pmatrix} \sin\phi_1 \\ \sin\phi_2 \\ \sin\phi_3 \\ ... \\ \sin\phi_m \end{pmatrix}$$

So there exist 2^m FPs if all the entries in the vector on the lhs are ≤ 1. In a work by Ermentrout and Kopell [94], it was proven that that if this is true, then one of the 2^m possible FP is a sink. Therefore, as long as the entries of the lhs are ≤ 1, there exists a global attractor.

We proceed and find the inverse of the square matrix. With the help of Mathematica I found the following patten:

$$\frac{-1}{m+1}\begin{pmatrix} \frac{m}{K_1} & \frac{m-1}{K_1} & \frac{m-2}{K_1} & \frac{m-3}{K_1} & ... & \frac{3}{K_1} & \frac{2}{K_1} & \frac{1}{K_1} \\ & \frac{2(m-1)}{K_2} & \frac{2(m-2)}{K_2} & \frac{2(m-3)}{K_2} & ... & \frac{2\cdot 3}{K_2} & \frac{2\cdot 2}{K_2} & \frac{2}{K_2} \\ & & \frac{3(m-2)}{K_3} & \frac{3(m-3)}{K_3} & ... & \frac{3\cdot 3}{K_3} & \frac{3\cdot 2}{K_3} & \frac{3}{K_3} \\ & & & & & & ... & ... \\ \frac{2}{K_{m-1}} & & & & & & & ... \\ \frac{1}{K_m} & \frac{2}{K_m} & & & & & & \frac{m}{K_m} \end{pmatrix} \begin{pmatrix} \Delta\omega_1 \\ \Delta\omega_2 \\ \Delta\omega_3 \\ ... \\ \Delta\omega_{m-1} \\ \Delta\omega_m \end{pmatrix} = \begin{pmatrix} \sin\phi_1 \\ \sin\phi_2 \\ \sin\phi_3 \\ ... \\ \sin\phi_{m-1} \\ \sin\phi_m \end{pmatrix} \tag{D.1}$$

the inverted matrix is symmetric about both diagonals.

We now derive the probability distribution for the value of each row. The following

point of view is taken. Let us imagine an ensemble of frequency layouts for a given layout of couplings. This will allows us to derive $P_j(s, K_j)$ - the probability of a value in the jth row for a given K_j. Then, we imagine an ensemble of couplings and obtain $P_j(s) = \int P_j(s, K)G(K)\, dK$. Here $G(K)$ denotes the distribution of Ks (distribution of Δs is denoted by $g(\Delta)$).

Each row has the form $S_j = \sum_{i=1}^{m} c_{j,i}\Delta_i$ where K_j is included in these cs. The distribution $P_j(s, K_j)$ is in principle given by $\int g(\Delta_1, \Delta_2, ...)\delta(S_j - s)\, d\Delta_1\, d\Delta_2...$, but this is a dead-end definition. However, its easy to calculate the characteristic function of this P_j first, and then inverse Fourier-Transform it. Define the characteristic function $\Phi_j = \int e^{ikS_j(\Delta_1,\Delta_2,...)}g(\Delta_1, \Delta_2, ...)\, d\Delta_1\, d\Delta_2....$ Its easy to see that this is indeed the characteristic function because upon expanding the exponential, we would obtain the series in which each coefficient contains the nth moment. Now, because all the Δs are IID, this Φ_j becomes $\Phi_j = \int e^{ikc_{j,1}\Delta_1}g(\Delta_1)\, d\Delta_1 \int e^{ikc_{j,2}\Delta_2}g(\Delta_2)\, d\Delta_2....$ So we need to calculate

$$\int_{-\infty}^{\infty} e^{ikc\Delta}g(\Delta)\, d\Delta = \int_{-\infty}^{\infty} e^{ikc\Delta}\frac{\lambda/\pi}{\lambda^2 + \Delta^2}.$$

