University of Reading
Department of Mathematics and Statistics

Stochastic Resonance for a Model with Two Pathways

Tommy Liu

Thesis submitted for the Degree of Doctor of Philosophy
September 2016

Abstract

In this thesis we consider stochastic resonance for a diffusion with drift given by a potential, which has two metastable states and two pathways between them. Depending on the direction of the forcing the height of the two barriers, one for each path, will either oscillate alternating or in synchronisation.

We consider a simplified model given by discrete and continuous time Markov Chains with two states. This was done for alternating and synchronised wells. The invariant measures are derived for both cases and shown to be constant for the synchronised case. A PDF for the escape time from an oscillatory potential is reviewed.

Methods of detecting stochastic resonance are presented, which are linear response, signal-to-noise ratio, energy, out-of-phase measures, relative entropy and entropy. A new statistical test called the conditional Kolmogorov-Smirnov test is developed, which can be used to analyse stochastic resonance.

An explicit two dimensional potential is introduced, the critical point structure derived and the dynamics, the invariant state and escape time studied numerically.

The six measures are unable to detect the stochastic resonance in the case of synchronised saddles. The distribution of escape times however not only shows a clear sign of stochastic resonance, but changing the direction of the forcing from alternating to synchronised saddles an additional resonance at double the forcing frequency starts to appear. The conditional KS test reliably detects the stochastic resonance even for forcing quick enough and for data so sparse that the stochastic resonance is not obvious directly from the histogram of escape times.

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Tommy Liu

Acknowledgement

I would like to thank Tobias Kuna for supervising this thesis; Valerio Lucarini for co-supervising; Tristan Pryer, Horatio Boedihardjo, Martin Kolb and the late Professor Alexei Likhtman for being on the Monitoring Committee; Jochen Broecker and Ostap Hryniv for being the examiners on my viva; Peter Imkeller for helpful discussions; Pawel Stasiak for introducing me to the Meteorology Computer Clusters; Peta-Ann King and Sue Davis for their pastoral care; the EPSRC for funding and finally to my family and friends for their support over the years.

Tommy Liu
September 2016
University of Reading

Contents

List of Figures

Introduction

Outline of Problem

Consider the following problem. Let X_t^ϵ be the random variable describing the trajectory of a diffusion process in $\mathbb{R}^r$ where t is the time and ϵ^2 is the variance level. More precisely we consider processes described by the following type of stochastic differential equation

$$dX_t^\epsilon = b\left(X_t^\epsilon, t\right) dt + \epsilon\, dW_t$$

where $b : \mathbb{R}^r \times \mathbb{R} \longrightarrow \mathbb{R}^r$ and W_t is a Wiener process in $\mathbb{R}^r$. We suppose that the drift term b has the form

$$b(x,t) = -\nabla V_0(x) + F\cos\Omega t$$

where $F, x \in \mathbb{R}^r$ and $V_0 : \mathbb{R}^r \longrightarrow \mathbb{R}$ is called the unperturbed potential. We consider unperturbed potentials with two or more minimas (wells). Most importantly, we consider potentials where there are multiple pathways between the wells. To our knowledge systems with two pathways have not been studied in the context of stochastic resonance.

Consider the case $\Omega = 0$ and where the noise ϵ is very small. The particle will stay very close to one of the wells of the potential and will occasionally escape to the other well. The time of the actual transition from one well to the other is very short compare to the time it stays in any particular well.

Now consider the case where $\Omega > 0$. For particular choices of $\Omega > 0$ and $\epsilon > 0$, these transitions between the two wells will become synchronised with the driving frequency Ω. This is called stochastic resonance. Thus the term *noise induced synchronisation* was used for systems where the amplitude of the forcing F was not large [1, 2] (see also the discussions in [3]). New insights into the exact manner of these synchronised transitions will be studied in this thesis, which may be more appropriate in light of the results obtained in this thesis.

For small noise ϵ, one would expect that stochastic resonance depends only on the essential properties of the system, such as the height difference between the wells and the pathways for escape. We investigate what effects these multiple pathways have on the appearance of stochastic resonance. Varying F, Ω and ϵ should thus reveal the qualitative structure of the unperturbed potential V_0. In this thesis we test this paradigm by studying a two dimensional example with two wells and two independent pathways between them, see Chapter 5.

Historical Background

Stochastic resonance has attracted interest among mathematicians and physicist. An overview of the studies that have occurred in both physics and mathematics are given here.

Physical Background

Stochastic resonance was first observed in 1981 [4–6]. The first example [4] considered transitions between two metastable states to model the cyclic occurrences of ice ages. Since then many examples of stochastic resonance were found in optics [7–10], electronics [11–19], neuronal systems [20], quantum systems [21,22] and paddlefish [23,24]. Stochastic resonance could be thought of as quasi-deterministically periodic transition between two metastable states. For example, the climate of the Earth could be modelled by two states. There is a state corresponding to an Ice Age and another corresponding to the opposite of an Ice Age, a so-called "Hot Age". As the Earth's climate cyclically changes many times between Cold Ages and Hot Ages, its behaviour could be modelled by stochastic resonance.

A range of techniques for example linear response [25,26], signal-to-noise ratio [27,28] and distribution of escape times [28–30] were used to define, analyse and study stochastic resonance. These techniques along with other examples of stochastic resonance are reviewed in the long overview paper by Gammaitoni, Hänggi, Jung and Marchesoni [31]. We will evaluate the usefulness of some of these techniques for our problem, see Chapter 7.

Mathematical Background

There are various mathematical studies of stochastic resonance. These often involve different orders of approximations for small noise levels. The first and second order of approximations are discussed below. Adiabatic large deviation is also presented.

In the first leading order of approximation, a key element of study is to control the escape times from the wells as given by the so called large deviation theory, see the monograph of Freidlin and Wentzell [32]. The distribution of the exit time was derived by Day in [33] and by Galves, Kifer, Olivieri and Vares [34–36]. To go beyond leading order has been much more difficult for the transition problem between two wells as WKB theory could up to now not be rigorously applied.

The next order of approximation was rigorously derived by Bovier, Eckhoff, Gayrard, Klein [37] and Berglund and Gentz [38] using techniques from potential theory. Berglund and Gentz in a series of papers studied the situation of low, non-quadratic barriers and drifts not given by autonomous potentials [3,38]. A review of different techniques used to derive Kramers' formula can be found in the review paper [39].

In [40] Friedlin considered stochastic resonance in the adiabatic regime. This means the diffusion can effectively be described by a Markov process which describes the jumps between wells. This problem was revisited by Hermann, Imkeller and Pavlyukevich, see Chapter 4 in [41] and references therein, to derive results uniformly for varying time scale

to identify the optimal resonance point asymptotically for small noise even outside the adiabatic regime leading to different logarithmic corrections including the famous cycling effect discovered by Day [42], see also [43] for the connection with stochastic resonance. Escape time outside of adiabatic regime is studied in [44].

As mentioned above in leading order the transitions of the diffusion process X_t^ϵ between the wells can be approximated by a two state Markov Chain $Y_t^\epsilon = \pm 1$ which have been studied [41,45–47]. Further comparative studies of the stochastic resonance for the diffusion case X_t^ϵ versus the Markov Chain Y_t^ϵ case were done by Hermann, Imkeller, Pavlyukevich and Peithmann in [48–51]. A collection of papers on comparative studies between stochastic resonance in diffusion and Markov Chains can be found in the monograph [41]. One of the main conclusions in [41, 49–51] is rigorously showing that using linear response and signal-to-noise ratio to analyse stochastic resonance in the diffusion case X_t^ϵ gives a different result to analysing the Markov Chain case $Y_t^\epsilon = \pm 1$ with the same techniques even asymptotically in the small noise limit. Other common methods used to study stochastic resonance include invariant measures and Fourier transforms. We consider six measures of stochastic resonance frequently used and considered by Pavlyukevich in his thesis [41, 45] which are linear response, signal-to-noise ratio, energy, out-of-phase measure, relative entropy and entropy.

In this thesis we will study stochastic resonance on a two dimensional toy model, in both the diffusion and Markov Chain cases, and where there are two independent pathways between the wells going through two different saddles. The escape times and the six measures of stochastic resonance introduced above are studied.

Summary of Research

In Chapter 1 we review the first model in which stochastic resonance was observed, that is, we are considering the unperturbed potential

$$V_0(x) = \frac{x^4}{4} - a\frac{x^2}{2}$$

and the corresponding SDE

$$dX_t^\varepsilon = [-\nabla V_0 + F\cos(\Omega t)]\,dt + \epsilon\,dW_t.$$

In one dimension the escape time can be explicitly computed as the solution to an ODE and using Laplace method asymptotic formulas can be derived. In Chapter 2 a review of large deviation theory and results concerning escape times are given. In Chapter 2.2 further results, based on potential theory, are given and the analogue of Kramers' formula for our case is presented. In Chapter 3 discrete and continuous time Markov Chains are considered. The associated invariant measures and the relaxation time to this invariant measure is derived for alternating and synchronised wells. The probability density function of escape times is derived as well. In Chapter 4 the six measures used to analyse stochastic resonance mentioned above are introduced. Furthermore, methods used to study escape

times are given and in particular a new version of the Kolmogorov-Smirnov test suitable for this problem is discussed. In Chapter 5 the main model under consideration in this thesis is studied, which has two wells and two saddles. The two wells are connected through to independent pathways each through one of the saddles. Due to its form, we nicknamed it the Mexican Hat Toy Model

$$V_0(x,y) = \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 \quad \text{where} \quad r = \sqrt{x^2+y^2}.$$

We rigorously derive the qualitative structure of the potential with and without external forcing. In Chapter 6 the numerical methods used to simulate the associated SDE

$$\begin{aligned} dx &= \left[-\frac{\partial V_0}{\partial x} + F_x \cos \Omega t \right] dt + \epsilon \, dw_x \\ dy &= \left[-\frac{\partial V_0}{\partial y} + F_y \cos \Omega t \right] dt + \epsilon \, dw_y \end{aligned}$$

are discussed and non rigorous estimates of all relevant error sources are given necessary to be confident about the precision of the simulation needed. The dw_x and dw_y are x and y components of the two dimensional Wiener processes. The numerical algorithm used is the Euler method [52] which is sufficiently accurate for our purposes. In Chapter 7 the results from simulating the SDE are presented and interpreted. The six measures are studied and the quality of the approximation by the aforementioned Markov chains is tested using the Kolomogorov-Smirnov test developed. The results were repeated in a sparse data context.

In Chapter 7 the main findings and conclusions of this thesis are presented. It is shown that the six measures are unable to detect stochastic resonance in the case of synchronised saddles. The six measures show no sharp signature as the saddles change from alternating to synchronised saddles. This is due to the fact that the invariant measures are constant for synchronised saddles. By contrast, not only did the distribution of escape times show a signature for stochastic resonance with synchronised saddles; the distribution of escape times did show a clear signature as the saddles change from alternating to synchronised, by exhibiting signatures which we call the Single, Intermediate and Double Frequency. The newly developed conditional Kolomogorov-Smirnov test was shown to be a good method to analyse the statistics of the escape times.

This thesis then finishes with a conclusion of the results obtained. In Appendix A the conventions used are collected. In Appendix B the methods used to calculate the Fourier transforms and the escape times are explained.

Chapter 1

Stochastic Resonance

The earliest known and simplest example of stochastic resonance is reviewed. This was done in 1981 [4]. Properties about its escape times are derived. Estimates for the resonance noise level ϵ_{res} are given. The techniques involved include a review of Laplace method. This study only works for ϵ in the small noise approximation. Only one dimensional systems will be studied in this Chapter. Deducing properties about the underlying potential is trivial.

1.1 Laplace Method

The main technique used to study exit times in one dimension is the so called Laplace Method. For completeness and to get a better understanding of the mechanism we are going to study, a proof will be provided later on.

Theorem 1.1. *(Laplace Method) Let $f : [a, b] \to \mathbb{R}$ be twice differentiable on $[a, b]$. Let $x_0 \in (a, b)$ be unique such that $f(x_0) = \max_{x\in(a,b)} f(x)$. Assuming $f''(x)$ is continuous on $[a, b]$ with $f'(x_0) = 0$ and $f''(x_0) < 0$ then*

$$\lim_{n\to\infty} \left(\frac{\int_a^b e^{nf(x)}dx}{e^{nf(x_0)}\sqrt{\frac{2\pi}{n(-f''(x_0))}}} \right) = 1.$$

A Corollary follows from Laplace Method as a special case of Theorem 1.1.

Corollary 1.2. *Let $f : [a, b] \to \mathbb{R}$ be twice differentiable on $[a, b]$. Let $x_0 = a$ or $x_0 = b$ be unique such that $f(x_0) = \max_{x\in[a,b]} f(x)$. Assuming $f''(x)$ is continuous on $[a, b]$ with $f'(x_0) = 0$ and $f''(x_0) < 0$ then*

$$\lim_{n\to\infty} \left(\frac{\int_a^b e^{nf(x)}dx}{\frac{1}{2}e^{nf(x_0)}\sqrt{\frac{2\pi}{n(-f''(x_0))}}} \right) = 1.$$

We recall Taylor's Remainder Theorem which is needed in the proofs.

Theorem 1.3. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is $(n+1)$ times differentiable on $\mathbb{R}$. Let $x, a \in \mathbb{R}$, with $x > a$ then f can be expressed as*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^n(a)}{n!}(x-a)^n + R_{n+1}(x)$$

where R_{n+1} the remainder can be expressed as

$$\textit{Integral Form} \quad R_{n+1}(x) = \frac{1}{n!}\int_a^x (x-t)^n f^{n+1}(t)dt$$

$$\textit{Lagrange Form} \quad R_{n+1}(x) = \frac{f^{n+1}(\xi)}{(n+1)!}(x-a)^{n+1} \ , \ \xi \in [a, x].$$

The following simple Lemma is also needed in the proof of Laplace Method.

Lemma 1.4. *Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous. Let x_0 be a unique maximum such that $f(x_0) = \max_{x\in[a,b]} f(x)$, then for any fixed $\delta > 0$, there exists an $\eta > 0$ such that for any $s \notin (x_0 - \delta, x_0 + \delta)$ we have*

$$\eta \leq f(x_0) - f(s).$$

Now we review proofs of the methods needed.

Proof of Lemma 1.4. We know that x_0 is the unique maximum, which means

$$0 < f(x_0) - f(s)$$

for any $s \notin (x_0 - \delta, x_0 + \delta)$. This means the infinum of the set is bounded by zero

$$\inf\{f(x_0) - f(s) : s \notin (x_0 - \delta, x_0 + \delta)\} \geq 0.$$

Suppose that the infinum of the set is zero

$$\inf\{f(x_0) - f(s) : s \notin (x_0 - \delta, x_0 + \delta)\} = 0$$

and yet all elements of the set are strictly greater than zero. This means some members would be arbitrarily close to zero,

$$0 < f(x_0) - f(s) < \epsilon$$

where ϵ is arbitrarily small. There exists a sequence

$$(s_n)_{n\geq 1} \subset [a, b]\backslash(x_0 - \delta, x_0 + \delta) = [a, x_0 - \delta] \cup [x_0 + \delta, b]$$

such that

$$0 < f(x_0) - f(s_n) < \frac{1}{n} \Longrightarrow f(x_0) < f(s_n) + \frac{1}{n}.$$

But this sequence is in a compact set, which must have a subsequence which converges to a member $s' \in [a, x_0 - \delta] \cup [x_0 + \delta, b]$, that is

$$f(x_0) \leq f(s')$$

which contradicts the fact x_0 is the unique maximum. This implies that

$$\inf \{f(x_0) - f(s) : s \notin (x_0 - \delta, x_0 + \delta)\} > 0$$

so the $\eta > 0$ as in the assertion of the Lemma must exist. □

Proof of Theorem 1.1. A differentiable function is also a continuous function. Since $f(x_0) = \max_{x \in [a,b]} f(x)$ we can say $f'(x_0) = 0$. Using the Taylor's Remainder Theorem we can rewrite $f(x)$ for $x \in [x_0, x_0 + \delta]$ for some $\delta > 0$ and $\xi \in [x_0, x]$ as

$$f(x) = f(x_0) + \frac{f''(\xi)}{2}(x - x_0)^2.$$

We can obtain an upper and lower bound for $f''(\xi)$ by exploiting its continuity on $[a, b]$. Since $x \in [x_0, x_0 + \delta]$ we must also have $\xi \in [x_0, x_0 + \delta]$. So

$$|\xi - x_0| \leq \delta.$$

For any $\epsilon > 0$ and for a sufficiently small δ, we can have

$$|f''(\xi) - f''(x_0)| < \epsilon.$$

This means we can say

$$\begin{aligned} -\epsilon < f''(\xi) - f''(x_0) < \epsilon \\ f''(x_0) - \epsilon < f''(\xi) < f''(x_0) + \epsilon \end{aligned}$$

which gives

$$f(x) \leq f(x_0) + \frac{1}{2}(f''(x_0) + \epsilon)(x - x_0)^2 \tag{1.1}$$

$$f(x) \geq f(x_0) + \frac{1}{2}(f''(x_0) - \epsilon)(x - x_0)^2. \tag{1.2}$$

We start with the lower bound for $f(x)$ as in Equation 1.1

$$\begin{aligned} \int_a^b e^{nf(x)} dx &\geq \int_{x_0 - \delta}^{x_0 + \delta} e^{nf(x)} dx \\ &\geq e^{nf(x_0)} \int_{x_0 - \delta}^{x_0 + \delta} e^{\frac{n}{2}(f''(x_0) - \epsilon)(x - x_0)^2} dx \\ &= e^{nf(x_0)} \frac{1}{\sqrt{n(-f''(x_0) + \epsilon)}} \int_{-\delta\sqrt{n(-f''(x_0)+\epsilon)}}^{+\delta\sqrt{n(-f''(x_0)+\epsilon)}} e^{-\frac{1}{2}y^2} dy \end{aligned}$$

where we have made a transformation using

$$y = \sqrt{n(-f''(x_0) + \epsilon)}(x - x_0).$$

Dividing both sides by $e^{nf(x_0)}\sqrt{\frac{2\pi}{n(-f''(x_0))}}$ gives

$$\left(\frac{\int_a^b e^{nf(x)}dx}{e^{nf(x_0)}\sqrt{\frac{2\pi}{n(-f''(x_0))}}} \right) \geq \frac{1}{\sqrt{2\pi}} \sqrt{\frac{-f''(x_0)}{-f''(x_0) + \epsilon}} \int_{-\delta\sqrt{n(-f''(x_0)+\epsilon)}}^{+\delta\sqrt{n(-f''(x_0)+\epsilon)}} e^{-\frac{1}{2}y^2} dy. \tag{1.3}$$