We can see (from the contour integral for example) that the answer is $e^{-kc\lambda}$. Hence, the characteristic function for jth row is

$$\Phi_j = e^{-k\lambda\Sigma_j},$$

where $\Sigma_j = \sum_{i=1}^{m} c_{j,i}$ So, since the characteristic function of $g(\Delta)$ is $e^{-k\lambda}$, then $e^{-k\lambda\Sigma_j}$ is a characteristic function of

$$P_j(s, K_j) = \frac{(\lambda\Sigma_j)/\pi}{(\lambda\Sigma_j)^2 + s^2}.$$

We see from the matrix above that Σ_j is given by
$\frac{1}{K_j(m+1)}\left[j\,(1 + 2 + ... + (m - j)) + (m - j + 1)\,(1 + 2 + ... + j)\right] = \frac{j(m-j+1)}{2K_j}$. We should now compute $P_j(s) \equiv \int P_j(s, K)G(K)\, dK$. This is not easy. However, what we are after is the probability that a cluster of size m does not break. This means that $|s|$ is less then 1 in every bond. In other words, this probability is given by $\Pi_{j=1}^{m}\int_{-1}^{1} P_j(s)\, ds$.

Let us take the s-integral first, and then deal with the K-integral. So, exchanging the order of integration this way, the probability that a cluster of size m *does not* break is

$$\left(\int_0^\infty dKG(K)\int_{-1}^{1}\frac{(\lambda\Sigma_1)/\pi}{(\lambda\Sigma_1)^2+s^2}\,ds\right)...\left(\int_0^\infty dKG(K)\int_{-1}^{1}\frac{(\lambda\Sigma_m)/\pi}{(\lambda\Sigma_m)^2+s^2}\,ds\right)$$

$$\begin{aligned}
&= \left(\frac{4\mu}{\pi^2}\right)^m\prod_{j=1}^{m}\left(\int_0^\infty\frac{\arctan\left(\frac{2K}{\lambda j(m-j+1)}\right)}{\mu^2+K^2}\,dK\right)\\
&= \left(\frac{4}{\pi^2}\right)^m\prod_{j=1}^{m}\left(\int_0^\infty\frac{\arctan\left(a_jx\right)}{1+x^2}\,dx\right)=f\left(m,\frac{\mu}{\lambda}\right),
\end{aligned}$$

where $a_j=\frac{\mu}{\lambda}\frac{2}{j(m-j+1)}$. This is a bit complicated, so let us first bound it: obtain lower and upper limits. Note that the coefficient a_j is smallest in the middle of the cluster and biggest at the edges. Moreover, the area under $\frac{\arctan(ax)}{1+x^2}$ is smaller for smaller a (and approaches ≈ 2.47 as $a\to\infty$). So the lower bound can be obtained by setting $j=m/2$ and the upper bound by setting all $j=1$ for example. So, the upper limit is:

$$u(m,\mu/\lambda)=\left(\frac{4}{\pi^2}\right)^m\left(\int_0^\infty\frac{\arctan\left(\frac{\mu}{\lambda}\frac{2}{m}x\right)}{1+x^2}\,dx\right)^m,$$

and the lower limit is

$$\begin{aligned}
l(m,\mu/\lambda) &= \left(\frac{4}{\pi^2}\right)^m\left(\int_0^\infty\frac{\arctan\left(\frac{\mu}{\lambda}\frac{2}{m(m+2)}x\right)}{1+x^2}\,dx\right)^m\\
&\approx \left(\frac{4}{\pi^2}\right)^m\left(\int_0^\infty\frac{\arctan\left(\frac{\mu}{\lambda}\frac{2}{m^2}x\right)}{1+x^2}\,dx\right)^m,
\end{aligned}$$

where in the last term it was assumed that $m\gg 1$. Both expressions are functions of μ/λ and m. We plot these functions below.