Using Lemma 1.4 we can say that for any fixed δ, there exists an $\eta > 0$ such that for any $s \notin (x_0 - \delta, x_0 + \delta)$ we have

$$\eta \leq f(x_0) - f(s).$$

So we can proceed with

$$\begin{aligned}
\int_a^b e^{nf(x)}dx &= \int_a^{x_0-\delta} e^{nf(x)}dx + \int_{x_0-\delta}^{x_0+\delta} e^{nf(x)}dx + \int_{x_0+\delta}^b e^{nf(x)}dx \\
&\leq \int_a^{x_0-\delta} e^{n(f(x_0)-\eta)}dx + \int_{x_0-\delta}^{x_0+\delta} e^{nf(x)}dx + \int_{x_0+\delta}^b e^{n(f(x_0)-\eta)}dx \\
&= (x_0 - \delta - a)e^{n(f(x_0)-\eta)} + (b - x_0 - \delta)e^{n(f(x_0)-\eta)} + \int_{x_0-\delta}^{x_0+\delta} e^{nf(x)}dx \\
&= (b - a - 2\delta)e^{n(f(x_0)-\eta)} + \int_{x_0-\delta}^{x_0+\delta} e^{nf(x)}dx.
\end{aligned}$$

Now we use the upper bound for $f(x)$ from Equation 1.2. So

$$\begin{aligned}
\int_a^b e^{nf(x)}dx &\leq (b - a)e^{n(f(x_0)-\eta)} + e^{nf(x_0)} \int_{x_0-\delta}^{x_0+\delta} e^{\frac{n}{2}(f''(x_0)+\epsilon)(x-x_0)^2}dx \\
&\leq (b - a)e^{n(f(x_0)-\eta)} + e^{nf(x_0)} \int_{-\infty}^{+\infty} e^{\frac{n}{2}(f''(x_0)+\epsilon)(x-x_0)^2}dx \\
&= (b - a)e^{n(f(x_0)-\eta)} + e^{nf(x_0)} \sqrt{\frac{2\pi}{n(-f''(x_0) - \epsilon)}}
\end{aligned}$$

where ϵ is chosen small enough so that $(f''(x_0) + \epsilon) < 0$ is still negative. Now divide both sides by $e^{nf(x_0)}\sqrt{\frac{2\pi}{n(-f''(x_0))}}$ which gives

$$\left(\frac{\int_a^b e^{nf(x)}dx}{e^{nf(x_0)}\sqrt{\frac{2\pi}{n(-f''(x_0))}}} \right) \leq \left((b - a)e^{-n\eta}\sqrt{\frac{n(-f''(x_0))}{2\pi}} + \sqrt{\frac{-f''(x_0)}{-f''(x_0) - \epsilon}} \right). \tag{1.4}$$

Now using the other bound for $\left(\frac{\int_a^b e^{nf(x)}dx}{e^{nf(x_0)}\sqrt{\frac{2\pi}{n(-f''(x_0))}}}\right)$ from Equation 1.3 together with Equation 1.4 gives

$$\frac{1}{\sqrt{2\pi}}\sqrt{\frac{-f''(x_0)}{-f''(x_0)+\epsilon}}\int_{-\delta\sqrt{n(-f''(x_0)+\epsilon)}}^{+\delta\sqrt{n(-f''(x_0)+\epsilon)}} e^{-\frac{1}{2}y^2}dy \leq \left(\frac{\int_a^b e^{nf(x)}dx}{e^{nf(x_0)}\sqrt{\frac{2\pi}{n(-f''(x_0))}}}\right)$$
$$\leq \left((b-a)e^{-n\eta}\sqrt{\frac{n(-f''(x_0))}{2\pi}}+\sqrt{\frac{-f''(x_0)}{-f''(x_0)-\epsilon}}\right).$$

Now we can take the limit as $n \to +\infty$ which gives

$$\sqrt{\frac{-f''(x_0)}{-f''(x_0)+\epsilon}} \leq \lim_{n\to\infty}\left(\frac{\int_a^b e^{nf(x)}dx}{e^{nf(x_0)}\sqrt{\frac{2\pi}{n(-f''(x_0))}}}\right) \leq \sqrt{\frac{-f''(x_0)}{-f''(x_0)-\epsilon}}$$

after noting that $\lim_{n\to 0} e^{-n\eta}\sqrt{n} = 0$ and $\lim_{n\to\infty}\int_{-\delta\sqrt{n(-f''(x_0)+\epsilon)}}^{+\delta\sqrt{n(-f''(x_0)+\epsilon)}} e^{-\frac{1}{2}y^2}dy = \sqrt{2\pi}$. Since ϵ can be chosen to be arbitrarily small using the Sandwich Theorem gives

$$\lim_{n\to\infty}\left(\frac{\int_a^b e^{nf(x)}dx}{e^{nf(x_0)}\sqrt{\frac{2\pi}{n(-f''(x_0))}}}\right) = 1.$$

□

Proof of Corollary 1.2. If $x_0 = a$ then the proof is the same but with a few adjustments. In other words, the interval $[x_0 - \delta, a]$ does not need to be considered as it is outside the region of integration.

$$\int_{x_0-\delta}^{x_0+\delta} \to \int_{x_0}^{x_0+\delta}$$
$$\int_a^{x_0-\delta} \to 0$$
$$\int_{x_0+\delta}^{b} \to \int_{x_0+\delta}^{b}$$

and the resulting computation would give the extra factor of $\frac{1}{2}$ after using

$$\int_{x_0}^{x_0+\delta} e^{\frac{n}{2}(f(x_0)\pm\epsilon)}dy = \frac{1}{2}\int_{x_0-\delta}^{x_0+\delta} e^{\frac{n}{2}(f(x_0)\pm\epsilon)}dy.$$

A similar argument holds for $x_0 = b$. Note that the computation shows that only a small neighbourhood of x_0 is relevant asymptotically and that the average term is exponentially small in n. □

1.2 One Dimensional Potential

The potential we are interested in is

$$V_0 = \frac{x^4}{4} - a\frac{x^2}{2}$$

where $a > 0$. When this is given a driving frequency it is[1]

$$\begin{aligned} V_t &= V_0 - Fx\cos\Omega t \\ &= \frac{x^4}{4} - a\frac{x^2}{2} - Fx\cos\Omega t \end{aligned}$$

which when the forcing is zero, the potential has two wells at $x_0 = \pm\sqrt{a}$. The SDE we want to study is

$$\begin{aligned} \frac{dx}{dt} &= -\nabla V_t + \epsilon\frac{dw}{dt} \\ dx &= \left[x(a-x^2) + F\cos\Omega t\right]dt + \epsilon\, dw \end{aligned}$$

where w is a one dimensional Wiener process. Consider a realisation of the trajectory $x(t)$ starting at $x(0) = y$. Its escape time from the left well τ_1 and right well τ_2 are defined as

$$\tau_1(y) = \inf\{t : x(t) = 0 \quad \text{and} \quad x(0) = y\} \quad \text{where} \quad y \in (-\infty, 0) \tag{1.5}$$

$$\tau_2(y) = \inf\{t : x(t) = 0 \quad \text{and} \quad x(0) = y\} \quad \text{where} \quad y \in (0, +\infty). \tag{1.6}$$

Note that the trajectory $x(t)$ is related to the escape times τ_1 and τ_2. Define a new quantity by

$$f_n^i(y) = \langle \tau_i(y)^n \rangle$$

with $i = 1, 2$ for the two wells and $n = 1, 2, \ldots$. Note $\langle\cdot\rangle$ denotes the mean average over all realisations. Also note that the nth moment is being used here. In [4] a method by Gihman and Skorohod [53] was used to derive the following equation

$$\frac{1}{2}\epsilon^2\frac{d^2}{dy^2}f_n^i(y) - V'\frac{d}{dy}f_n^i(y) = -nf_{n-1}^i(y) \tag{1.7}$$

where V' is a shorthand for $V' = \nabla V_0$ if we are escaping from the potential described by V_0. Note that the potential is frozen in the case of V_0. Similarly V' is a shorthand for $V' = \nabla V_t$ if we are escaping from the potential described by V_t. The following boundary conditions are

$$f_n^i(0) = 0, \quad \frac{d}{dy}f_n^1(-\infty) = 0 \quad \text{and} \quad \frac{d}{dy}f_n^2(+\infty) = 0.$$

[1] See Appendix A for a full explanation of the notation used for the potentials.

Having $f_n^i(0) = 0$ is appropriate since being at $y = 0$ means it is in neither well and so has already escaped at $t = 0$ anyway. We can see how $\frac{d}{dy}f_n^1(-\infty) = 0$ and $\frac{d}{dy}f_n^2(+\infty) = 0$ make sense by considering f_1^2 as an example. If the particle starts at $x(0) = y$, where y is a very large positive number $y \gg N\sqrt{a}$ (where $N \gg 1$) then with the equilibrium point being an attractor, it would more or less deterministically slide towards $x = +\sqrt{a}$. We call the time it takes for it to travel to $x = +\sqrt{a}$, τ'. If the particle starts somewhere further beyond y say $x = y + \delta$, where $\delta > 0$, it would also slide down to $x = +\sqrt{a}$ almost deterministically. We call this new time to get to $x = +\sqrt{a}$, τ''. Intuitively, we would expect $\tau' \approx \tau''$ so $\frac{d}{dy}f_1^2(-\infty) = 0$.

The aim now is to solve Eqn 1.7 for different cases. These are for $F = 0$ and $F \neq 0$, in the small noise approximation.

1.2.1 One Dimensional Potential - Case $F = 0$

The potential is stationary and does not depend on time. It is symmetric at $x = 0$ so we must have

$$f_1^1(-y) = f_1^2(y).$$

For simplicity we denote the following

$$f = f_1^1 \quad \text{and} \quad I = \frac{df}{dy}$$

which rewrites the differential equation as

$$\frac{1}{2}\epsilon^2\frac{dI}{dy} - V_0'I = -1$$

where $V_0' = \nabla V_0$. Using an integrating factor gives

$$\frac{d}{dy}\left(I \times \exp\left\{\int_0^y -\frac{2}{\epsilon^2}V_0'(s)\,ds\right\}\right) = -\frac{2}{\epsilon^2}\exp\left\{\int_0^y -\frac{2}{\epsilon^2}V_0'(s)\,ds\right\}$$

which gives

$$\frac{d}{dy}\left(I \times \exp\left\{-\frac{2}{\epsilon^2}V_0(y)\right\}\right) = -\frac{2}{\epsilon^2}\exp\left\{-\frac{2}{\epsilon^2}V_0(y)\right\}.$$

We know that $I(-\infty) = 0$ so integrating we have

$$I(y) \times \exp\left\{-\frac{2}{\epsilon^2}V_0(y)\right\} - I(-\infty) \times \exp\left\{-\frac{2}{\epsilon^2}V_0(-\infty)\right\} = -\frac{2}{\epsilon^2}\int_{-\infty}^y \exp\left\{-\frac{2}{\epsilon^2}V_0(s)\right\} ds$$

and proceeding we have

$$I(y) \times \exp\left\{-\frac{2}{\epsilon^2}V_0(y)\right\} = -\frac{2}{\epsilon^2}\int_{-\infty}^y \exp\left\{-\frac{2}{\epsilon^2}V_0(s)\right\} ds$$

$$I(y) = -\frac{2}{\epsilon^2}\exp\left\{+\frac{2}{\epsilon^2}V_0(y)\right\}\int_{-\infty}^{y}\exp\left\{-\frac{2}{\epsilon^2}V_0(s)\right\}\,ds.$$

We know that $f(0) = 0$ so integrating we have

$$f(y) - f(0) = -\frac{2}{\epsilon^2}\int_0^y \exp\left\{+\frac{2}{\epsilon^2}V_0(u)\right\}\int_{-\infty}^{u}\exp\left\{-\frac{2}{\epsilon^2}V_0(s)\right\}\,ds\,du$$

which we rewrite as

$$f(y) = -\frac{2}{\epsilon^2}\int_0^y \exp\left\{+\frac{2}{\epsilon^2}V_0(u)\right\}g(u)\,du \tag{1.8}$$

$$\text{where}\quad g(u) = \int_{-\infty}^{u}\exp\left\{-\frac{2}{\epsilon^2}V_0(s)\right\}\,ds.$$

Up to now the methods we have used for solving $f(\cdot)$ are exact and the boundary conditions on $f(\cdot)$ have also been kept. There are no approximations to our approach so far. Recall that $y < 0$. We seek an approximate solution for $f(y)$ in the region $y \in [-\sqrt{a}, 0]$. Now we use Laplace Method in the small noise limit (small ϵ) to evaluate the integrals in Equation 1.8. Note that

$$\max_{u\in[-\infty,0]}\left(-\frac{2}{\epsilon^2}V_0(u)\right) = -\frac{2}{\epsilon^2}V_0\left(-\sqrt{a}\right).$$

Using the Laplace Method for small ϵ gives the approximation

$$g(u) \approx \begin{cases} \sqrt{\frac{\pi\epsilon^2}{2a}}\exp\left\{\frac{a^2}{2\epsilon^2}\right\} & \text{if} \quad u \in (-\sqrt{a}, 0] \\ \frac{1}{2}\sqrt{\frac{\pi\epsilon^2}{2a}}\exp\left\{\frac{a^2}{2\epsilon^2}\right\} & \text{if} \quad u = -\sqrt{a} \end{cases}$$

This means $g(u) \approx \sqrt{\frac{\pi\epsilon^2}{2a}}\exp\left\{\frac{a^2}{2\epsilon^2}\right\}$ almost everywhere with respect to the Lebesgue measure on $[-\sqrt{a}, 0]$. This approximates $f(\cdot)$ to

$$f(y) \approx +\frac{2}{\epsilon^2}\sqrt{\frac{\pi\epsilon^2}{2a}}\exp\left\{\frac{a^2}{2\epsilon^2}\right\}\int_y^0 \exp\left\{+\frac{2}{\epsilon^2}V_0(s)\right\}\,ds$$

where we have switched the limits of the integral. We use Laplace Method again after noting that

$$\max_{s\in[-\sqrt{a},0]}\left(+\frac{2}{\epsilon^2}V_0(s)\right) = +\frac{2}{\epsilon^2}V_0(0)$$

here the maximum is on the edge on the boundary meaning we would need an extra factor of $\frac{1}{2}$. So

$$\begin{aligned} f(y) &\approx +\frac{2}{\epsilon^2}\sqrt{\frac{\pi\epsilon^2}{2a}}\exp\left\{\frac{a^2}{2\epsilon^2}\right\}\frac{1}{2}\sqrt{\frac{\pi\epsilon^2}{a}} \\ &= \frac{1}{\sqrt{2}}\frac{\pi}{a}\exp\left\{\frac{a^2}{2\epsilon^2}\right\}. \end{aligned}$$

1.2.2 One Dimensional Potential - Case $F \neq 0$

The potential is now oscillating. We aim to do a similar calculation to the static potential case. We make an approximation by assuming that the amplitude of the oscillations F is small enough such that there will always be two distinct wells. The positions of the critical points and wells will be very close to the static potential case. We can just focus on one well $x_0(t)$ on the left of the hill which is dependent on the time t. The method is similar to what we have used in the $F = 0$ case. We have

$$f(y) = -\frac{2}{\epsilon^2}\int_0^y \exp\left\{+\frac{2}{\epsilon^2}V_t(u)\right\}\int_{-\infty}^u \exp\left\{-\frac{2}{\epsilon^2}V_t(s)\right\}\,ds\,du \tag{1.9}$$

$$= +\frac{2}{\epsilon^2}\int_y^0 \exp\left\{+\frac{2}{\epsilon^2}V_t(u)\right\} g_t(u)\,du$$

$$\text{where} \quad g_t(u) = \int_{-\infty}^u \exp\left\{-\frac{2}{\epsilon^2}V_t(s)\right\}\,ds$$

since $y < 0$ the limits of the integral may be switched. Notice we have approximated the situation by assuming that the hill moves very little away from $x = 0$ which is what makes Equation 1.9 valid. We seek a solution for $f(y)$ in the region $y \in [x_0(t), 0]$. For small ϵ, we can use Laplace's Method to approximate the integrals in Equation 1.9. Note that

$$\max_{u\in[-\infty,0]}\left(-\frac{2}{\epsilon^2}V_t(u)\right) = -\frac{2}{\epsilon^2}V_t\left(x_0(t)\right).$$

Now using the Laplace's Method gives

$$g_t(u) \approx \begin{cases} \exp\left\{-\frac{2}{\epsilon^2}V_t\left(x_0(t)\right)\right\}\sqrt{\frac{\pi\epsilon^2}{V_t''(x_0(t))}} & \text{if} \quad u \in (x_0(t), 0] \\ \frac{1}{2}\exp\left\{-\frac{2}{\epsilon^2}V_t\left(x_0(t)\right)\right\}\sqrt{\frac{\pi\epsilon^2}{V_t''(x_0(t))}} & \text{if} \quad u = x_0(t) \end{cases}$$

where $V_t'' = \nabla^2 V_t$. In other words $g_t \approx \exp\left\{-\frac{2}{\epsilon^2}V_t\left(x_0(t)\right)\right\}$ almost everywhere on $[x_0(t), 0]$ with respect to the Lebesgue measure. So

$$f(y) \approx +\frac{2}{\epsilon^2}\exp\left\{-\frac{2}{\epsilon^2}V_t\left(x_0(t)\right)\right\}\sqrt{\frac{\pi\epsilon^2}{V_t''\left(x_0(t)\right)}}\int_y^0 \exp\left\{+\frac{2}{\epsilon^2}V_t(u)\right\}\,du.$$

Now

$$\max_{u\in[x_0(t),0]}\left(+\frac{2}{\epsilon^2}V_t(u)\right) = +\frac{2}{\epsilon^2}V_t(0)$$

where the maximum is on the boundary of $[x_0(t), 0]$ meaning we would need an extra factor of $\frac{1}{2}$. So

$$f(y) \approx +\frac{2}{\epsilon^2}\exp\left\{-\frac{2}{\epsilon^2}V_t\left(x_0(t)\right)\right\}\sqrt{\frac{\pi\epsilon^2}{V_t''\left(x_0(t)\right)}}\frac{1}{2}\sqrt{\frac{\pi\epsilon^2}{-V_t''(0)}}$$

which gives

$$f(y) \approx \frac{\pi}{\sqrt{aV_t''\left(x_0(t)\right)}} \exp\left\{-\frac{2}{\epsilon^2} V_t\left(x_0(t)\right)\right\}.$$

Finding $f(y)$ is very hard so we consider when the oscillations are very slow, that is for very small Ω. At the two extremes we have

$$dx = [x(a - x^2) + F]dt \, + \, \epsilon \, dw \quad \text{when } t = 0 \tag{1.10}$$

$$dx = [x(a - x^2) - F]dt \, + \, \epsilon \, dw \quad \text{when } t = \frac{\pi}{\Omega}. \tag{1.11}$$

We also assume that the oscillations are so small, the now time dependent equilibrium point does not differ much from the time independent case $x_0 = -\sqrt{a}$. We solve $f(y)$ for the case of Equation 1.10. The case for Equation 1.11 is similar. Let $x_0(t)$ be approximated and denoted with a new notation by

$$x_0(t) = z_0 + \delta = s$$

where $z_0 = -\sqrt{a}$. We seek an expression for δ by

$$\begin{aligned} [s(a - s^2) + F]dt = 0 \quad &\text{from Equation 1.10} \\ s(a - s^2) &= -F \\ \delta &\approx \frac{-F}{a - 3z_0^2} = +\frac{F}{2a} \end{aligned}$$

after ignoring terms of higher order than δ^2. Progressing further gives

$$\begin{aligned} -\frac{2}{\epsilon^2} V_{t=0}(s) &= -\frac{2}{\epsilon^2}\left\{\frac{s^4}{4} - a\frac{s^2}{2} - Fs\right\} \\ &\approx -\frac{2}{\epsilon^2}\left[V_0(z_0) + \delta(z_0^3 - az_0 - F) - Fz_0\right] \end{aligned} \tag{1.12}$$

again after ignoring terms of higher order than δ^2. Equation 1.12 is now approximated by

$$\begin{aligned} -\frac{2}{\epsilon^2} V_{t=0}(s) &\approx -\frac{2}{\epsilon^2}\left[V_0(z_0) + \delta(z_0^3 - az_0 - F) - Fz_0\right] \\ &= \frac{a^2}{2\epsilon^2}\left\{1 + \frac{4F^2}{2a^3} - \frac{4F}{a^{\frac{3}{2}}}\right\} \\ &\approx \frac{a^2}{2\epsilon^2}\left\{1 - \frac{4F}{a^{\frac{3}{2}}}\right\} \end{aligned}$$

after assuming F^2 is small. We make another approximation by

$$\sqrt{aV_{t=0}''\left(x_0(t)\right)} \approx \sqrt{aV_0''(z_0)}$$

$$= \sqrt{a[3z_0^2 - a]}$$
$$= a\sqrt{2}$$

after noting that $z_0 = -\sqrt{a}$. So $f(y)$ for Equation 1.10 and 1.11 are

$$f(x_0(t)) = \frac{\pi}{a\sqrt{2}} \exp\left\{\frac{a^2}{2\epsilon^2}\left(1 - \frac{4F}{a^{\frac{3}{2}}}\right)\right\} \quad \text{when } t = 0$$

$$f(x_0(t)) = \frac{\pi}{a\sqrt{2}} \exp\left\{\frac{a^2}{2\epsilon^2}\left(1 + \frac{4F}{a^{\frac{3}{2}}}\right)\right\} \quad \text{when } t = \frac{\pi}{\Omega}.$$

We can see how the solution make physical sense because when $t = \frac{\pi}{\Omega}$ the left well is lower, and so the probability to escape is lower and the time to escape would also be longer. Now that all of our calculations are done for both the time independent and time dependent case we can compare them.

1.2.3 Conclusion and Resonance Condition ϵ_{res}

We have effectively reviewed f_1^1 in the limit of small noise, which is $\langle\tau\rangle$ the averaged escape time for small ϵ. Comparing them more clearly here gives

$$F = 0 \quad \langle\tau\rangle = \frac{1}{\sqrt{2}}\frac{\pi}{a}\exp\left\{\frac{a^2}{2\epsilon^2}\right\}$$

$$F \neq 0 \quad \langle\tau\rangle = \frac{\pi}{a\sqrt{2}} \exp\left\{\frac{a^2}{2\epsilon^2}\left(1 - \frac{4F}{a^{\frac{3}{2}}}\right)\right\} \quad \text{when} \quad t = 0$$

$$\langle\tau\rangle = \frac{\pi}{a\sqrt{2}} \exp\left\{\frac{a^2}{2\epsilon^2}\left(1 + \frac{4F}{a^{\frac{3}{2}}}\right)\right\} \quad \text{when} \quad t = \frac{\pi}{\Omega}.$$

For the $F \neq 0$ case if we impose

$$\langle\tau\rangle = \frac{\pi}{\Omega} \quad \text{for} \quad t = 0 \tag{1.13}$$

$$\langle\tau\rangle = \frac{\pi}{\Omega} \quad \text{for} \quad t = \frac{\pi}{\Omega} \tag{1.14}$$

and solve for the noise in both cases (that is solving Equation 1.13 and 1.14 for ϵ) we have

$$\epsilon_1 = a\left(\frac{1 - 4F/a^{3/2}}{2\ln(2\sqrt{2}a/\Omega)}\right)^{1/2} \quad \text{for} \quad t = 0$$

$$\epsilon_2 = a\left(\frac{1 + 4F/a^{3/2}}{2\ln(2\sqrt{2}a/\Omega)}\right)^{1/2} \quad \text{for} \quad t = \frac{\pi}{\Omega}$$

then the resonance condition ϵ_{res} should be inside the interval $\epsilon_{res} \in [\epsilon_1, \epsilon_2]$. For the example below it just so happen that $[\epsilon_1, \epsilon_2] = [0.18, 0.31]$ and $\epsilon_{res} \approx 0.26$, which gives the trajectory

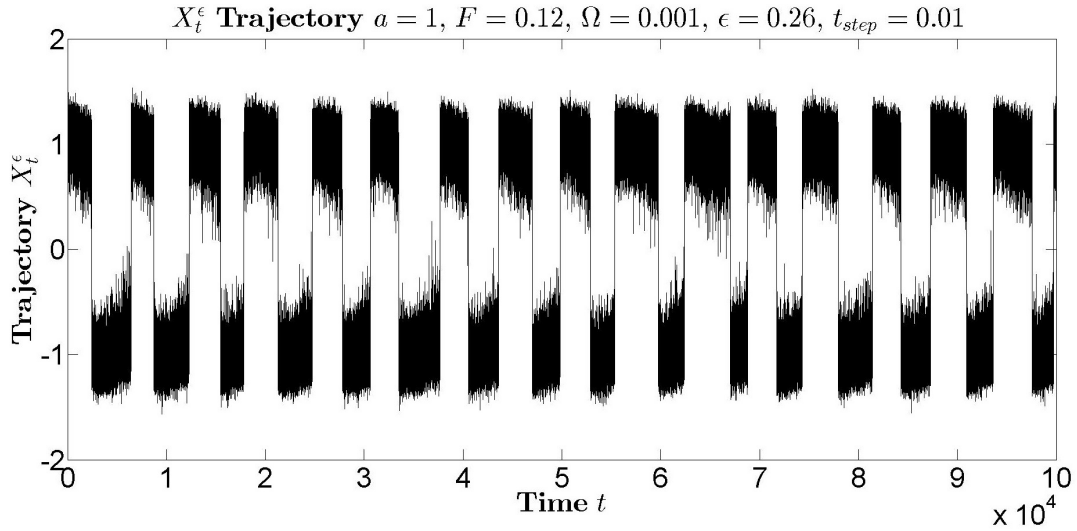

Figure 1.1: This trajectory is exhibiting quasi-determinism. It is near stochastic resonance.

We can increase and decrease the noise away from $\epsilon_{res} \approx 0.26$, and transitions will occur more frequently or less frequently as we move away from resonance.

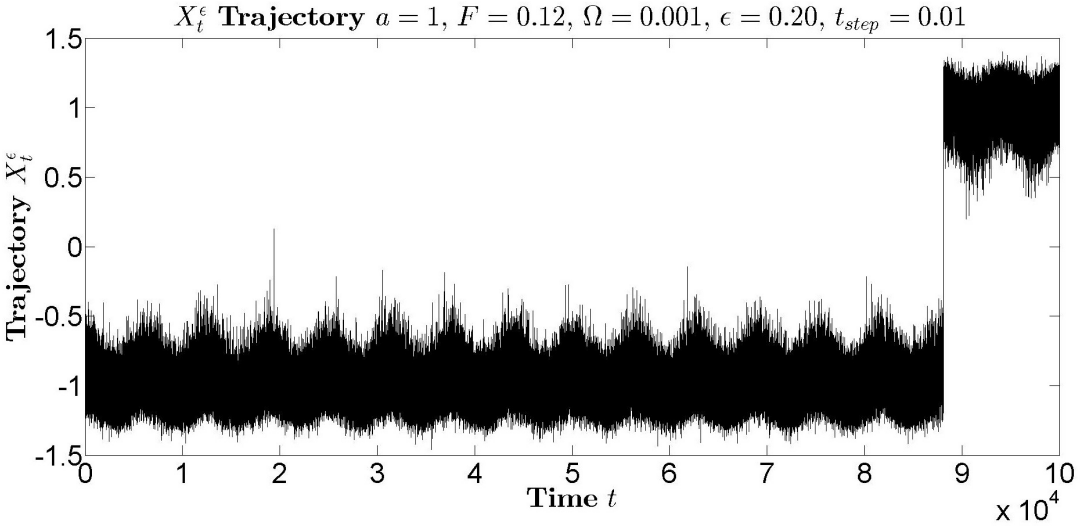

Figure 1.2: Transitions occur irregularly and are rare.

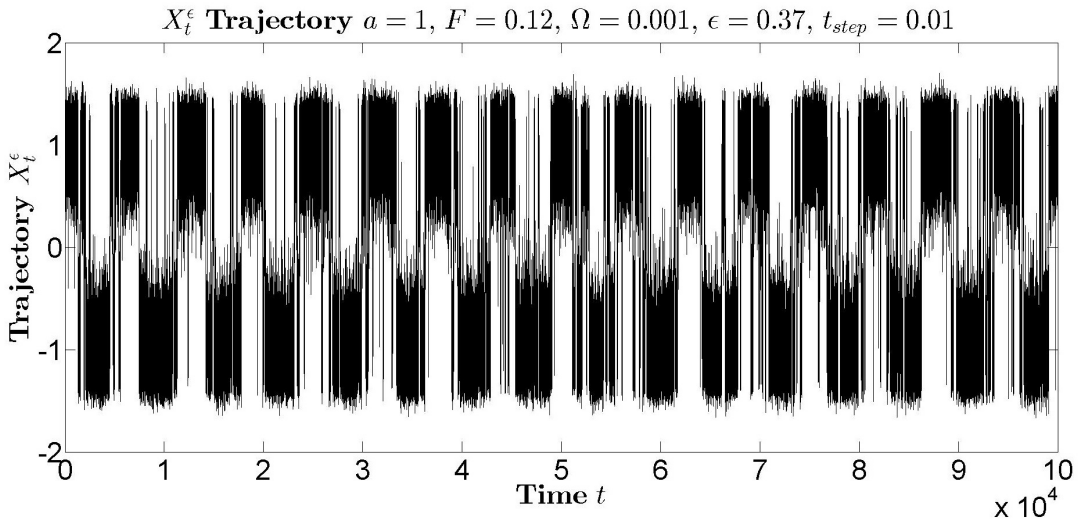

Figure 1.3: Transitions occur very often.

1.3 Remarks on One Dimensional Potential

Notice that this is a very crude way to study the system. The oscillating potential is being approximated by a frozen static potential. This is precisely the adiabatic approximation. Not only does it assume small noise, small forcing and adiabatic time development, the exact positions of the critical points (the two wells and the hill) were not calculated. All the calculations assumed that the hill was near $x = 0$. This means the forcing is assumed to be small enough such that the hill does not move far away from $x = 0$. Benzi et al's definition of the escape time is so crude it will be used only as a rough guide.

Also note it is hard to tell if Figures 1.1, 1.2 and 1.3 show any regularity or not. As we shall see, regularity shows itself in the distribution of the escape times.

Chapter 2

Theoretical Escape Time from a Well of a Static Potential

We consider the theoretical escape time of a particle from a well of a static potential. This is done in two parts. The first part is Freidlin-Wentzell theory or large deviation and the second part is Kramers' formula which is derived using potential theory. For small noise levels large deviation is considered and for higher noise levels potential theory is used.

2.1 Freidlin-Wentzell Theory and Large Deviation

A review of the major results of the Freidlin-Wentzell theory is presented which is found in [32]. This is done by considering stochastic systems converging to the deterministic limit for small noise, action functional for Wiener processes, action functional for general processes and the main theorems concerning the escape time.

2.1.1 Stochastic Processes

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(\mathbb{R}^r, \mathcal{B})$ be a measure space on $\mathbb{R}^r$. Let T be an indexing set. For $\omega \in \Omega$ and $t \in T$ define a mapping $\Omega \times T \longrightarrow \mathbb{R}^r$ by

$$X_t(\omega) : \Omega \times T \longrightarrow \mathbb{R}^r$$

where $X_t(\omega) \in \mathbb{R}^r$ is called a stochastic process on $\mathbb{R}^r$. The probability measure defined on $A \in \mathcal{F}$ is denoted by $P(A)$. But if this probability measure can depend on a value $x \in \mathbb{R}^r$, we will often put this dependence explicitly into the notation

$$P(A, x) = P_x(A).$$

We will only consider Markov processes in this thesis. There are further technical properties a Markov process has to fulfil.[2] Intuitively this can be understood in the following way;

[2]For more details see page 20 in [32].

the Ω can be thought of as the set of all trajectories of a stochastic process; the T can be thought of as the set of time, for example $T = [0, \infty)$; the $\mathbb{R}^r$ can be thought as the space in which the trajectory is in, then $X_t(\omega)$ is a trajectory in $\mathbb{R}^r$ with continuous time $t \geq 0$.

2.1.2 Deterministic Limit

Consider the following system. We have the r-dimensional real space $\mathbb{R}^r$. Let $x_t \in \mathbb{R}^r$ be a time dependent variable in $\mathbb{R}^r$. Let $b(x_t)$ be a function $b : \mathbb{R}^r \to \mathbb{R}^r$ on x_t. We then let

$$dx_t = b(x_t)dt \tag{2.1}$$

which can be seen as a system of r differential equations for each of the elements of x_t. When we consider random systems we denote the (random) variable by X_t^ϵ with values in $\mathbb{R}^r$. The stochastic processes we consider in this thesis are diffusion processes. More precisely we consider random dynamical systems which are solutions of the following system of stochastic differential equations

$$dX_t^\epsilon = b(X_t^\epsilon)dt + \epsilon\sigma(X_t^\epsilon)dw_t \tag{2.2}$$

where ϵ is the noise level, w_t is a l-dimensional Wiener process and $\sigma(X_t^\epsilon)$ is a function on X_t^ϵ returning a $l \times r$ matrix.

The first circle of results in the book of Freidlin and Wentzell are about how the solutions of the random system Equation 2.2 approximate the solutions of the deterministic system Equation 2.1. For example we know that

$$\lim_{\epsilon \to 0} X_t^\epsilon = x_t.$$

But the exact manner of this limit and the conditions under which $X_t^\epsilon \to x_t$ is reached is documented in Freidlin-Wentzell.[3] For example we have[4]

Theorem 2.1. *Suppose that the coefficients of Equation 2.2 satisfy a Lipschitz condition and a growth condition given by*

$$\sum_i \left[b_i(x) - b_i(y)\right]^2 + \sum_{i,j} \left[\sigma_{ij}(x) - \sigma_{ij}(y)\right]^2 \leq K^2|x - y|^2 \tag{2.3}$$

$$\sum_i \left[b_i(x)\right]^2 + \sum_{i,j} \left[\sigma_{ij}(x)\right]^2 \leq K^2(1 + |x|^2) \tag{2.4}$$

then for all $t > 0$ and $\delta > 0$ we have

$$E\left|X_t^\epsilon - x_t\right|^2 \leq \epsilon^2 a(t) \quad \textit{and} \quad \lim_{\epsilon \to 0} P\left\{\max_{0 \leq s \leq t} |X_s^\epsilon - x_s| > \delta\right\} = 0$$

where $a(t)$ is a monotone increasing function, which is expressed in terms of $|x|$ and K.

[3] See pages 44-59 in [32].

[4] Adapted from Theorem 1.2 page 45 of [32].

Theorem 2.1 can be explained in another way. Intuitively as $\epsilon \to 0$ we would expect to be back in the deterministic system x_t. Note that Equation 2.3 is the Lipschitz condition and 2.4 is the growth condition.

There is also a stochastic analogue of Taylor's Remainder's Theorem where it can be shown that X_t^ϵ admits the following decomposition[5]

$$X_t^\epsilon = X_t^{(0)} + \epsilon X_t^{(1)} + \cdots + \epsilon^k X_t^{(k)} + R_{k+1}^\epsilon(t)$$

where the remainder is bounded by new functions

$$\sup_{0\le t\le T} \left|R_{k+1}^\epsilon(t)\right| < C(\omega)\epsilon^{k+1} \quad \text{and} \quad P\left\{C(\omega) < \infty\right\} = 1$$

and the $X_t^{(k)}$ are solutions of stochastic differential equations.

2.1.3 Action Functional for Wiener processes

The second part of Freidlin-Wentzell has a new setting.[6] Let $b(X_t^\epsilon) = 0$, $\sigma(X_t^\epsilon) = 1$ and w_t be a r-dimensional Wiener process, that is to say

$$dX_t^\epsilon = \epsilon\, dw_t$$

where we have reduced the system to a r-dimensional Wiener process. Let $C_{T_1T_2} = C_{T_1T_2}(\mathbb{R}^r)$ denote the set of all continuous paths in $\mathbb{R}^r$ starting at time T_1 and ending at T_2. On this set we define a metric by

$$\rho_{T_1T_2}(\psi,\varphi) = \sup_{T_1\le t\le T_2} |\psi_t - \varphi_t|.$$

We define a new functional by

$$S(\varphi) = S_{T_1T_2}(\varphi) = \frac{1}{2}\int_{T_1}^{T_2} |\dot\varphi_s|^2 ds$$

for absolutely continuous (and differentiable) φ_t. If φ_t is not absolutely continuous or if the integral is divergent, we set $S(\varphi) = +\infty$. We define the action functional by

$$I_{T_1T_2}^\epsilon(\varphi) = \epsilon^{-2} S_{T_1T_2}(\varphi)$$

and $S_{T_1T_2}(\varphi)$ will be called the normalized action functional. The paths should be interpreted as points, that is elements of the functional space of paths, that is each point is itself a path. The distance between these points, and hence paths, is given by the metric just defined. We define a new set

$$\Phi(s) = \{\varphi \in C_{0T} \text{ such that } \varphi_0 = 0 \text{ and } S_{0T}(\varphi) \le s\}$$

[5]See Theorem 2.1 page 52 of [32].

[6]See pages 70-79 in [32].

which can be shown to be compact in the uniform topology.[7] Now the SDE

$$dX_t^\epsilon = \epsilon\, dw_t$$

cannot be solved pathwise as in the deterministic case. The solution is a randomly chosen path out of an infinitude of possible paths. This is described by the probability of the path having certain properties. Also X_t^ϵ is self-similar and non-differentiable, but may be approximated by differentiable functions φ_t. The next major theorems in Freidlin-Wentzell show that for any $\delta > 0$ and $\gamma > 0$ we have[8]

$$P\left\{\rho_{0T}(X_t^\epsilon, \varphi_t) < \delta\right\} \geq \exp\left\{-\epsilon^{-2}\left[S_{0T}(\varphi) + \gamma\right]\right\}$$

and for any $\delta > 0$, $\gamma > 0$, $s_0 > 0$ with $s < s_0$ we have[9]

$$P\left\{\rho_{0T}(X_t^\epsilon, \Phi(s)) \geq \delta\right\} \leq \exp\left\{-\epsilon^{-2}(s - \gamma)\right\}.$$

These two statements may be interpreted as a Laplace type theorem in function spaces. A physical interpretation is that this gives an asymptotic description (in small ϵ) for the probability that the path X_t^ϵ is near to φ_t, that is

$$P\left\{\rho(X_t^\epsilon, \varphi) < \delta\right\} \approx \exp\left\{-\epsilon^{-2} S(\varphi)\right\}.$$

In the next section we develop the action functional for more general processes.