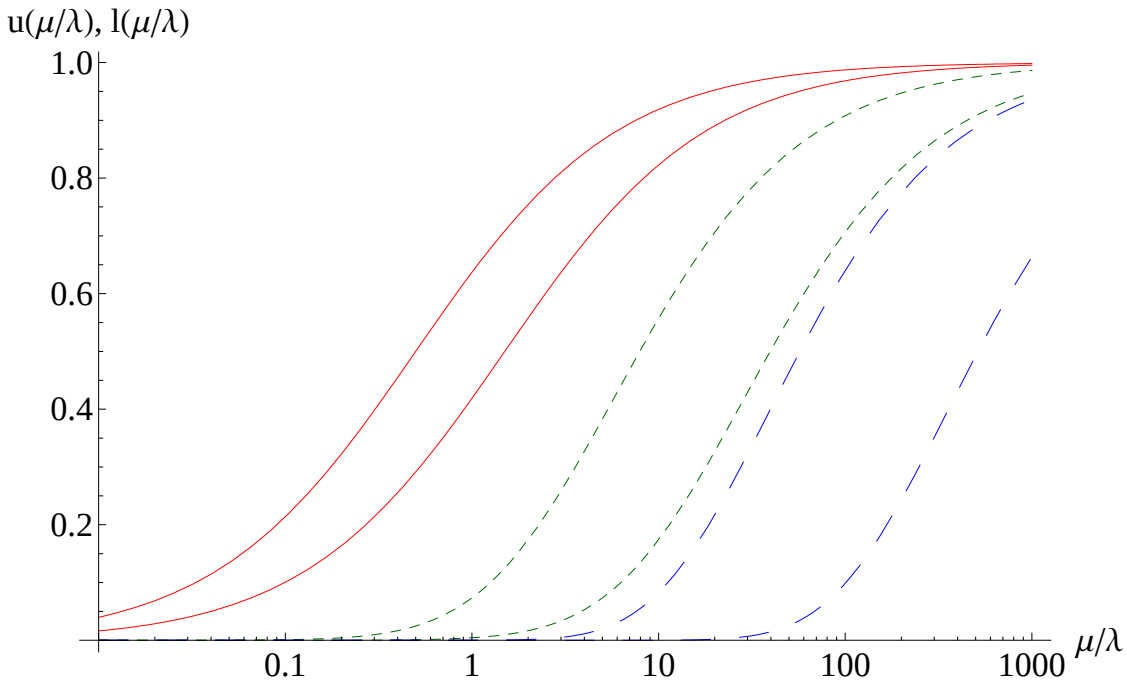


Figure D.1: The functions $u(m, \mu/\lambda)$ and $l(m, \mu/\lambda)$ plotted versus μ/λ for three values of m: 1 (solid line), 3 (short dashed curve) and 7 (long dashed curve). In each case, the upper and lower curve corresponds, respectively to functions u and l. Notice the logarithmic scale of the x-axis.

Bibliography

[1] R. Zwanzig, *Nonequilibrium Statistical Mechanics*, Oxford: Oxford University Press, 2001.

[2] E. M. Lifshitz and L. P. Pitayevskii, *Physical Kinetics*, Oxford: Pergamon Press, 1981.

[3] M. C. Cross and P. C. Hohenberg, *Rev. Mod. Phys.* **65**, 851 (1993).

[4] H.W.C. Postma, I. Kozinsky, A. Husain, and M. L. Roukes, *Appl. Phys. Lett.* **86**, 223105 (2005).

[5] R. Lifshitz and M. C. Cross, *Nonlinear Dynamics of Nanomechanical and Micromechanical Resonators* in *Reviews of Nonlinear Dynamics and Complexity* **1**, H. G. Schuster, editor, Weinhein: Wiley, 2008.

[6] I. Kozinsky, H. W. C. Postma, O. Kogan, A. Husain, and M. L. Roukes. *Phys. Rev. Lett.* **99**, 207201 (2007).

[7] M. Roukes, *Phys. World* **14**, 25 (2001).

[8] M. LaHaye, O. Buu, B. Camarota, and K. Schwab, *Science* **304**, 74 (2004).

[9] I. Katz, A. Retzker, R. Straub, and R. Lifshitz, *Phys. Rev. Lett.* **99**, 040404 (2007).