2.1.4 Action Functional for General processes

So far the action theory was developed for just one particular example of a stochastic process, that is the Wiener process. Now we develop an action theory for a general stochastic process described by Equation 2.2.[10]

Before we do that we state a list of properties a functional should have so that we can consider a suitable action functional. We state these properties in a more general context. Let (X, ρ) be a metric space with metric ρ. On the σ-algebra of its Borel subsets let μ^h be a family of probability measures depending on a parameter $h > 0$. Let $\lambda(h)$ be a positive function going to $+\infty$ as $h \downarrow 0$. Let $S(x)$ be a function such that $S : X \to [0, \infty]$. We say that $\lambda(h)S(x)$ is an action function if the following holds.

(0) the set $\Phi(s) = \{x : S(x) \leq s\}$ is compact for every $s \geq 0$.
(I) for any $\delta > 0$, any $\gamma > 0$ and any $x \in X$ there exists an $h_0 > 0$ such that

$$\mu^h\{y : \rho(x, y) < \delta\} \geq \exp\{-\lambda(h)[S(x) + \gamma]\}$$

[7] See Lemma 2.1 page 77 of [32].
[8] See Theorem 2.1 page 74 in [32].
[9] See Theorem 2.2 page 74 in [32].
[10] See pages 79-92 of [32].

for all $h \leq h_0$

(II) for any $\delta > 0$, any $\gamma > 0$ and any $s > 0$ there exists an $h_0 > 0$ such that

$$\mu^h\{y : \rho(y, \Phi(s)) \geq \delta\} \leq \exp\{-\lambda(h)(s - \gamma)\}$$

for all $h \leq h_0$.

$S(x)$ and $\lambda(h)$ will be called the normalized action functional and normalizing coefficient.

The results given in Chapter 2.1.3 show that the functional considered there has all the above properties, where $X = C_{T_1T_2}(\mathbb{R}^r)$, $S(\varphi) = S_{T_1T_2}(\varphi)$, $\lambda(h) = \epsilon^{-2}$ and $\epsilon = h$. Thus we can see how $X = C_{T_1T_2}(\mathbb{R}^r)$ with $I^{\epsilon}_{T_1T_2}$ is an action functional since it satisfied all three properties. Doubtless that there will be many other systems which satisfy all these three properties as well. This higher level of abstraction would allow us to prove powerful theorems.

2.1.5 Main Theorems

The action functional was given for a diffusion with drift term zero i.e. $b(X^{\epsilon}_t) = 0$. We now put this term back in to consider the equation

$$dX^{\epsilon}_t = b(X^{\epsilon}_t)dt + \epsilon\, dw_t$$

where w_t is a r-dimensional Wiener process. It can be shown that letting[11]

$$S_{0T}(\varphi) = \frac{1}{2}\int_0^T |\dot{\varphi}_s - b(\varphi_s)|^2 ds$$

with $\lambda(h) = \epsilon^{-2}$ and $h = \epsilon$ satisfy the three properties of the action functional. The action functional then allows us to compute asymptotically different probabilities. For example, let $D \subset \mathbb{R}^r$ be a region of space in $\mathbb{R}^r$ and let

$$H_D(t, x) = \{\varphi \in C_{0T}(\mathbb{R}^r) : \varphi_0 = x,\ \varphi_t \in D \cup \partial D\}$$
$$\overline{H}_D(t, x) = \{\varphi \in C_{0T}(\mathbb{R}^r) : \varphi_0 = x,\ x_s \notin D \text{ for some } \ s \in [0, t]\}$$

then it can be shown that[12]

$$\lim_{\epsilon\to 0} \epsilon^2 \ln P_x\{X^{\epsilon}_t \in D\} = -\min_{\varphi\in H_D(t,x)} S_{0T}(\varphi) \tag{2.5}$$

$$\lim_{\epsilon\to 0} \epsilon^2 \ln P_x\{\tau^{\epsilon} \leq t\} = -\min_{\varphi\in \overline{H}_D(t,x)} S_{0T}(\varphi) \tag{2.6}$$

where $\tau^{\epsilon} = \min\{t : X^{\epsilon}_t \notin D\}$ is the escape time from D. This theorem gives us the leading term for the probabilities leaving this region of space D. The H_D is the set of all paths

[11] See Theorem 1.1 page 104 of [32].

[12] See Theorem 1.2 page 105 of [32].

that stay in D and its boundary. The $\overline{H}_D$ is the set of all paths that leave D at some time. Equation 2.5 is thus the probability of remaining in D and Equation 2.6 is the probability of the escape time being less than t.

So far the above results hold for a general region D. Now we want to consider the case where D is the vicinity of a well, that is the region near and around a metastable state. This means D is attracted to a point inside of D. Without loss of generality we can choose this point, which is the position of the well, to be zero $x_{well} = 0$. But the minimiser of $S_{0T}(\varphi)$ is very difficult to compute explicitly using the usual differential equations. Let

$$f(t, x, y) = \min_{\substack{\varphi_0 = x \\ \varphi_t = y}} S_{0t}(\varphi).$$

The Hamilton-Jacobi equations are given by

$$\frac{\partial f(t,x,y)}{\partial t} + \frac{1}{2} \left| \nabla_y f(t,x,y) \right|^2 + (b(y), \nabla_y f(t,x,y)) = 0 \tag{2.7}$$

where ∇_y is the gradient operator in the variable y and note that

$$\min_{\varphi \in H_D(t,x)} S_{0t}(\varphi) = \min_{y \in D \cup \partial D} f(t,x,y)$$

$$\min_{\varphi \in \overline{H}_D(t,x)} S_{0t}(\varphi) = \min_{\substack{0 \le s \le t \\ y \notin D}} f(s,x,y)$$

and the solution to Equation 2.7 would be closely related to Equation 2.5 and 2.6.[13] Let us introduce the so-called quasipotential

$$\tilde{V}(x,y) = \inf\{S_{T_1 T_2}(\varphi) : \varphi \in C_{T_1 T_2}(\mathbb{R}^r),\ \varphi_{T_1} = x,\ \varphi_{T_2} = y\}$$

which is the least action over all paths which starts at x and ends at y. Suppose that the drift term can be written as the gradient of a potential V

$$b(x) = -\nabla V(x)$$

then it can be shown that[14]

$$\tilde{V}(0,x) = 2V(x). \tag{2.8}$$

Note that this only holds for points $x \in \mathbb{R}^r$ such that $V(x) \le \min_{y \in \partial D} V(y)$, that is for points lower than the exit point. Now suppose that there exists a unique point $y_0 \in \partial D$ for which $\tilde{V}(0,y_0) = \min_{y \in \partial D} \tilde{V}(0,y)$ then[15]

$$\lim_{\epsilon \to 0} P_x\{\rho(X_s^\epsilon, y_0) < \delta\} = 1 \quad \text{where} \quad s = \inf\{t : X_t^\epsilon \in \partial D\} \tag{2.9}$$

for every $\delta > 0$ and any $x \in D$. This means X_t^ϵ will exit near points of least height in the small noise limit.

[13] See pages 105-108 of [32].
[14] See Theorem 3.1 page 118 of [32].
[15] See Theorem 2.1 page 108 of [32].

2.1.6 Remarks on Freidlin-Wentzell and Large Deviation Theory

The main Theorems of Freidlin-Wentzell were developed on a precise and rigorous mathematical setting. It would be appropriate to interpret what they mean in a more physical setting. Consider Equation 2.8 and 2.9. Equation 2.8 gives an easier way to calculate the quasipotential, because the quasipotential is related in a very simple way to the height of the potential. Equation 2.9 means that the particle will escape whilst travelling through a path which gives the least height, or interpreted in another way, a path of least action. Thus, one of the main conclusion of the Fredlin-Wentzell theory is that the particle will tend to escape following close to a path which gives the least distance to climb out of a well.

2.2 Kramers' Formula and Potential Theory

Let $V : \mathbb{R}^r \longrightarrow \mathbb{R}$. Let $x \in \mathbb{R}^r$ be a well and $z_i \in \mathbb{R}^r$ be saddles labelled by $i = 1 \dots n$. The saddles would be gateways providing a passage for escape from the well. Define

$$\Delta V_i = V(z_i) - V(x)$$

which is the height difference between the well and the ith saddle. For small noise ϵ, an approximate expression can be estimated for the escape time of the particle going through the ith saddle. In the smallest order of the noise ϵ the mean exit time, as described in the previous section, is given by[16]

$$\tau_i = e^{+2\Delta V_i/\epsilon^2}.$$

Inverting this gives the escape rate

$$R_i = e^{-2\Delta V_i/\epsilon^2}$$

and the total escape rate would be to sum over all the saddles

$$R = \sum_{i=1}^{n} R_i = \sum_{i=1}^{n} e^{-2\Delta V_i/\epsilon^2}.$$

The order correction is done by adding a coefficient called Kramers' coefficient and the resulting corrected rate is called Kramers' rate

$$R_i = k_i e^{-2\Delta V_i/\epsilon^2}$$

where

$$k_i = \frac{\sqrt{|\nabla^2 V(x)|}}{2\pi} \frac{|\lambda(z_i)|}{\sqrt{\|\nabla^2 V(z_i)\|}}$$

[16] See Theorem 4.1 and 4.2 on pages 124-127 of [32].

where $|\nabla^2(x)|$ denotes the determinant of the Hessian of the potential at the well x, $\|\nabla^2 V(z_i)\|$ denotes the modulus of the determinant of the potential at the saddle z_i and $|\lambda(z_i)|$ denotes the minimum eigenvalue of the Hessian of the potential at the saddle z_i. This gives the escape rate in the next order of approximation to be

$$R = \sum_{i=1}^{n} R_i = \sum_{i=1}^{n} k_i e^{-2\Delta V_i/\epsilon^2}$$

which is rewritten as

$$R = \frac{\sqrt{|\nabla^2 V(x)|}}{2\pi} \sum_{i=1}^{n} \frac{|\lambda(z_i)|}{\sqrt{\|\nabla^2 V(z_i)\|}} \exp\left\{\frac{-2\left(V(z_i) - V(x)\right)}{\epsilon^2}\right\}.$$

The last order of approximation for higher noise ϵ is done by bounding the error on Kramers' coefficient. This is

$$k_i = \frac{\sqrt{|\nabla^2 V(x)|}}{2\pi} \frac{|\lambda(z_i)|}{\sqrt{\|\nabla^2 V(z_i)\|}} \left(\frac{1}{1 + \mathcal{O}\left(\frac{\epsilon^2}{2} \ln \frac{\epsilon^2}{2}\right)}\right)$$

which is rewritten as

$$\frac{1}{k_i} = \frac{2\pi}{\sqrt{|\nabla^2 V(x)|}} \frac{\sqrt{\|\nabla^2 V(z_i)\|}}{|\lambda(z_i)|} \left[1 + \mathcal{O}\left(\frac{\epsilon^2}{2} \ln \frac{\epsilon^2}{2}\right)\right]. \tag{2.10}$$

We conclude with a few words on the derivation of Kramers' formula and the bound on its error. Equation 2.10 was derived rigorously using techniques from potential theory instead of large deviations. Part of the technique involves the escape time being expressed in terms of a partial differential equation, similar to Equation 1.7 for example. This derivation was done in [37] which is beyond the scope of this thesis. This Chapter reviewed the escape rates of a particle from a static well, which will be relevant when we consider escape rates from an oscillating well.

Chapter 3

Theoretical Escape Time from a Well of an Oscillatory Potential

Stochastic resonance usually involves studying transitions between two stable states. Studying a stochastic differential equation in multidimensional real space can be complicated. It would be useful to simplify stochastic resonance down to a Markov Chain with transitions between two states $+1$ and -1, then we try to model stochastic resonance with a two state Markov Chain. This is done for both discrete and continuous time Markov Chains with two states, for both alternating and synchronised saddles.

3.1 Markov Chain Reduction

Let $V : \mathbb{R}^r \longrightarrow \mathbb{R}$ be a potential with two wells. This potential is subjected to a periodic forcing F with frequency Ω and perturbed by noise ϵ, which is described by the SDE[17]

$$\dot{X}_t^\epsilon = -\nabla V + F\cos(\Omega t) + \epsilon \dot{W}_t \tag{3.1}$$

where W_t is a Wiener process in $\mathbb{R}^r$ and $F \in \mathbb{R}^r$. We call X_t^ϵ the diffusion case. The X_t^ϵ can be reduced to a Markov Chain Y_t^ϵ on $\{-1, +1\}$

$$X_t^\epsilon \longrightarrow Y_t^\epsilon$$

in the following way. In what follows we will assume that the diffusion X_t^ϵ is continuous in time and space. Let $w_l(t)$ denote the position of the left well at time t and $w_r(t)$ the position of the right well at time t. Note that $w_l(t)$ and $w_r(t)$ are also continuous in time. Let $R \in \mathbb{R}$ be constant. The reduction from the X_t^ϵ to the Markov Chain Y_t^ϵ is

$$Y_t^\epsilon = \begin{cases} -1 & \text{if} \quad |X_t^\epsilon - w_l(t)| \leq R \\ +1 & \text{if} \quad |X_t^\epsilon - w_r(t)| \leq R \\ Z & \text{if} \quad |X_t^\epsilon - w_l(t)| > R \quad \text{and} \quad |X_t^\epsilon - w_r(t)| > R \end{cases}$$

[17] See Appendix A for how the forcing is denoted.

where Z is given by

$$Z = \begin{cases} -1 & \text{if} \quad s_2 < s_1 \\ +1 & \text{if} \quad s_1 < s_2 \end{cases}$$

where s_1 and s_2 are given by

$$\begin{aligned} s_1 &= \max_{u<t} \{u : |X_u^\epsilon - w_l(u)| \leq R\} \\ s_2 &= \max_{u<t} \{u : |X_u^\epsilon - w_r(u)| \leq R\} . \end{aligned}$$

When $Y_t^\epsilon = -1$ we say the particle is in the left well and when $Y_t^\epsilon = +1$ we say the particle is in the right well. Only when it enters the other well would Y_t^ϵ change sign. When the condition $|X_u^\epsilon - w_l(u)| \leq R$ is satisfied we say the particle is covered by the left well. When the condition $|X_u^\epsilon - w_r(u)| \leq R$ is satisfied we say the particle is covered by the right well. Note that R is chosen small enough such that it is impossible for the particle to be covered by both wells at any time, that is

$$\{x \in \mathbb{R}^r \, : \, |X_t^\epsilon - w_l(t)| < R \quad \text{and} \quad |X_t^\epsilon - w_r(t)| < R\} = \emptyset$$

for all times $t \geq 0$. This means that s_1 is the most recent time the particle is covered by the left well and s_2 is the most recent time the particle is covered by the right well. Notice that if initially at $t = 0$, the particle is covered by neither well then Y_t^ϵ cannot be derived nor defined by the above definitions. In this case either $Y_0^\epsilon = -1$ or $Y_0^\epsilon = +1$ is chosen depending on what initial conditions are required. In other words if $Y_0^\epsilon = -1$ is chosen as the initial condition then the particle is covered by the left well for $t < 0$. If $Y_0^\epsilon = +1$ is chosen as the initial condition then the particle is covered by the right well for $t < 0$.

The escape time from the left to right well τ_{-1+1} and from the right to left well τ_{+1-1} are defined in the following way[18]

$$\begin{aligned} \tau_{-1+1} &= \mu\left(\{t : Y_t^\epsilon = -1\}\right) \quad \text{where} \quad \{t : Y_t^\epsilon = -1\} \quad \text{is an interval} \\ \tau_{+1-1} &= \mu\left(\{t : Y_t^\epsilon = +1\}\right) \quad \text{where} \quad \{t : Y_t^\epsilon = +1\} \quad \text{is an interval} \end{aligned}$$

where μ denotes the Lebesgue measure. In other words the time spent being in the state $Y_t^\epsilon = -1$ is τ_{-1+1} and the time spent being in the state $Y_t^\epsilon = +1$ is τ_{+1-1}. These intervals will always be closed intervals. The process Y_t^ϵ has two states, hence each sample is a piecewise constant function. The length of each piece is the escape time τ_{-1+1} or τ_{+1-1}. Note that τ_{-1+1} and τ_{+1-1} are random times and random variables.

For the diffusion the escape time can be explained in the following way. Each well is surrounded by a circle with a constant radius R which moves with the well. A particle is said to have entered the left well if it enters the region covered by the radius R over the left well. The particle is then said to have entered the right well when it enters the region covered by R in the right well. The time difference between entering the left well

[18]See Appendix B.1 for details of the actual use of R in the measurement of the escape times.

and entering the right well is defined to be the escape time from left to right τ_{-1+1}. A similar argument is said for τ_{+1-1}. This also means the escape times in the Markov Chain is the same as the diffusion trajectory X_t^ϵ by definition. Notice that the diffusion X_t^ϵ has to be defined first before the Markov Chain Y_t^ϵ which is a derived quantity.

Notice that all of our reasoning in deriving Y_t^ϵ only assumes that X_t^ϵ is a continuous time process in $\mathbb{R}^r$. We did not check whether Y_t^ϵ satisfy the strict definitions of a continuous time Markov Chain. If Y_t^ϵ is a Markov Chain it should also satisfy the Markov property, that is

$$\begin{aligned} P\left(Y_t^\epsilon = i | Y_{t_1}^\epsilon = -1, Y_{t_2}^\epsilon = +1\right) &= P\left(Y_t^\epsilon = i | Y_{t_1}^\epsilon = -1\right) \\ &= P\left(Y_t^\epsilon = i | Y_{t_2}^\epsilon = +1\right) \end{aligned}$$

for any $0 \leq t_1 < t_2 < t$. Again we stress that the only assumption we made when deriving Y_t^ϵ from X_t^ϵ is that X_t^ϵ is a continuous time process in $\mathbb{R}^r$, which is not sufficient for Y_t^ϵ to be a Markov Chain nor for Y_t^ϵ to satisfy the Markov property. But throughout the rest of this thesis the diffusion X_t^ϵ will be a Markov process, which means Y_t^ϵ should be a good approximation to a Markov Chain.[19]

A discrete time and continuous time Markov Chain model for Equation 3.1 are studied in the following sections.

3.2 Discrete Time Markov Chain

Let the time be discrete. This to say time t belongs to

$$t \in \{0, 1, 2, \ldots\}.$$

The Markov Chain is a time dependent stochastic process which can take values $+1$ or -1

$$Y_t = \pm 1.$$

At time t the probability of Y_t jumping from -1 to $+1$ is denoted by $p_{-1+1}(t)$; the probability of jumping from $+1$ to -1 is denoted by $p_{+1-1}(t)$; the probability of staying in -1 is denoted by $p_{-1-1}(t)$; and the probability of staying in $+1$ is denoted by $p_{+1+1}(t)$. Notice that they have the following properties for all time t

$$\begin{aligned} p_{-1-1}(t) + p_{-1+1}(t) &= 1 \\ p_{+1+1}(t) + p_{+1-1}(t) &= 1. \end{aligned}$$

A transition matrix can be defined as

$$P_t := \begin{pmatrix} p_{-1-1}(t) & p_{-1+1}(t) \\ p_{+1-1}(t) & p_{+1+1}(t) \end{pmatrix}.$$

[19] Whether Y_t^ϵ is a Markov Chain for a Markov process X_t^ϵ, or for X_t^ϵ described by an SDE requires proof. This is an open question.