[10] I. Katz, R. Lifshitz, A. Retzker, and R. Straub, *New J. Phys.* **10**, 125023 (2008).

[11] D. Wahyu-Utami, and A. A. Clerk *Phys. Rev. A*, **78**, 042323 (2008).

[12] I. Kozinsky, *Nonlinear Nanoelectromechanical Systems*, Ph.D. Thesis, Caltech, (2007).

[13] H. A. Kramers, *Physica* **VII, 4**, 284 (1940).

[14] J. Kurkijarvi, *Phys. Rev. B* **6**, 832 (1972).

[15] R. H. Victora, *Phys. Rev. Lett.* **63**, 457 (1989).

[16] M. I. Dykman, and M. A. Krivoglaz, *Physica* **104A**, 480-494 (1980).

[17] O. A. Tretiakov, *Phys. Rev. B* **67**, 073303 (2003).

[18] K. A. Wiesenfeld, and E. Knobloch, *Phys. Rev. A* **26**, 2946 (1982).

[19] R. S. Maier and D. L. Stein, *Phys. Rev. Lett.* **86**, 3942 (2001).

[20] D. Ryvkine, M. I. Dykman, and B. Golding, *Phys. Rev. E* **69**, 061102 (2004).

[21] M. I. Dykman, I. B. Schwartz, and M. Shapiro, *Phys. Rev. E* **72** 021102 (2005).

[22] J. S. Aldrige and A. N. Cleland, *Phys. Rev. Lett.* **94**, 156403 (2005).

[23] R. Almog, S. Zaitsev, O. Shtempluck, and E. Buks, *Appl. Phys. Lett.* **90**, 013508 (2007).

[24] C. A. Stambaugh and H. B. Chan, *Phys. Rev. B* **73**, 172302 (2006).

[25] A.N. Cleland and M.L. Roukes, *J. Appl. Phys.* **92**, 2758 (2002).

[26] I. Siddiqi, R. Vijay, F. Pierre, C. M. Wilson, M. Metcalfe, C. Rigetti, L. Frunzio, and M. H. Devoret, *Phys. Rev. Lett.* **93**, 207002 (2004).

[27] Kihwan Kim, Myoung-Sun Heo, Ki-Hwan Lee, Hyoun-Jee Ha, Kiyoub Jang, Heung-Ryoul Noh, and Wonho Jhe, *Phys. Rev. A* 72, 053402 (2005).

[28] B. Abdo, E. Segev, O. Shtempluck, and E. Buks, *J. App. Phys.* **101**, 083909 (2007).

[29] R. Graham and H. Haken, *Z. Physik* **243**, 289-302 (1971).

[30] R. Graham and T. Tél, *Phys. Rev. A* **31**, 1109 (1985).

[31] M. I. Dykman, M. M. Millonas, V. N. Smelyanskiy, *Phys. Lett. A* **195**, 53-58 (1994).

[32] R. Graham and T. Tél, *Phys. Rev. A* **33**, 1322 (1986).

[33] A. Pikovsky, *Synchronization: A Universal Concept in Nonlinear Science*, New York: Cambridge University Press, 2001.

[34] F. Varela, J. P. Lachaux, E. Rodriguez, J. Martinerie, *Nature Rev. Neurosci.* **2**, 229 (2001).

[35] A. K. Engel and W. Singer, *Trends Cogn. Sci.* **5**, 16 (2001).

[36] M. Rabinovich, P. Varona, A. Selverston, and H. Abarbanel, *Rev. Mod. Phys.* **78**, 1213 (2006).

[37] K. Wiesenfeld, *Physica B* **222**, 315 (1996).

[38] B. R. Trees, V. Saranathan, and D. Stroud, *Phys. Rev. E* **71**, 016215 (2005).

[39] B. R. Trees, S. Natu, and D. Stroud, *Phys. Rev. B* **72**, 214524 (2005).

[40] A. G. Vladimirov, G. Kozyreff, and P. Mandel, *Europhys. Lett.* **61**, 613 (2003).