At every point in time it is possible to define a state probability, that is the probability of the trajectory being -1 or $+1$,

$$P\left(Y_t = -1\right) = \nu_-(t)$$
$$P\left(Y_t = +1\right) = \nu_+(t).$$

Notice that the state probability satisfy the following condition for all time t

$$\nu_-(t) + \nu_+(t) = 1.$$

The two $\nu_-(t)$ and $\nu_+(t)$ can be written compactly in vector notation

$$\nu(t) = \begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix}.$$

The state probability at time $t+1$ can be expressed in terms of the last time t, that is

$$\nu(t+1) = P_t^\dagger \nu(t)$$
$$\begin{pmatrix} \nu_-(t+1) \\ \nu_+(t+1) \end{pmatrix} = \begin{pmatrix} p_{-1-1}(t) & p_{+1-1}(t) \\ p_{-1+1}(t) & p_{+1+1}(t) \end{pmatrix} \begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix}$$

where $P_t^\dagger$ denote the transpose of the matrix P_t. This means if the initial value of the state probability is known at $t = 0$, then the future behaviour of the state probability can be described by computing all subsequent values of $\nu(t)$, that is

$$\nu(t) = \prod_{i=0}^{i=t-1} P_i^\dagger \nu(0).$$

The main aim for the rest of our studies of the Markov Chain is to compute the state probability for various transition matrices. When the wells of the potential are oscillating such that one well is higher than the other, we model using $p \neq q$. When the wells of the potential are oscillating such that both wells are always at the same height as each other, we model using $p = q$.

3.2.1 Discrete Time Markov Chain - Alternating Saddles $p \neq q$

We want to study a system with periodic elements. The transition matrix would change periodically in time. Let the period be

$$T = 2m$$

where m is an integer. Let the time t be written in the form

$$t = NT + n$$

where N is an integer number of periods. The transition matrix would vary periodically according to

$$\text{For} \quad n = mod(t,T) \in T_1 \quad P_t = P_1 = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

$$\text{For} \quad n = mod(t,T) \in T_2 \quad P_t = P_2 = \begin{pmatrix} 1-q & q \\ p & 1-p \end{pmatrix}$$

where

$$T_1 = \{0, 1, \ldots, m-1\}$$
$$T_2 = \{m, m+1, \ldots, 2m-1\}.$$

We interpret $Y_t = -1$ as being in the left well and $Y_t = +1$ as being in the right well. We also interpret p as the probability of escape from a shallow well and q as the probability of escape from a deep well. The transition matrix varying periodically in time can be used to model the periodic forcing being applied to the potential. The following Theorem derives the state probabilities.

Theorem 3.1. *Let the time be $t = NT + n$. Let $\lambda = 1 - p - q$. The state probability at time t is*

$$\textit{For} \quad n \in T_1 \quad \nu = \frac{1}{p+q}\begin{pmatrix} q \\ p \end{pmatrix} - \frac{p-q}{p+q} \times \frac{\lambda^n}{1+\lambda^m}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$
$$+ \frac{\nu_+(0)(p+q\lambda^m) - \nu_-(0)(q+p\lambda^m)}{1+\lambda^m}\lambda^{2mN+n}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\textit{For} \quad n \in T_2 \quad \nu = \frac{1}{p+q}\begin{pmatrix} p \\ q \end{pmatrix} + \frac{p-q}{p+q} \times \frac{\lambda^{n-m}}{1+\lambda^m}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$
$$+ \frac{\nu_+(0)(p+q\lambda^m) - \nu_-(0)(q+p\lambda^m)}{1+\lambda^m}\lambda^{2mN+n}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Proof. Notice that the eigenvectors and eigenvalues of the transpose matrix $P_1^\dagger$ are

$$\lambda_1 = 1 - p - q \qquad v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 1 \qquad v_2 = \begin{pmatrix} q \\ p \end{pmatrix}$$

which also spans the $\mathbb{R}^2$ space. For short we call $\lambda = \lambda_1$. This means an arbitrary vector can be expressed as a linear combination of the eigenvectors of $P_1^\dagger$. This is

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{p+q}\left\{(qy - px)\begin{pmatrix} -1 \\ 1 \end{pmatrix} + (x+y)\begin{pmatrix} q \\ p \end{pmatrix}\right\}.$$

Now consider m application of the $P_1^\dagger$ matrix on the arbitrary vector.

$$\begin{aligned}(P_1^\dagger)^m \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{p+q}\left\{(qy-px)\lambda^m \begin{pmatrix} -1 \\ 1 \end{pmatrix} + (x+y)\begin{pmatrix} q \\ p \end{pmatrix}\right\} \\ &= \frac{1}{p+q}\begin{pmatrix} q+p\lambda^m & q(1-\lambda^m) \\ p(1-\lambda^m) & p+q\lambda^m \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} 1-p' & p' \\ q' & 1-q' \end{pmatrix}^\dagger \begin{pmatrix} x \\ y \end{pmatrix}\end{aligned}$$

where

$$p' = \frac{p(1-\lambda^m)}{p+q} \quad \text{and} \quad q' = \frac{q(1-\lambda^m)}{p+q}$$

with a similar expression for the other transition matrix

$$(P_2^\dagger)^m \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1-q' & q' \\ p' & 1-p' \end{pmatrix}^\dagger \begin{pmatrix} x \\ y \end{pmatrix}.$$

Now denote a new matrix by

$$\begin{aligned}P_{tot} &= (P_2^\dagger)^m (P_1^\dagger)^m \\ &= \begin{pmatrix} 1-q' & p' \\ q' & 1-p' \end{pmatrix}\begin{pmatrix} 1-p' & q' \\ p' & 1-q' \end{pmatrix}\end{aligned}$$

where P_{tot} has eigenvalues and eigenvectors

$$\xi_1 = 1 \qquad e_1 = \begin{pmatrix} 1-q' \\ 1-p' \end{pmatrix}$$

$$\xi_2 = (1-p'-q')^2 \qquad e_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

and these eigenvectors span the $\mathbb{R}^2$ space

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{1+\lambda'}\left\{(x+y)\begin{pmatrix} 1-q' \\ 1-p' \end{pmatrix} + [y(1-q') - x(1-p')]\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$$

where we have denoted

$$\lambda' = 1-p'-q'.$$

Now consider N applications of the matrix P_{tot} on the initial value of the state probability

$$P_{tot}^N \nu(0) = \frac{1}{1+\lambda'}\left\{\begin{pmatrix} 1-q' \\ 1-p' \end{pmatrix} + [\nu_+(0)(1-q') - \nu_-(0)(1-p')]\,\lambda'^{2N}\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$$

and express the results in terms of the eigenvectors of $P_1^\dagger$

$$P_{tot}^N \nu(0) = \frac{1}{1+\lambda'} \times \frac{1}{p+q} \left\{ [q(1-p') - p(1-q')] \begin{pmatrix} -1 \\ 1 \end{pmatrix} + (1+\lambda') \begin{pmatrix} q \\ p \end{pmatrix} \right\}$$
$$+ \frac{[\nu_+(0)(1-q') - \nu_-(0)(1-p')]\, \lambda'^{2N}}{1+\lambda'} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

If $n \in T_1$ consider

$$(P_1^\dagger)^n P_{tot}^N \nu(0) = \frac{1}{1+\lambda'} \times \frac{1}{p+q} \left\{ [q(1-p') - p(1-q')]\, \lambda^n \begin{pmatrix} -1 \\ 1 \end{pmatrix} + (1+\lambda') \begin{pmatrix} q \\ p \end{pmatrix} \right\}$$
$$+ \frac{[\nu_+(0)(1-q') - \nu_-(0)(1-p')]\, \lambda'^{2N}}{1+\lambda'} \lambda^n \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$
$$= \frac{1}{p+q} \begin{pmatrix} q \\ p \end{pmatrix} - \frac{p-q}{p+q} \times \frac{\lambda^n}{1+\lambda^m} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$
$$+ \frac{\nu_+(0)(p+q\lambda^m) - \nu_-(0)(q+p\lambda^m)}{1+\lambda^m} \lambda^{2mN+n} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

If $n \in T_2$ consider

$$(P_2^\dagger)^{n-m} (P_1^\dagger)^m P_{tot}^N \nu(0) = (P_2^\dagger)^{n-m} \left[\frac{1}{p+q} \begin{pmatrix} q \\ p \end{pmatrix} - \frac{p-q}{p+q} \times \frac{\lambda^m}{1+\lambda^m} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right.$$
$$\left. + \frac{\nu_+(0)(p+q\lambda^m) - \nu_-(0)(q+p\lambda^m)}{1+\lambda^m} \lambda^{2mN+m} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right]$$

and we express $(q,p)^\dagger$ in terms of the eigenvectors of $P_2^\dagger$

$$(P_2^\dagger)^{n-m} (P_1^\dagger)^m P_{tot}^N \nu(0) = (P_2^\dagger)^{n-m} \left[\frac{p-q}{p+q} \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \frac{1}{p+q} \begin{pmatrix} p \\ q \end{pmatrix} - \frac{p-q}{p+q} \times \frac{\lambda^m}{1+\lambda^m} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right.$$
$$\left. + \frac{\nu_+(0)(p+q\lambda^m) - \nu_-(0)(q+p\lambda^m)}{1+\lambda^m} \lambda^{2mN+m} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right]$$
$$= \frac{1}{p+q} \begin{pmatrix} p \\ q \end{pmatrix} + \frac{p-q}{p+q} \times \frac{\lambda^{n-m}}{1+\lambda^m} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$
$$+ \frac{\nu_+(0)(p+q\lambda^m) - \nu_-(0)(q+p\lambda^m)}{1+\lambda^m} \lambda^{2mN+n} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

This completes the proof. □

3.2.2 Discrete Time Markov Chain - Synchronised Saddles $p = q$

Let the period be

$$T = 4m$$

where m is an integer. Let the time t be written in the form

$$t = NT + n$$

where N is an integer number of periods. The transition matrix would vary periodically according to

$$\text{For} \quad n = mod(t,T) \in T_1 \quad P_1 = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$$

$$\text{For} \quad n = mod(t,T) \in T_2 \quad P_2 = \begin{pmatrix} 1-q & q \\ q & 1-q \end{pmatrix}$$

$$\text{For} \quad n = mod(t,T) \in T_3 \quad P_3 = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$$

$$\text{For} \quad n = mod(t,T) \in T_4 \quad P_4 = \begin{pmatrix} 1-q & q \\ q & 1-q \end{pmatrix}$$

where

$$\begin{aligned} T_1 &= \{0, 1, \ldots, m-1\} \\ T_2 &= \{m, m+1, \ldots, 2m-1\} \\ T_3 &= \{2m.2m+1, \ldots, 3m-1\} \\ T_4 &= \{3m, 3m+1, \ldots, 4m-1\} \end{aligned}$$

where again p should be interpreted as the probability of escape from a shallow well and q from a deep well. The following Theorem derives the state probabilities.

Theorem 3.2. *Let the time be $t = NT + n$. The state probability at time t is*

$$\textit{For} \quad n \in T_1 \quad \nu = \frac{1}{2}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (1-2p)^{2mN+n}(1-2q)^{2mN}[\nu_+(0) - \nu_-(0)] \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$$

$$\textit{For} \quad n \in T_2 \quad \nu = \frac{1}{2}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (1-2p)^{2mN+m}(1-2q)^{2mN+(n-m)}[\nu_+(0) - \nu_-(0)] \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$$

$$\textit{For} \quad n \in T_3 \quad \nu = \frac{1}{2}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (1-2p)^{2mN+m+(n-2m)}(1-2q)^{2mN+m}[\nu_+(0) - \nu_-(0)] \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$$

$$\textit{For} \quad n \in T_4 \quad \nu = \frac{1}{2}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (1-2p)^{2mN+2m}(1-2q)^{2mN+m+(n-3m)}[\nu_+(0) - \nu_-(0)] \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$$

Proof. Notice that the transpose of the matrix $P_1^\dagger$ has the following eigenvalues and eigenvectors

$$\lambda_1 = 1 - 2p \quad v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 1 \qquad\quad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

These eigenvectors span the space, which means any vectors can be expressed as a linear combination of them

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2}\left\{(y-x)\begin{pmatrix} -1 \\ 1 \end{pmatrix} + (x+y)\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}.$$

This means the initial values of the state probability ν can be expressed in terms of the eigenvectors of $P_1^\dagger$. Denote the matrix

$$P_{tot} = (P_2^\dagger)^m (P_1^\dagger)^m (P_2^\dagger)^m (P_1^\dagger)^m$$

which is the total transition matrix in one period. Proceeding we have

$$P_{tot}^N \nu(0) = \frac{1}{2}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix} + (1-2p)^{2mN}(1-2q)^{2mN}[\nu_+(0) - \nu_-(0)]\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$$

with the added condition $\nu_-(0) + \nu_+(0) = 1$. Now note the following

$$\text{For} \quad n \in T_1 \quad \nu(t) = (P_1^\dagger)^n P_{tot}^N \nu(0)$$

$$\text{For} \quad n \in T_2 \quad \nu(t) = (P_2^\dagger)^{n-m}(P_1^\dagger)^m P_{tot}^N \nu(0)$$

$$\text{For} \quad n \in T_3 \quad \nu(t) = (P_1^\dagger)^{n-2m}(P_2^\dagger)^m (P_1^\dagger)^m P_{tot}^N \nu(0)$$

$$\text{For} \quad n \in T_4 \quad \nu(t) = (P_2^\dagger)^{n-3m}(P_1^\dagger)^m (P_2^\dagger)^m (P_1^\dagger)^m P_{tot}^N \nu(0)$$

This completes the proof. □

3.2.3 Discrete Time Markov Chain - Invariant Measures, Relaxation Time and Fourier Transform

We consider a discrete Markov Chain on $\{-1, +1\}$, that is

$$Y_t^\epsilon = \pm 1$$

and the time is

$$t \in \{0, 1, 2, \ldots\}$$

and the probabilities for being in $Y_t^\epsilon = -1$ or $Y_t^\epsilon = +1$ at time t are given by the state probabilities $\nu_\pm(t)$

$$P\left(Y_t^\epsilon = -1\right) = \nu_-(t) \quad \text{and} \quad P\left(Y_t^\epsilon = +1\right) = \nu_+(t).$$

The probabilities of transitions occurring as given in the transition matrices changes with period T. After a very long time the state probabilities $\nu_\pm(\cdot)$ should not depend on the initial state probabilities $\nu_\pm(0)$. At time infinity $\nu_\pm(\cdot)$ should also be cyclic on $[0, T]$. Let the time be given by $t = NT + n$ where N is a discrete number of periods. This leads us to define the invariant measure as the state probabilities in the limit as $N \longrightarrow \infty$

$$\overline{\nu}(n) := \lim_{N \longrightarrow \infty} \nu(NT + n)$$

and since $\overline{\nu}(\cdot)$ is periodic on $[0, T]$ it should also satisfy

$$\overline{\nu}(t + T) = \prod_{i=t}^{i=t+T-1} P_i^\dagger \overline{\nu}(t)$$

where P_i are the transition matrices, that is to say $\overline{\nu}$ is invariant over one period of application of the transition matrices. This brings us to the following.

Corollary 3.3. *For the state probabilities in Theorem 3.1 the invariant measures are*

$$\textit{For} \quad t \in T_1 \quad \overline{\nu}(t) = \frac{1}{p+q} \begin{pmatrix} q \\ p \end{pmatrix} - \frac{p-q}{p+q} \times \frac{\lambda^t}{1+\lambda^m} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\textit{For} \quad t \in T_2 \quad \overline{\nu}(t) = \frac{1}{p+q} \begin{pmatrix} p \\ q \end{pmatrix} + \frac{p-q}{p+q} \times \frac{\lambda^{t-m}}{1+\lambda^m} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Corollary 3.4. *For the state probabilities in Theorem 3.2 the invariant measures are*

$$\overline{\nu}(t) = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

The proof is easy and omitted. The fact that the $\overline{\nu}$ are invariant over one period of application of the transition matrices follow from the proof of the Theorems.

The rate of convergence to the invariant measure would depend on the value of p and q themselves. Define the relaxation time T_{relax} as the first time $t = T_{relax}$ such that

$$\left|\overline{\nu}\left(T_{relax}\right) - \nu\left(T_{relax}\right)\right| \leq e^{-1}$$

which is a measure of the rate of convergence to the invariant measure.

Consider the averaged Markov Chain over many realisations. This is related to the invariant measure by

$$\langle Y_t^\epsilon \rangle = \overline{\nu}_-(t)(-1) + \overline{\nu}_+(t)(+1) = \overline{\nu}_+(t) - \overline{\nu}_-(t).$$

The Fourier Transform of the averaged Markov Chain $\langle \tilde{Y}^\epsilon_\omega \rangle$ is often studied (see Chapter 4), that is

$$\langle \tilde{Y}^\epsilon_\omega \rangle = \mathcal{F}\left(\langle Y^\epsilon_t \rangle\right).$$

We can Fourier Transform both the alternating saddle $p \neq q$ case and the synchronised saddle $p = q$ case. This brings us to the following.

Corollary 3.5. *For the Markov Chain in Theorem 3.1 the Fourier Transform of the averaged trajectory is*

$$\langle \tilde{Y}^\epsilon_\omega \rangle = \frac{1}{T}\frac{p-q}{p+q}\left(1-e^{-i\pi\omega}\right)\left\{\frac{1-e^{-i\pi\omega}}{1-e^{-i\pi\omega/m}} - \frac{2}{1+\lambda^m}\frac{1-\lambda^m e^{-i\pi\omega}}{1-\lambda e^{-i\pi\omega/m}}\right\}.$$

Proof. Notice that

$$\overline{\nu}_-(t+m) = \overline{\nu}_+(t) \quad \text{and} \quad \overline{\nu}_+(t+m) = \overline{\nu}_-(t)$$

so we have

$$\begin{aligned}
\langle \tilde{Y}^\epsilon_\omega \rangle &= \mathcal{F}\left(\langle Y^\epsilon_t \rangle\right) \\
&= \frac{1}{T}\sum_{t=0}^{2m-1} \langle Y^\epsilon_t \rangle e^{-2\pi i\omega t/2m} \\
&= \frac{1}{T}\sum_{t=0}^{m-1} \langle Y^\epsilon_t \rangle e^{-2\pi i\omega t/2m} + \frac{1}{T}\sum_{t=m}^{2m-1} \langle Y^\epsilon_t \rangle e^{-2\pi i\omega t/2m} \\
&= \frac{1}{T}\sum_{t=0}^{m-1} \left[\overline{\nu}_+(t) - \overline{\nu}_-(t)\right] e^{-2\pi i\omega t/2m} + \frac{1}{T}\sum_{t=0}^{m-1} \left[\overline{\nu}_-(t) - \overline{\nu}_+(t)\right] e^{-2\pi i\omega (t+m)/2m} \\
&= \frac{1}{T}\sum_{t=0}^{m-1} \left[\overline{\nu}_+(t) - \overline{\nu}_-(t)\right] \left(e^{-2\pi i\omega t/2m} - e^{-2\pi i\omega (t+m)/2m}\right) \\
&= \frac{1}{T}\left(1-e^{-i\pi\omega}\right)\sum_{t=0}^{m-1} \left[\overline{\nu}_+(t) - \overline{\nu}_-(t)\right] e^{-i\pi\omega t/m} \\
&= \frac{1}{T}\left(1-e^{-i\pi\omega}\right)\frac{p-q}{p+q}\sum_{t=0}^{m-1} \left(1 - 2\frac{\lambda^t}{1+\lambda^m}\right) e^{-i\pi\omega t/m} \\
&= \frac{1}{T}\frac{p-q}{p+q}\left(1-e^{-i\pi\omega}\right)\left\{\frac{1-e^{-i\pi\omega}}{1-e^{-i\pi\omega/m}} - \frac{2}{1+\lambda^m}\frac{1-\lambda^m e^{-i\pi\omega}}{1-\lambda e^{-i\pi\omega/m}}\right\}.
\end{aligned}$$

This completes the proof. □

Corollary 3.6. *For the Markov Chain in Theorem 3.2 the Fourier Transform of the averaged trajectory is*

$$\langle \tilde{Y}^\epsilon_\omega \rangle = 0.$$

Again the proof is trivial and omitted. If we study the Fourier Transform at $\omega = 1$, this would be the same as studying the driving frequency, which is the frequency at which the transition matrices are changing. The physical intuition is that one has the most significant response at this frequency.

3.3 Continuous Time Markov Chain

Let the time be continuous. This is to say time t belongs to

$$t \in \mathbb{R}.$$

The Markov Chain is a time dependent stochastic process with values -1 or $+1$,

$$Y_t = \pm 1.$$

Let p and q be real functions

$$p : \mathbb{R} \longrightarrow \mathbb{R} \quad \text{and} \quad q : \mathbb{R} \longrightarrow \mathbb{R}$$

and p and q are periodic on $[0, T]$

$$p(t + T) = p(t) \quad \text{and} \quad q(t + T) = q(t).$$

Let $A \subseteq [0, T]$ be a subset of the interval $[0, T]$. The probability of Y_t transiting from $Y_t = -1$ to $Y_t = +1$ for the times in A, $t \in A$, is denoted by

$$p_{-1+1}(A).$$

Similarly the probability of Y_t transiting from $Y_t = +1$ to $Y_t = -1$ for the times in A, $t \in A$, is denoted by

$$p_{+1-1}(A).$$

The probability of Y_t staying at -1 in the time $t \in A$ is given by

$$p_{-1-1}(A) = 1 - p_{-1+1}(A).$$

Similarly the probability of Y_t staying at $+1$ in the time $t \in A$ is given by

$$p_{+1+1}(A) = 1 - p_{+1-1}(A).$$

If A is a small time interval $A = [t, t + \delta t]$ then the following infinitesimal representation can be made

$$\begin{aligned} p_{-1+1}([t, t + \delta t]) &= p(t)\delta t \\ p_{+1-1}([t, t + \delta t]) &= q(t)\delta t. \end{aligned}$$

Now we consider a small change in the state probabilities at times t and $t+\delta t$.

$$\begin{pmatrix} \nu_-(t+\delta t) \\ \nu_+(t+\delta t) \end{pmatrix} = \begin{pmatrix} p_{-1-1}([t,t+\delta t]) & p_{-1+1}([t,t+\delta t]) \\ p_{+1-1}([t,t+\delta t]) & p_{+1+1}([t,t+\delta t]) \end{pmatrix}^\dagger \begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix}$$

$$= \begin{pmatrix} 1-p(t)\delta t & p(t)\delta t \\ q(t)\delta t & 1-q(t)\delta t \end{pmatrix}^\dagger \begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix}$$

then

$$\nu(t+\delta t)-\nu(t) = \begin{pmatrix} \nu_-(t+\delta t) \\ \nu_+(t+\delta t) \end{pmatrix} - \begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix}$$

$$= \begin{pmatrix} -p(t)\delta t & p(t)\delta t \\ q(t)\delta t & -q(t)\delta t \end{pmatrix}^\dagger \begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix}$$

$$\frac{\nu(t+\delta t)-\nu(t)}{\delta t} = \begin{pmatrix} -p(t) & p(t) \\ q(t) & -q(t) \end{pmatrix}^\dagger \begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix}$$

which in the limit of small δt leads to a differential equation describing the behaviour of $\nu(t)$

$$\frac{d\nu}{dt} = Q^\dagger \nu \tag{3.2}$$

where the infinitesimal generator Q is defined as

$$Q = \begin{pmatrix} -p(t) & p(t) \\ q(t) & -q(t) \end{pmatrix}.$$

Note that the transpose of Q is taken in Equation 3.2. The aim now is to derive the state probability by solving this differential equation for various forms of p and q. The extra conditions we use are

$$\nu_-(t)+\nu_+(t) = 1$$
$$\nu'_-(t)+\nu'_+(t) = 0$$

for all times t and the initial conditions at $t=0$ are $\nu_-(0)$ and $\nu_+(0)$.

3.3.1 Continuous Time Markov Chain - Alternating Saddles $p \neq q$

Notice that p may be interpreted as the probability of escape from the left well and q as the probability of escape from the right well. If p and q are cyclic over $[0,T]$, then this can be interpreted as modelling a potential with periodic forcing in continuous time.

Theorem 3.7. *Let $p \neq q$ and $t \geq 0$. The state probabilities are given by*

$$\nu_-(t) = \frac{\nu_-(0) + \int_0^t q(s) \exp\left\{\int_0^s p(u) + q(u)\, du\right\}\, ds}{\exp\left\{\int_0^t p(u) + q(u)\, du\right\}}$$

$$\nu_+(t) = \frac{\nu_+(0) + \int_0^t p(s) \exp\left\{\int_0^s p(u) + q(u)\, du\right\}\, ds}{\exp\left\{\int_0^t p(u) + q(u)\, du\right\}}$$

Proof. The differential equations we want to solve are given by

$$\begin{aligned}\frac{d}{dt}\begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix} &= \begin{pmatrix} -p(t) & p(t) \\ q(t) & -q(t) \end{pmatrix}^\dagger \begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix} \\ &= \begin{pmatrix} -p(t) & q(t) \\ p(t) & -q(t) \end{pmatrix} \begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix}\end{aligned}$$

which gives

$$\begin{aligned}\frac{d\nu_-}{dt} &= -p\nu_- + q\nu_+ \\ \frac{d\nu_+}{dt} &= p\nu_- - q\nu_+\end{aligned}$$

and by using $\nu_- + \nu_+ = 1$ we get

$$\frac{d\nu_-}{dt} + (p+q)\nu_- = q \tag{3.3}$$

$$\frac{d\nu_+}{dt} + (p+q)\nu_+ = p. \tag{3.4}$$

We will only solve for $\nu_-(t)$. The case for $\nu_+(t)$ is similar. Equation 3.3 can easily be solved with an integrating factor

$$\frac{d}{dt}\left\{\nu_-(t) \exp\left\{\int_0^t p(u) + q(u)\, du\right\}\right\} = q(t) \exp\left\{\int_0^t p(u) + q(u)\, du\right\}$$

and proceeding we have

$$\nu_-(t) \exp\left\{\int_0^t p(u) + q(u)\, du\right\} - \nu_-(0) = \int_0^t q(s) \exp\left\{\int_0^s p(u) + q(u)\, du\right\}\, ds$$

which rearranges to give

$$\nu_-(t) = \frac{\nu_-(0) + \int_0^t q(s) \exp\left\{\int_0^s p(u) + q(u)\, du\right\}\, ds}{\exp\left\{\int_0^t p(u) + q(u)\, du\right\}}.$$

This completes the proof. □

3.3.2 Continuous Time Markov Chain - Synchronised Saddles $p = q$

If $p = q$ for continuous time, then this can be modelled as both wells of the potential always being at the same height but moving together.

Theorem 3.8. *Let $p = q$ and $t \geq 0$. The state probabilities are given by*

$$\nu_-(t) = \frac{1}{2} - \frac{\nu_+(0) - \nu_-(0)}{2} \exp\left\{-2\int_0^t p(s)\,ds\right\}$$
$$\nu_+(t) = \frac{1}{2} + \frac{\nu_+(0) - \nu_-(0)}{2} \exp\left\{-2\int_0^t p(s)\,ds\right\}$$

Proof. The differential equations we have to solve are given by

$$\frac{d}{dt}\begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix} = p(t)\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} \nu_-(t) \\ \nu_+(t) \end{pmatrix}$$

which gives

$$\frac{d\nu_-}{dt} = p\left(\nu_+ - \nu_-\right) \tag{3.5}$$
$$\frac{d\nu_+}{dt} = p\left(\nu_- - \nu_+\right) \tag{3.6}$$

and subtracting Equation 3.5 away from Equation 3.6 leads to

$$\frac{d}{dt}\left(\nu_+ - \nu_-\right) = -2p\left(\nu_+ - \nu_-\right).$$

Denote the difference by

$$z(t) = \nu_+(t) - \nu_-(t)$$

which gives the differential equation we need to solve to

$$\frac{dz}{dt} = -2pz$$
$$\int_{z(0)}^{z(t)} \frac{dz}{z} = -2\int_0^t p(s)\,ds$$
$$\ln(z(t)) - \ln(z(0)) = -2\int_0^t p(s)\,ds$$
$$z(t) = z(0)\exp\left\{-2\int_0^t p(s)\,ds\right\}$$

and using $z = \nu_+ - \nu_-$ and $\nu_- + \nu_+ = 1$ rearranges the solution to

$$\nu_-(t) = \frac{1}{2} - \frac{\nu_+(0) - \nu_-(0)}{2} \exp\left\{-2\int_0^t p(s)\,ds\right\}$$

$$\nu_+(t) = \frac{1}{2} + \frac{\nu_+(0) - \nu_-(0)}{2} \exp\left\{-2\int_0^t p(s)\,ds\right\}.$$

This completes the proof. □

3.3.3 Continuous Time Markov Chain - Invariant Measures and Fourier Transform

As in the discrete time case we can compute the corresponding invariant measures.

Corollary 3.9. *For the state probabilities in Theorem 3.7 the invariant measures are*

$$\overline{\nu}_-(t) = \frac{\int_0^t p(s)g(s)\,ds}{g(t)} + \frac{\int_0^T p(s)g(s)\,ds}{g(t)\,(g(T)-1)}$$

$$\overline{\nu}_+(t) = \frac{\int_0^t q(s)g(s)\,ds}{g(t)} + \frac{\int_0^T q(s)g(s)\,ds}{g(t)\,(g(T)-1)}$$

where

$$g(t) = \exp\left\{\int_0^t p(u) + q(u)\,du\right\}.$$

Proof. We derive the invariant measure for $\overline{\nu}_-(t)$. The case for $\overline{\nu}_+(t)$ is similar. Consider the fact that $p(\cdot)$ and $q(\cdot)$ are cyclic on $[0, T]$ and let i be an integer, then the following integral can be rewritten as

$$\begin{aligned}
\int_{iT}^{(i+1)T} p(s)g(s)\,ds &= \int_0^T p(s)g(iT+s)\,ds \\
&= \int_0^T p(s)g(iT)g(s)\,ds \\
&= g(iT)\int_0^T p(s)g(s)\,ds \\
&= g(T)^i \int_0^T p(s)g(s)\,ds.
\end{aligned}$$

Let the time be given by $NT + t$ where N is an integer number of periods. This means the following integral can be written as

$$\int_0^{NT+t} p(s)g(s)\,ds = \int_{NT}^{NT+t} p(s)g(s)\,ds + \sum_{i=0}^{N-1} \int_{iT}^{(i+1)T} p(s)g(s)\,ds$$

$$= \int_0^t p(s)g(NT+s)\,ds + \int_0^T p(s)g(s)\,ds \sum_{i=0}^{N-1} g(T)^i$$

$$= g(T)^N \int_0^t p(s)g(s)\,ds + \int_0^T p(s)g(s)\,ds \sum_{i=0}^{N-1} g(T)^i.$$

So the state probability is equal to

$$\nu_-(NT+t) = \frac{\nu_-(0) + \int_0^{NT+t} q(s) \exp\left\{\int_0^s p(u)+q(u)\,du\right\}\,ds}{\exp\left\{\int_0^{NT+t} p(u)+q(u)\,du\right\}}$$

$$= \frac{\nu_-(0) + g(T)^N \int_0^t p(s)g(s)\,ds + \int_0^T p(s)g(s)\,ds \sum_{i=0}^{N-1} g(T)^i}{g(T)^N g(t)}$$

$$= \frac{\nu_-(0)}{g(T)^N g(t)} + \frac{\int_0^t p(s)g(s)\,ds}{g(t)} + \frac{\int_0^T p(s)g(s)\,ds}{g(t)} \frac{1}{g(T)^N} \sum_{i=0}^{N-1} g(T)^i$$

$$= \frac{\nu_-(0)}{g(T)^N g(t)} + \frac{\int_0^t p(s)g(s)\,ds}{g(t)} + \frac{\int_0^T p(s)g(s)\,ds}{g(t)} \frac{1}{g(T)-1}\left(1 - \frac{1}{g(T)^N}\right).$$

Letting $N \longrightarrow \infty$ gives the required result. □

Corollary 3.10. *For the state probabilities in Theorem 3.8 the invariant measures are*

$$\overline{\nu}(t) = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The proof is trivial and omitted. Similar to the discrete time case we can also study the Fourier Transform of the averaged Markov Chain.

Corollary 3.11. *For the Markov Chain in Theorem 3.7 the Fourier Transform of the averaged Markov Chain is*

$$\mathcal{F}\left(\langle Y_t^\epsilon \rangle\right) = \int_{-\infty}^{+\infty} \left(\frac{\int_0^t \left[q(s)-p(s)\right] g(s)\,ds}{g(t)} + \frac{\int_0^T \left[q(s)-p(s)\right] g(s)\,ds}{g(t)\left(g(T)-1\right)} \right) e^{-i2\pi\omega t}\,dt.$$

Corollary 3.12. *For the Markov Chain in Theorem 3.8 the Fourier Transform of the averaged Markov Chain is*

$$\mathcal{F}\left(\langle Y_t^\epsilon \rangle\right) = 0.$$

3.4 Probability Density Function of Escape Times

The escape rates from the left to right are denoted by $R_{-1+1}(\cdot)$ and right to left escape rates are denoted by $R_{+1-1}(\cdot)$. The PDFs for the escape times are given by the Theorem below.

Theorem 3.13. *Let u be the time of entry into a well, then the PDFs for the escape occurring at time $t > u$ are*

$$p_-(t,u) = R_{-1+1}(t) \exp\left\{ -\int_u^t R_{-1+1}(s)\, ds \right\}$$

$$p_+(t,u) = R_{+1-1}(t) \exp\left\{ -\int_u^t R_{+1-1}(s)\, ds \right\}$$

where $p_-(t,u)$ is for left to right and $p_+(t,u)$ is for right to left.

Proof. We consider escaping from the left well. The right well is similar. Divide the time interval $[u,t]$ into many small time intervals

$$\delta t = \frac{t-u}{N}.$$

Similar to how we derived the invariant measures we want to derive the probability of escape in a very small time interval $[t, t+\delta t]$. This is given by

$$\begin{aligned} p_{-1+1}([t,t+\delta t]) &= p(t)\delta t \\ &= 1 - e^{-R_{-1+1}(t)\delta t} \\ &\approx R_{-1+1}(t)\delta t \end{aligned}$$

which is valid for small δt. Large deviations allow us to say even more about the escape time τ_{-1+1} and τ_{+1-1}. Theorem 1 in [35] shows that it is an exponentially distributed random variable. The probability of staying in the left well is given by

$$\begin{aligned} p_{-1-1}([t,t+\delta t]) &= 1 - p_{-1+1}([t,t+\delta t]) \\ &= 1 - p(t)\delta t \\ &= 1 - \left(1 - e^{-R_{-1+1}(t)\delta t}\right) \\ &= e^{-R_{-1+1}(t)\delta t}. \end{aligned}$$

We want to know the probability of escaping in the time interval $[t, t+\delta t]$ given that the particle has entered at u and stayed up to time t. This is given by

$$p_{-1-1}([u,t])p_{-1+1}([t,t+\delta t]) = \prod_{i=1}^{N} p_{-1-1}\left([u+(i-1)\delta t, u+i\delta t]\right) p_{-1+1}([t,t+\delta t])$$

$$
\begin{aligned}
&= \prod_{i=1}^{N} \exp\left\{-R_{-1+1}\left(u+(i-1)\delta t\right)\delta t\right\} p_{-1+1}([t,t+\delta t]) \\
&= \exp\left\{\sum_{i=1}^{N} -R_{-1+1}\left(u+(i-1)\delta t\right)\delta t\right\} p_{-1+1}([t,t+\delta t]) \\
&= \exp\left\{-\int_{u}^{t} R_{-1+1}(s)\,ds\right\} R_{-1+1}(t)\delta t.
\end{aligned}
$$

This completes the proof. □

3.4.1 Normalised Time Probability Density Function of Escape Times

The period of the forcing is T and we can make a change of variables to normalised time

$$t^{norm} = \frac{t^{real}}{T}$$

which measures time in how many periods have elapsed. This rearranges the PDFs to

$$
\begin{aligned}
p_{-}(t,u) &= TR_{-1+1}(Tt)\exp\left\{-T\int_{u}^{t} R_{-1+1}(Ts)\,ds\right\} \\
p_{+}(t,u) &= TR_{+1-1}(Tt)\exp\left\{-T\int_{u}^{t} R_{+1-1}(Ts)\,ds\right\}
\end{aligned}
$$

where u, t and s are in normalised time. Note that $R_{-1+1}(\cdot)$ and $R_{+1-1}(\cdot)$ always have their arguments in real time and $R_{-1+1}(\cdot)$ and $R_{+1-1}(\cdot)$ always give the averaged number of transitions per real unit time. Expressing the PDF in normalised time can be found in [41].