[41] J. F. Heagy, T. L. Carroll, and L. M. Pecora, *Phys. Rev. E* **50**, 1874 (1994).

[42] S. K. Esser, S. L. Hill, and G. Tononi, *Sleep* **30**, 1617 (2007).

[43] J. J. Collins, and I. N. Stewart, *J. Nonlinear Sci.* **3**, 349 (1993).

[44] L. Glass, *Nature* **410**, 277 (2001).

[45] J. A. Acerbón, L. L. Bonilla, C. J. Pérez-Vicente, F. Ritort, and R. Spigler, *Rev. Mod. Phys.* **77**, 137-185 (2005).

[46] Y. Kuramoto, in *International Symposium on Mathematical Problems in Theoretical Physics, Lecture Notes in Physics*, **39**, 420, H. Araki, editor, New York: Springer, 1975.

[47] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, Berlin: Springer, 1984.

[48] B. C. Daniels, S. T. M. Dissanayake, and B. R. Trees, *Phys. Rev. E* **67**, 026216 (2003).

[49] F. Rogister, and R. Roy, *Phys. Rev. Lett.* **98**, 104101 (2007).

[50] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group*, Reading: Addison-Wesley, 1992.

[51] S. H. Strogatz, and R. E. Mirollo, *J. Phys. A: Math. Gen.* **21**, L699-L705 (1988).

[52] S. H. Strogatz, R. E. Mirollo, *Physica D* **31**, 143-168 (1988).

[53] M. I. Dykman, and M. A. Krivoglaz, *Theory of Nonlinear Oscillator Interacting with a Medium* in Soviet Physics Reviews, **5**, 265-441, I. M. Khalatnikov, editor, New York: Harwood Academic, 1984.

[54] M. I. Freidlin, and A. D. Wentzell, *Random perturbations of dynamical systems*, New York: Springer, 1998.

[55] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, Berlin: Springer-Verlag, 2004.

[56] P. J. Holmes, and D. A. Rand, *Journal of Sound and Vibration*, **44**, 237-253 (1976).

[57] P. Holmes, and D. Rand, *International Journal of Non-linear Mechanics*, **15**, 449-458 (1980).

[58] A. H. Nayfeh, and D. T. Mook, *Nonlinear oscillations*, New York: John Wiley, 1979.

[59] A. H. Nayfeh, *Perturbation methods*, New York: John Wiley & Sons, 2000.

[60] N. N. Bogoliubov, and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Nonlinear Oscillations*, New York: Gordon and Breach Science Publishers, 1961.

[61] F. Verhulst, *Nonlinear Differential Equations and Dynamical Systems*, Berlin: Springer-Verlag, 1990.

[62] L. D. Landau, and E. M. Lifshitz, *Mechanics*, Oxford: Pergamon Press, 1989.

[63] O. Kogan, *Phys. Rev. E* **76**, 037203 (2007).

[64] R. K. Pathria, *Statistical Mechanics*, Oxford: Butterworth-Heinemann, 1996.

[65] M.I. Dykman, and M. A. Krivoglaz, *Sov. Phys. JETP* **50**, 30 (1979).

[66] R. Graham, *Z. Physik B* **26** 281-290 (1977).

[67] M. I. Dykman, *Phys. Rev. A.* **42**, 2020 (1990).

[68] M. I. Dykman in *Stochastic and Chaotic Dynamics in the Lakes*, D. S. Broomhead, E. A. Luchinskaya, P. V. E. McClintock, and T. Mullin, editors, New York: Melville, 1999.

[69] M. I. Dykman, and V. N. Smelyanskiy, *Superlattices and Microstructures* **23**, 495 (1998).

[70] R. Phythian, *J. Phys. A: Math. Gen.* **10**, 777 (1977).

[71] B. Jouvert, and R. Phythian, *Phys. Rev. A* **19**, 1350 (1979).

[72] R. P. Feynman, and A. R. Hibbs, *Quantum Mehcanics and Path Integrals*, New York: McGraw-Hill, 1965.