3.4.2 Perfect Phase Approximation of Probability Density Function of Escape Times

The PDF for the escape times derived in Theorem 3.13 had to differentiate between left and right escapes and are conditioned on the time u of entrance into the well. Suppose now that t is the escape time from any well, which does not differentiate between left and right escape. Note that t is the actual time it takes to escape from a well and is not a time coordinate. The PDF for t is given by

$$p_{tot}(t) = \frac{1}{2}\int_{0}^{T} p_{-}(t+u,u)m_{-}(u) + p_{+}(t+u,u)m_{+}(u)\,du.$$

This is because after a long time has elapsed we would expect that many transitions would have occurred between left and right. The number of transitions escaping from the left

and right should be roughly the same. The $m_-(u)$ is a PDF for the time of entrance into the left well and the $m_+(u)$ is a PDF for the time of entrance into the right well. We may not have explicit expressions for $m_-(u)$ and $m_+(u)$. We derive an approximate expression for p_{tot} without an explicit expressions for $m_-(u)$ and $m_+(u)$. Let $m_-(u)$ and $m_+(u)$ be approximated by

$$m_-(u) \approx \delta\left(u - T/2\right)$$

$$m_+(u) \approx \frac{1}{2}\delta\left(u\right) + \frac{1}{2}\delta\left(u - T\right)$$

where $\delta(\cdot)$ is the Dirac delta function. This approximation is used because in the SDEs which we will simulate, the times when transition into the left well is greatest is at half the period $u = \frac{T}{2}$ and the times when transition into the right well is greatest is at $u = 0$ and $u = T$. Due to the fact that $m_-(u)$ and $m_+(u)$ are probabilities a factor of $\frac{1}{2}$ is used in $m_+(u)$. Progressing we have

$$\begin{aligned}
p_{tot}(t) &= \frac{1}{2}\int_0^T p_-(t+u,u)m_-(u) + p_+(t+u,u)m_+(u)\,du \\
&\approx \frac{1}{2}\int_0^T p_-(t+u,u)\delta\left(u - T/2\right) + p_+(t+u,u)\left(\frac{1}{2}\delta\left(u\right) + \frac{1}{2}\delta\left(u-T\right)\right)du \\
&= \frac{1}{2}\left\{p_-(t+T/2,T/2) + \frac{1}{2}p_+(t,0) + \frac{1}{2}p_+(t+T,T)\right\} \\
&= \frac{1}{2}\left\{p_-(t+T/2,T/2) + \frac{1}{2}p_+(t,0) + \frac{1}{2}p_+(t+0,0)\right\} \\
&= \frac{1}{2}\left\{p_-(t+T/2,T/2) + p_+(t,0)\right\} \\
&= p_+(t,0).
\end{aligned}$$

This is because for the simulations which we are going to do, the Kramers' rate satisfy $R_{-1+1}(t) = R_{+1-1}(t+T/2)$ (see later in Chapter 5 for the geometry of the Mexican Hat Toy Model which justifies this). Thus the following approximation

$$p_{tot} \approx p_+(t,0)$$

is only valid for the simulations we do, and not for a general potential. We call this way of approximating $m_-(u)$ and $m_+(u)$ the perfect phase approximation.

3.5 Adiabatic Large Deviation

We have to stress that this thesis is built on three approximations, which form the backbone of all the research presented. These are small noise approximation, adiabatic approximation and perfect phase approximation.

Perfect phase approximation only works for small noise. This is because the noise is so small the particle will only escape when the maximum probability to escape has arrived. When the minimum probability to escape is present it will almost never escape. This is the idea behind the perfect phase approximation.

Notice one subtlety behind all the theory presented in this Chapter. The derivations involved probabilities of escape p and q and the escape rates R_{-1+1} and R_{+1-1}. But it was assumed that p, q, R_{-1+1} and R_{+1-1} are accurately known no matter how large or small the noise level ϵ is and no matter how fast or slow the driving frequency Ω is. But such ideal expressions for p, q, R_{-1+1} and R_{+1-1} are not known.

When we come to do the analysis in Chapter 7, the p_{tot} is calculated with the approximation $p_{tot} \approx p_+(t, 0)$. When the rates R_{-1+1} and R_{+1-1} are needed they are calculated using Kramers' formula as though it is escape from a static potential in the small noise limit. This means an oscillatory potential is being approximated by a static potential which is the adiabatic approximation.

In the paper [54] the adiabatic approximation was justified in the small noise, slow forcing limit using time dependent large deviation theory, that is, it was shown asymptotically the escape times are given by the adiabatic approximation. This result is only for the leading term, whether the analogue result holds for the Kramers' rate is unknown.

Chapter 4

Theory of Analysis of Stochastic Resonance

We present different criteria that have been used to define stochastic resonance. This includes the six measures, which are linear response, signal-to-noise ratio, energy, out-of-phase measures, relative entropy and entropy. A new statistical test called the conditional Kolmogorov-Smirnov test is introduced.

4.1 Six Measures of Stochastic Resonance

We introduce six possible criteria of measuring how close a process is to exhibiting stochastic resonance [41, 45]. These six criteria are closely related to linear response [25, 26], signal-to-noise ratio [27, 28] and distribution of escape times [28–30]. We call them the six measures denoted by M_1, M_2, M_3, M_4, M_5 and M_6. Recall that the SDE we want to study is

$$\dot{X}_t^\epsilon = -\nabla V_0 + F\cos\Omega t + \epsilon\,\dot{W}_t$$

where $V_0 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is the unperturbed potential, F is the forcing, Ω is the forcing frequency, ϵ is the noise level and W_t is a Wiener process in two dimensions, which when rewritten into separate components are

$$\begin{aligned} dx &= \left[-\frac{\partial V_0}{\partial x} + F_x\cos\Omega t\right]dt + \epsilon\; dw_x \\ dy &= \left[-\frac{\partial V_0}{\partial y} + F_y\cos\Omega t\right]dt + \epsilon\; dw_y \end{aligned}$$

where w_x and w_y are two independent Wiener processes. The solution to these equations is the trajectory in two dimensions

$$X_t^\epsilon = (x_t, y_t)\,.$$

This diffusion can be reduced to a Markov Chain on $\{-1,+1\}$ denoted by

$$Y_t^\epsilon = \pm 1$$

where by definition of the Markov Chain the escape times are the same as the diffusion case (see Chapter 3). The probability of the Markov Chain being in one state at time t is given by the state probabilities

$$P\left(Y_t^\epsilon = -1\right) = \nu_-(t) \quad \text{and} \quad P\left(Y_t^\epsilon = +1\right) = \nu_+(t).$$

In what follows we will consider so large times, that the relaxation time has effectively elapsed for both the diffusion and Markov Chain, in other words the state probability would have effectively converged to the invariant measure $\overline{\nu}$. This means that over one period $T = 2\pi/\Omega$ of the forcing, the invariant measures will have the properties

$$\overline{\nu}_\pm(t) = \overline{\nu}_\pm(t+T) \quad \text{and} \quad \overline{\nu}_\pm(t) = \overline{\nu}_\mp(t+T/2).$$

We obtain the averaged trajectories given by

$$\langle X_t^\epsilon \rangle = E\left(X_t^\epsilon\right) \quad \text{and} \quad \langle Y_t^\epsilon \rangle = E\left(Y_t^\epsilon\right)$$

which are the trajectories obtained after averaging over many realisations. Notice that $\langle Y_t^\epsilon \rangle$ is related to the invariant measures by

$$\langle Y_t^\epsilon \rangle = \overline{\nu}_+(t) - \overline{\nu}_-(t).$$

We introduce the Out-of-Phase Markov Chain defined by

$$\overline{Y}_t^\epsilon = \begin{cases} 0 & \text{if} \quad Y_t^\epsilon = -1 \quad \text{and} \quad mod(t,T) \leq T/2 \\ 1 & \text{if} \quad Y_t^\epsilon = -1 \quad \text{and} \quad mod(t,T) > T/2 \\ 1 & \text{if} \quad Y_t^\epsilon = +1 \quad \text{and} \quad mod(t,T) \leq T/2 \\ 0 & \text{if} \quad Y_t^\epsilon = +1 \quad \text{and} \quad mod(t,T) > T/2 \end{cases}$$

and similarly the averaged Out-of-Phase Markov Chain is defined by

$$\left\langle \overline{Y}_t^\epsilon \right\rangle = E\left(\overline{Y}_t^\epsilon\right).$$

Define two new functions by

$$\phi^-(t) = \begin{cases} 1 & \text{if} \quad mod(t,T) \leq T/2 \\ 0 & \text{if} \quad mod(t,T) > T/2 \end{cases}$$

$$\phi^+(t) = \begin{cases} 0 & \text{if} \quad mod(t,T) \leq T/2 \\ 1 & \text{if} \quad mod(t,T) > T/2. \end{cases}$$

The following trajectories are Fourier transformed

$$\tilde{x}(\omega) = \mathcal{F}\left(\langle x_t \rangle\right) = \langle \mathcal{F}\left(x_t\right) \rangle$$

$$\tilde{Y}(\omega) = \mathcal{F}\left(\langle Y_t^\epsilon \rangle\right) = \langle \mathcal{F}\left(x_t\right) \rangle.$$

The linear response is defined as the intensity of the Fourier Transform at the driving frequency Ω [20]

$$X_{lin} = \left| \tilde{x}\left(\frac{\Omega}{2\pi}\right) \right| \quad \text{and} \quad Y_{lin} = \left| \tilde{Y}\left(\frac{\Omega}{2\pi}\right) \right|.$$

Now we can define the six measures. For the diffusion case only M_1 and M_2 are defined.

$$M_1 = \frac{1}{F} X_{lin}$$
$$M_2 = \frac{1}{\epsilon F} X_{lin}$$

where F is the magnitude of the forcing. For the Markov Chain M_1, M_2, M_3, M_4, M_5 and M_6 are all defined as

$$M_1 = \frac{1}{F} Y_{lin}$$
$$M_2 = \frac{1}{\epsilon F} Y_{lin}$$
$$M_3 = \int_0^T \langle Y_t^\epsilon \rangle^2 \, dt$$
$$M_4 = \int_0^T \left\langle \overline{Y}_t^\epsilon \right\rangle dt$$
$$M_5 = \int_0^T \phi^-(t) \ln\left(\frac{\phi^-(t)}{\overline{\nu}_-(t)}\right) + \phi^+(t) \ln\left(\frac{\phi^+(t)}{\overline{\nu}_+(t)}\right) dt$$
$$M_6 = \int_0^T -\overline{\nu}_-(t) \ln \overline{\nu}_-(t) - \overline{\nu}_+(t) \ln \overline{\nu}_+(t) \, dt.$$

Note that in the definition of the six measures it is assumed that the process has relaxed to equilibrium. We give a few physical interpretation of the six measures M_1, M_2, M_3, M_4, M_5 and M_6. The M_1 is the intensity of the driving frequency Ω in the spectrum of the Fourier transform. The M_2 is sometimes called signal-to-noise ratio as it compares this intensity to the noise level ϵ. The M_3 is sometimes called the energy. The M_4 is sometimes called the out-of-phase measure since it measures the amount of time the Markov Chain spends in the "wrong" well. The M_5 and M_6 are sometimes called relative entropy and entropy respectively, since they measure how far away the invariant measures are from being constant. If the invariant measures are constant then these six measures will also be constant. Thus it can be understood that these six measures is a measure of how far away the invariant measures are from being constant. M_6 measures how non-constant the invariant measure is. M_5 is extremal if the invariant measure is constant.

[20] See Appendix B.2 for how the linear response is calculated numerically.

4.2 Statistical Tests

We will measure the escape time for many consecutive transitions. This will result in a collection of measurements of escape times

$$\tau_1, \tau_2, \ldots, \tau_n.$$

A new method for analysing such a collection of measurements is presented.

4.2.1 Kolmogorov-Smirnov Test

First we recall results about the Kolmogorov-Smirnov statistic and the Kolmogorov-Smirnov test [55]. Let

$$\xi_1, \xi_2, \ldots, \xi_n$$

be n independently and identically distributed real random variables. Each ξ_i is distributed with PDF $f(\cdot)$ as in

$$P(\xi_i \in A) = \int_A f(s)\, ds$$

and distributed with CDF $F(\cdot)$ as in

$$P(\xi_i \leq x) = F(x) = \int_{-\infty}^{x} f(s)\, ds.$$

Define a function by $F_n(\cdot)$ by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(-\infty,x]}(\xi_i)$$

where $\mathbf{1}_A$ is the indicator function for a set A. We may think of $\xi_1, \xi_2, \ldots, \xi_n$ as n empirical or numerical realisations of the same random variable ξ. The $F_n(x)$ is therefore an approximation to the CDF of ξ that is empirically found using $\xi_1, \xi_2, \ldots, \xi_n$, therefore $F_n(\cdot)$ is called the empirical CDF. Consider the supremum metric on the space of real continuous functions. Consider the distance between the real and the empirical CDF in this metric.

$$\begin{aligned} D_n &= \|F_n - F\|_\infty \\ &= \sup_{x \in \mathbb{R}} \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(-\infty,x]}(\xi_i) - F(x) \right| \end{aligned}$$

where D_n is called the Kolmogorov-Smirnov statistic or KS statistic. Intuitively we would expect D_n to tend to zero as n increases, that is

$$\lim_{n \longrightarrow \infty} D_n = 0$$

if the ξ_i are distributed by $F(\cdot)$. There are times when we experimentally obtain n values of a random variable $\xi_1, \xi_2, \ldots, \xi_n$, and want to test whether they are distributed by a CDF $F(\cdot)$. We define what we mean by the null hypothesis.

Definition 4.1. *Let $\xi_1, \xi_2, \ldots, \xi_n$ be n real random variables. The null hypothesis is that each ξ_i is independently distributed with CDF $F(x)$.*

We want to know how large or small D_n needs to be before deciding whether to reject the null hypothesis. The following Theorem offers a remarkable answer to this problem.

Theorem 4.2. *Suppose the null hypothesis is true, then the distribution of D_n depends only on n.*

Notice that D_n is in itself a real random variable. The PDF and CDF of D_n is a function of n only, and will be the same whatever $F(\cdot)$ is. This distribution is called the KS distribution and tables are available upto $n = 100$. There is a Theorem which describes the asymptotic behaviour of the KS distribution [56, 57].[21]

Theorem 4.3. *In the limit $n \longrightarrow \infty$, $\sqrt{n}D_n$ is asymptotically Kolmogorov distributed with the CDF*

$$Q(x) = 1 - 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2x^2}$$

that is to say

$$\lim_{n\longrightarrow\infty} P(\sqrt{n}D_n \leq x) = Q(x).$$

4.2.2 Conditional Kolmogorov-Smirnov Test

Let $\zeta_1, \zeta_2, \ldots, \zeta_n$ be n iid real random variables. They are n empirical observations of a random variable ζ. Now suppose that each of the $\xi_1, \xi_2, \ldots, \xi_n$ is conditioned and dependent on the corresponding $\zeta_1, \zeta_2, \ldots, \zeta_n$. This means a conditional PDF $f(\cdot, \cdot)$ gives the probability

$$P(\xi_i \in A \,|\, \zeta_i) = \int_A f(s, \zeta_i)\, ds$$

and the conditional CDF $F(\cdot, \cdot)$ is

$$P(\xi_i \leq x \,|\, \zeta_i) = F_{\zeta_i}(x) = \int_{-\infty}^{x} f(s, \zeta_i)\, ds.$$

But $\xi_1, \xi_2, \ldots, \xi_n$ are empirical measurements of the same random variable ξ. The PDF for ξ is given by

$$P(\xi \in A) = \int_A \int_{-\infty}^{+\infty} f(s, u)m(u)\, ds\, du$$

[21] There appears to be topographical errors in the literature for the limiting function. Some sources cite $Q_1 = 1 - 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2x^2}$ (see [58–60]) and some cite $Q_2 = 1 - 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-k^2x^2}$ (see [56, 57]). But the proof of Theorem 1 in [57] shows $Q = Q_1$. Nevertheless in this thesis we use $Q = Q_1$ which actually gives smaller and more conservative values of the metric D_n that are needed.

and the CDF for ξ is

$$P(\xi \leq x) = F(x) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f(s,u)m(u)\,ds\,du$$

where $m(\cdot)$ is the PDF for ζ, that is

$$P(\zeta \in A) = \int_A m(s)\,ds.$$

In our context we have the problem that the random variables are not identically distributed under the null hypothesis. The $\xi_1, \xi_2, \ldots, \xi_n$ and $\zeta_1, \zeta_2, \ldots, \zeta_n$ are obtained experimentally and $F_{\zeta_i}(\xi_i)$ can be calculated but a PDF for ζ_i, that is $m(\cdot)$, has no easy expression. We still want to perform a statistics test that is similar to the KS test even in such situations where the distribution $m(\cdot)$ of ζ is unknown. First we define what we call the total null hypothesis and the conditional null hypothesis.

Definition 4.4. *Let $\xi_1, \xi_2, \ldots, \xi_n$ be n empirical observations of a random variable ξ. The total null hypothesis is that ξ is distributed with the CDF $F(\cdot)$. The conditional null hypothesis is that each ξ_i is distributed with the conditional CDF $F_{\zeta_i}(\cdot)$.*

A new statistical test is developed, which is similar to the KS test.

Theorem 4.5. *Suppose the conditional null hypothesis is true. Let $F_{\zeta_i}(\cdot)$ be continuous. Let S_n be the statistic given by*

$$S_n = \sup_{x \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{[0,x]} \left(F_{\zeta_i}(\xi_i) \right) - x \right|$$

then S_n is KS distributed.

Proof. Denote

$$Y_i = F_{\zeta_i}(\xi_i)$$

which means

$$\begin{aligned} P\left(Y_i \leq x\right) &= P\left(F_{\zeta_i}(\xi_i) \leq x\right) \\ &= P\left(\xi_i \leq F_{\zeta_i}^{-1}(x)\right) \\ &= F_{\zeta_i}\left(F_{\zeta_i}^{-1}(x)\right) \\ &= x \end{aligned}$$

and $0 \leq Y_i \leq 1$, so Y_i is uniformly distributed on $[0,1]$. Note that $F_{\zeta_i}(\cdot)$ is a function of one variable only. Let

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{[0,x]}(Y_i) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{[0,x]} \left(F_{\zeta_i}(\xi) \right)$$

where $F_n(\cdot)$ is the empirical CDF of a uniformly distributed random variable, computed using n observations. The statistic S_n is the suprenum metric

$$\begin{aligned} S_n &= \|F_n - x\|_\infty \\ &= \sup_{x\in[0,1]} \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{[0,x]} \left(F_{\zeta_i}(\xi_i)\right) - x \right|. \end{aligned}$$

So clearly S_n is KS distributed. □

We call S_n the conditional KS statistic. Compare this to the original KS statistic, which under the assumption of the total null hypothesis can be rewritten as

$$D_n = \sup_{x\in\mathbb{R}} \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(-\infty,x]}(\xi_i) - F(x) \right| = \sup_{x\in[0,1]} \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{[0,x]} \left(F(\xi_i)\right) - x \right|.$$

When both the total and conditional hypothesis are true D_n and S_n are KS distributed, that is

$$P\left(D_n \in A\right) = P\left(S_n \in A\right) \quad \text{and} \quad P\left(D_n \leq x\right) = P\left(S_n \leq x\right).$$

The subtlety here is that D_n and S_n are different objects, yet they have the same distribution. D_n is KS distributed under the total null hypothesis, whereas S_n is KS distributed under the conditional null hypothesis. This can be explained in another way. We have n experimental observations of a random variable ξ denoted by $\xi_1, \xi_2, \ldots, \xi_n$ and each are conditioned on observations of another random variable ζ denoted by $\zeta_1, \zeta_2, \ldots, \zeta_n$. The D_n is KS distributed if the random variable ξ is distributed by CDF $F(\cdot)$, but the S_n is KS distributed if each ξ_i is conditionally distributed by the CDF $F_{\zeta_i}(\cdot)$.