[73] L. S. Schulman, *Techniques and Applications of Path Integration*, Mineola: Dover, 2005.

[74] D. L. Stein, and R. S. Maier, *SIAM J. Appl. Math.* **57**, 752 (1997).

[75] V. P. Maslov, and M. V. Fedoriuk, *Semiclassical Approximations in Quanutm Mechanics*, Dordrecht: Reidel, 1981.

[76] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Berlin: Springer, 1978.

[77] A.P. Dmitriev, M.I. D'yakonov, *Zh. Eksp. Teor. Fiz.* **90**, 1430-1439, (1986); also in *Sov. Phys. JETP* **63**, 838 (1986).

[78] J. Guckenheimer, and P. Holmes, *Nonlinear Oscillators, Dynamical Systems and Bifurcations of Vector Fields*, New York: Springer-Verlag, 1987.

[79] V. Arnold, *Ordinary Differential Equations*, 3rd edition, New York: Springer, 1992.

[80] O. Kogan, *arXiv:0805.0972v2*. A shorter version co-authored with M. I. Dykman is to be submitted to *Phys. Rev. E.*

[81] V. A. Chinarov, M. I. Dykman, and V. N. Smelyanskiy, *Phys. Rev. E* **47**, 2448 (1993).

[82] R. S. Maier, and D. L. Stein, *Phys. Rev. Lett.* **85**, 1358 (2000).

[83] H. Whitney, *Ann. Math.* **62**, 374 (1955).

[84] V. I. Arnold, *Catastrophe theory*, Berlin: Springer, 1984.

[85] H. R. Jauslin, *Physica A* **144**, 179 (1987).

[86] M. V. Day, *Stochasitcs* **20**, 121 (1987).

[87] R. Graham and T. Tél, *Phys. Rev. Lett.* **52**, 9 (1984).

[88] V. N. Smelyanskiy, M. I. Dykman, and R. S. Maier, *Phys. Rev. E* **55**, 2369 (1996).

[89] O. Kogan, and M. I. Dykman, to be submitted.

[90] M. Bennett, M. F. Schatz, H. Rockwood, and K. Wiesenfeld, *Proc. Roy. Soc. Series A* **458**, 563 (2002).

[91] A. T. Winfree, *J. Theor. Biol.* **16**, 15 (1967).

[92] M. C. Cross, J. L. Rogers, R. Lifshitz, and A. Zumdieck, *Phys. Rev. E* **73**, 036205 (2006).

[93] S. H. Strogatz, *Physica D* **143**, 1-20 (2000).

[94] G. Ermentrout, and N. Kopell, *Siam J. Math. Anal.* **15**, 215 (1984).

[95] S. K. Ma, C. Dasgupta, and C. K. Hu, *Phys. Rev. Lett.* **43**, 1434 (1979).

[96] C. Dasgupta, and S. K. Ma, *Phys. Rev. B* **22**, 1305 (1980).

[97] D. S. Fisher, *Phys. Rev. B* **50**, 3799 (1994).

[98] D. S. Fisher, and A. P. Young, *Phys. Rev. B* **58**, 9131 (1998).

[99] E. Altman, Y. Kafri, A. Polkovnikov, and G. Refael, *Phys. Rev. Lett.* **93**, 150402 (2004).

[100] E. Altman, Y. Kafri, A. Polkovnikov, and G. Refael, *Phys. Rev. Lett.* **100**, 170402 (2008).

[101] O. Kogan, J. L. Rogers, M. C. Cross, G. Refael, submitted to *Phys. Rev. E.*; also on *arXiv:0810.3075v2*.

[102] K. Huang, *Statistical Mechanics*, New York: Wiley, 1963.

[103] T. Lee, G. Refael, M. C. Cross, O. Kogan, and J. L. Rogers, to appear.

[104] H. Sakaguchi, S. Shinomoto, and Y. Kuramoto, *Prog. Theor. Phys.* **77**, 1005 (1987).

[105] H. Daido, *Phy. Rev. Lett.* **61**, 231 (1988).