Chapter 5

Mexican Hat Toy Model

The main object of consideration of this project, which is called the Mexican Hat Toy Model, is now introduced. Let $a > 0$, $b > 0$ and $V_0 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a real function from the plane to the line. The unperturbed potential is defined as

$$V_0(x, y) = \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 \quad \text{where} \quad r = \sqrt{x^2 + y^2}.$$

Let $F_x, F_y \in \mathbb{R}$ be the forcing. The potential with forcing V_F is defined as

$$\begin{aligned} V_F(x, y) &= \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 + F_x x + F_y y \\ &= \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 + \mathbf{F} \cdot \mathbf{x} \\ &= V_0 + \mathbf{F} \cdot \mathbf{x} \end{aligned}$$

written more compactly in vector notation. When V_F is defined using $+\mathbf{F} \cdot \mathbf{x}$ we say positive forcing. Alternatively if V_F is defined using $-\mathbf{F} \cdot \mathbf{x}$ we say negative forcing. The V_F is defined with a positive forcing because for the rest of this Chapter we will study the critical points which are solutions to the simultaneous equations

$$\frac{\partial V_F}{\partial x} = 0 \quad \text{and} \quad \frac{\partial V_F}{\partial y} = 0.$$

The properties of the critical points will change as F is increased from zero, therefore it is convenient to define V_F with a positive forcing. The behaviour of the critical points are studied for different cases. The main aim is to find the positions and nature of the critical points for a range of parameter values. This is complex due to several cases to be considered and previewed in the following Theorem, which is one of the main conclusions of this Chapter. Although this Theorem only considers the case for non-negative forcing $F_x \geq 0$ and $F_y \geq 0$, the case for $F_x < 0$ and $F_y < 0$ is similar by considering the symmetry of the potential.

Theorem 5.1. *Let $F_x \geq 0$, $F_y \geq 0$, $a > 0$, $b > 0$. Note that definitions of constants are at the end. The positions and nature of the critical points of the Mexican Hat Toy Model $V_F(\cdot)$ are for the following range of parameters.*

For $F_x = 0$, $F_y = 0$ and $b < \frac{1}{2}$

$$\begin{array}{cc} (0,0) & \textit{hill} \\ (\pm\sqrt{1+2a},0) & \textit{well} \\ (0,\pm\sqrt{1-2b}) & \textit{saddle} \end{array}$$

For $F_x = 0$, $F_y = 0$ and $b \geq \frac{1}{2}$

$$\begin{array}{cc} (0,0) & \textit{saddle} \\ (\pm\sqrt{1+2a},0) & \textit{well} \end{array}$$

For $F_x > 0$, $F_y = 0$, $b < \frac{1}{2}$ and the following values of F_x

$$\textit{For any} \quad F_x > 0 \quad \left\{ \begin{array}{ll} (x_0,0) & \textit{well} \end{array} \right.$$

$$\textit{and if} \quad F_x < F_x^{sad} \left\{ \begin{array}{ll} (x_1,0) & \textit{well} \\ (x_2,0) & \textit{hill} \\ (x_{saddle},\pm y_{saddle}) & \textit{saddle} \end{array} \right.$$

$$\textit{or} \quad F_x^{sad} < F_x < F_x^{crit} \left\{ \begin{array}{lll} (x_1,0) & \textit{saddle} & \textit{for} \ \sqrt{1-2b} \in R_1 \\ (x_1,0) & \textit{well} & \textit{for} \ \sqrt{1-2b} \in R_2 \\ (x_2,0) & \textit{hill} & \textit{for} \ \sqrt{1-2b} \in R_1 \\ (x_2,0) & \textit{saddle} & \textit{for} \ \sqrt{1-2b} \in R_2 \\ (x_{saddle},\pm y_{saddle}) & \textit{nonexistent} & \end{array} \right.$$

$$\textit{or} \quad F_x = F_x^{sad} < F_x^{crit} \left\{ \begin{array}{lll} (x_1,0) & \textit{unidentified} & \textit{for} \ \sqrt{1-2b} \in R_1 \\ (x_1,0) & \textit{well} & \textit{for} \ \sqrt{1-2b} \in R_2 \\ (x_2,0) & \textit{hill} & \textit{for} \ \sqrt{1-2b} \in R_1 \\ (x_2,0) & \textit{unidentified} & \textit{for} \ \sqrt{1-2b} \in R_2 \\ (x_{saddle},\pm y_{saddle}) & \textit{unidentified} & \end{array} \right.$$

$$\textit{or} \quad F_x = F_x^{crit} \left\{ \begin{array}{ll} (x_1,0) & \textit{unidentified} \\ (x_2,0) & \textit{unidentified} \end{array} \right.$$

$$\textit{or} \quad F_x > F_x^{crit} \left\{ \begin{array}{ll} (x_1,0) & \textit{nonexistent} \\ (x_2,0) & \textit{nonexistent} \end{array} \right.$$

For $F_x > 0$, $F_y = 0$, $b \geq \frac{1}{2}$ and the following values of F_x

$$F_x < F_x^{crit} \begin{cases} (x_0, 0) & well \\ (x_1, 0) & well \\ (x_2, 0) & saddle \end{cases}$$

$$F_x > F_x^{crit} \begin{cases} (x_0, 0) & well \\ (x_1, 0) & nonexistent \\ (x_2, 0) & nonexistent \end{cases}$$

$$F_x = F_x^{crit} \begin{cases} (x_0, 0) & well \\ (x_1, 0) & unidentified \\ (x_2, 0) & unidentified \end{cases}$$

For $F_x = 0$, $F_y > 0$, $b \geq \frac{1}{2}$ and the following values of F_y

$$F_y < F_y^{crit} \quad \begin{cases} (0, y_1) & saddle \\ (0, y_2) & hill \end{cases}$$

$$F_y < F_y^{sad} \quad \begin{cases} (0, y_0) & saddle \\ (\pm x_{well}, y_{well}) & well \end{cases}$$

$$F_y > F_y^{crit} \quad \begin{cases} (0, y_1) & nonexistent \\ (0, y_2) & nonexistent \end{cases}$$

$$F_y > F_y^{sad} \quad \begin{cases} (0, y_0) & well \\ (\pm x_{well}, y_{well}) & nonexistent \end{cases}$$

$$F_y = F_y^{crit} \quad \begin{cases} (0, y_1) & unidentified \\ (0, y_2) & unidentified \end{cases}$$

$$F_y = F_y^{sad} \quad \begin{cases} (0, y_0) & unidentified \\ (\pm x_{well}, y_{well}) & unidentified \end{cases}$$

For $F_x = 0$, $F_y > 0$, $b > \frac{1}{2}$ and the following values of F_y

$$F_y < F_y^{sad} \quad \begin{cases} (0, y_0) & saddle \\ (\pm x_{well}, y_{well}) & well \end{cases}$$

$$F_y = F_y^{sad} \quad \{ \ (0, y_0) = (\pm x_{well}, y_{well}) \quad unidentified$$

$$F_y > F_y^{sad} \quad \begin{cases} (0, y_0) & well \\ (\pm x_{well}, y_{well}) & nonexistent \end{cases}$$

where

$$
\begin{aligned}
x_k &= -\frac{2}{\sqrt{3}}\sqrt{1+2a}\cos\left\{\frac{1}{3}\tan^{-1}\left(\frac{\sqrt{4(1+2a)^3-27F_x^2}}{F_x\sqrt{27}}\right)+\frac{2\pi}{3}k\right\}\\
y_k &= -\frac{2}{\sqrt{3}}\sqrt{1-2b}\cos\left\{\frac{1}{3}\tan^{-1}\left(\frac{\sqrt{4(1-2b)^3-27F_y^2}}{F_y\sqrt{27}}\right)+\frac{2\pi}{3}k\right\}\\
x_{saddle} &= \frac{F_x}{2(a+b)}\\
y_{saddle} &= \sqrt{(1-2b)-\left(\frac{F_x}{2(a+b)}\right)^2}\\
x_{well} &= \sqrt{(1+2a)-\left(\frac{F_y}{2(a+b)}\right)^2}\\
y_{well} &= \frac{-F_y}{2(a+b)}\\
F_x^{crit} &= \sqrt{\frac{4(1+2a)^3}{27}}\\
F_x^{sad} &= 2(a+b)\sqrt{1-2b}\\
F_y^{crit} &= \sqrt{\frac{4(1-2b)^3}{27}}\\
F_y^{sad} &= 2(a+b)\sqrt{1+2a}\\
R_1 &= \left(\frac{1}{\sqrt{3}}\sqrt{1+2a},\sqrt{1+2a}\right)\\
R_2 &= \left(0,\frac{1}{\sqrt{3}}\sqrt{1+2a}\right).
\end{aligned}
$$

The proof is given in a series of Lemmas for each of the six different cases. Theorem 5.4 proves the case for $F_x = 0$ and $F_y = 0$, Theorem 5.10 proves the case for $F_x > 0$, $F_y = 0$ and $b < \frac{1}{2}$, Theorem 5.11 proves the case for $F_x > 0$, $F_y = 0$ and $b \geq \frac{1}{2}$, Theorem 5.16 proves the case for $F_x = 0$, $F_y > 0$ and $b < \frac{1}{2}$ and Theorem 5.17 proves the case for $F_x = 0$, $F_y > 0$ and $b \geq \frac{1}{2}$. All the notation used will be consistent with this current Theorem 5.1. The following standard result is used.

Theorem 5.2. *Let $V : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be twice differentiable everywhere. Let H be the Hessian at a critical point (x_0, y_0). The nature of the critical point can be determined by*

$$
\begin{array}{lll}
\det H < 0 & \Rightarrow & \textit{saddle}\\
\det H > 0 \quad \textit{then} & \left\{\begin{array}{lll}
\textit{if} \quad \frac{\partial^2 V}{\partial x^2} > 0 & \Rightarrow & \textit{well}\\
\textit{if} \quad \frac{\partial^2 V}{\partial x^2} < 0 & \Rightarrow & \textit{hill}
\end{array}\right.
\end{array}
$$

We recall results about the cubic equation.

Theorem 5.3. *Let $a_3, a_2, a_1, a_0 \in \mathbb{R}$ where $a_3 \neq 0$. Consider the cubic equation*

$$a_3x^3 + a_2x^2 + a_1x + a_0 = 0$$

and its discriminant

$$\Delta = 18a_3a_2a_1a_0 - 4a_2^3a_0 + a_2^2a_1^2 - 4a_3a_1^3 - 27a_3^2a_0^2$$

then following statements hold

if $\Delta > 0$ *then the equation has 3 distinct real roots.*
if $\Delta = 0$ *then the equation has a multiple real root and all its roots are real.*
if $\Delta < 0$ *then the equation has 1 real root and 2 complex conjugate roots.*

and the three roots of the equations are

$$x_k = -\frac{1}{3a_3}\left(a_2 + e^{i\psi_k}C + e^{-i\psi_k}\frac{\Delta_0}{C}\right)$$

where

$$\begin{aligned}
\psi_0 &= 0, \quad \psi_1 = 2\pi/3, \quad \psi_2 = 4\pi/3 \\
C &= \sqrt[3]{\frac{\Delta_1 + \sqrt{-27\Delta}}{2}} \\
\Delta_0 &= a_2^2 - 3a_3a_1 \\
\Delta_1 &= 2a_2^3 - 9a_3a_2a_1 + 27a_3^2a_0.
\end{aligned}$$

5.1 Case $F_x = 0$ and $F_y = 0$

The case for no forcing $F = 0$ is considered first.

Theorem 5.4. *When $F_x = F_y = 0$ the critical points of the potential have the following properties. For $b < \frac{1}{2}$ the critical points and their nature are*

$b < \frac{1}{2}$	$(0, 0)$	*hill*
	$(\pm\sqrt{1+2a}, 0)$	*well*
	$(0, \pm\sqrt{1-2b})$	*saddle*

For $b \geq \frac{1}{2}$ the critical points and their nature are

$b \geq \frac{1}{2}$	$(0, 0)$	*saddle*
	$(\pm\sqrt{1+2a}, 0)$	*well*

The proof is trivial and omitted.

5.2 Case $F_x > 0$ and $F_y = 0$

When forcing is only in the x direction two cases are considered separately, that is for $b < \frac{1}{2}$ and $b \geq \frac{1}{2}$. The case for $F_x \leq 0$ is similar.

5.2.1 Case $F_x > 0$, $F_y = 0$ and $b < \frac{1}{2}$

When there is no forcing there are five critical points. Intuitively as forcing is increased the system could gradually start to deviate away from having five critical points. The critical points may collide and coincide. The structure of the following proofs are first determining the bounds on the critical points and then determining their nature. We have the following consequence which uses the solution and theory of the cubic equation with three real roots.

Theorem 5.5. *Let $F_x > 0$, $F_y = 0$ and $b < \frac{1}{2}$. Let F_x be bounded by*

$$F_x \leq F_x^{sad} \quad \textit{and} \quad F_x \leq F_x^{crit}$$

where

$$F_x^{sad} = 2(a+b)\sqrt{1-2b} \quad \textit{and} \quad F_x^{crit} = \sqrt{\frac{4(1+2a)^3}{27}}$$

then there are five critical points given by

$$\begin{aligned} &(x_k, 0) \quad k = 1, 2, 3 \\ &(x_{saddle}, \pm y_{saddle}) \end{aligned}$$

where

$$\begin{aligned} x_{saddle} &= \frac{F_x}{2(a+b)} \\ y_{saddle} &= \sqrt{(1-2b) - \left(\frac{F_x}{2(a+b)}\right)^2} \\ x_k &= -\frac{2}{\sqrt{3}}\sqrt{1+2a}\,\cos\left(\phi + k\frac{2\pi}{3}\right) \\ \phi &= \frac{1}{3}\tan^{-1}(l) \\ l &= \frac{\sqrt{4(1+2a)^3 - 27F_x^2}}{F_x\sqrt{27}} \\ k &= 0, \quad k = 1, \quad k = 2. \end{aligned}$$

Proof. The simultaneous equations to be solved are

$$\frac{\partial V_F}{\partial x} = x(x^2+y^2) - (1+2a)x + F_x = 0 \tag{5.1}$$

$$\frac{\partial V_F}{\partial y} = y(x^2 + y^2) - (1 - 2b)y = 0. \tag{5.2}$$

Equation 5.2 holds if either $(x^2 + y^2) - (1 - 2b) = 0$ or $y = 0$. The $(x^2 + y^2) - (1 - 2b) = 0$ case is considered first, which gives $(x^2 + y^2) = (1 - 2b)$. Substituting this into Equation 5.1 gives x as

$$x_{saddle} = \frac{F_x}{2(a + b)}$$

which when substituted back into $(x^2 + y^2) = (1 - 2b)$ gives y as

$$y_{saddle} = \pm\sqrt{(1 - 2b) - \left(\frac{F_x}{2(a + b)}\right)^2}.$$

For the $y = 0$ case, Equation 5.1 becomes

$$x^3 - (1 + 2a)x + F_x = 0$$

which is a cubic equation. Solving this cubic equation using the notation in Theorem 5.3 gives

$$\Delta_0 = 3(1 + 2a), \quad \Delta_1 = 27F_x, \quad \Delta = 4(1 + 2a)^3 - 27F_x^2$$

which gives[22]

$$\begin{aligned} C &= \sqrt[3]{\frac{27F_x + i\sqrt{27\Delta}}{2}} \\ &= \left[\left(\frac{27F_x}{2}\right)^2 + \left(\frac{\sqrt{27\Delta}}{2}\right)^2\right]^{1/6} \exp\left\{i(1/3)\tan^{-1}\left(\frac{\sqrt{27\Delta}}{27F_x}\right)\right\} \end{aligned}$$

which simplifies to

$$\begin{aligned} \left(\frac{27F_x}{2}\right)^2 + \left(\frac{\sqrt{27\Delta}}{2}\right)^2 &= \frac{1}{4}\left(27 \times 27F_x^2 + 27\Delta\right) \\ &= 27(1 + 2a)^3, \end{aligned}$$

after letting

$$\phi = (1/3)\tan^{-1}\left(\frac{\sqrt{27\Delta}}{27F_x}\right)$$

[22]After noting that $(x + iy)^{1/3} = (x^2 + y^2)^{1/6} \exp\left\{i(1/3)\tan^{-1}(y/x)\right\}$.

we have

$$\begin{aligned} C &= \sqrt{3(1+2a)}\; e^{i\phi} \\ \frac{\Delta_0}{C} &= \frac{3(1+2a)}{\sqrt{3(1+2a)}}\; e^{-i\phi} \\ &= \sqrt{3(1+2a)}\; e^{-i\phi} \end{aligned}$$

which gives the 3 solution as

$$\begin{aligned} x_k &= -\frac{1}{3}\sqrt{3(1+2a)}\left(e^{i(\phi+\psi_k)} + e^{-i(\phi+\psi_k)}\right) \\ &= -\frac{2}{3}\sqrt{3(1+2a)}\cos(\phi+\psi_k) \\ &= -\frac{2}{3}\sqrt{3(1+2a)}\cos\left\{\frac{1}{3}\tan^{-1}\left(\frac{\sqrt{4(1+2a)^3 - 27F_x^2}}{F_x\sqrt{27}}\right) + \psi_k\right\} \end{aligned}$$

where $k = 0, 1, 2$, $\psi_0 = 0$, $\psi_1 = \frac{2\pi}{3}$ and $\psi_2 = \frac{4\pi}{3}$. Now notice that y_{saddle} requires taking the square root of a real number. This means y_{saddle} will be real if and only if the argument under the square root is positive

$$\begin{aligned} y_{saddle} &= \pm\sqrt{(1-2b) - \left(\frac{F_x}{2(a+b)}\right)^2} \\ &\in \mathbb{R} \\ &\Leftrightarrow 0 \leq (1-2b) - \left(\frac{F_x}{2(a+b)}\right)^2 \\ &\Leftrightarrow F_x \leq F_x^{sad} \\ \text{where} \quad F_x^{sad} &= 2(a+b)\sqrt{1-2b}. \end{aligned}$$

Notice also how the argument inside the $\tan^{-1}(\cdot)$ function contains a square root as well, which is actually the square root of the discriminant. The three cubic roots x_k would be real if and only if the discriminant is positive

$$\begin{aligned} \sqrt{\Delta} &= \sqrt{4(1+2a)^3 - 27F_x^2} \\ &\in \mathbb{R} \\ &\Leftrightarrow 0 \leq 4(1+2a)^3 - 27F_x^2 \\ &\Leftrightarrow F_x \leq F_x^{crit} \\ \text{where} \quad F_x^{crit} &= \sqrt{\frac{4(1+2a)^3}{27}} \end{aligned}$$

which clearly puts bounds on the forces. This completes the proof. □

Next there is a simple but useful Lemma.

Lemma 5.6. *The function l (as in Theorem 5.5) is monotone in F_x for $F_x < F_x^{crit}$.*

Proof. We differentiate l with respect to F_x.

$$\begin{aligned}\frac{dl}{dF_x} &= \frac{d}{dF_x}\left(\frac{\sqrt{4(1+2a)^3 - 27F_x^2}}{F_x\sqrt{27}}\right) \\ &= \frac{1}{F_x\sqrt{27}}\left(4(1+2a)^3 - 27F_x^2\right)^{-\frac{1}{2}}(-2F_x)\,\frac{1}{2}\,27 \\ &\quad + \frac{\sqrt{4(1+2a)^3 - 27F_x^2}}{F_x\sqrt{27}}\,\left(\frac{-1}{F_x^2}\right) \\ &\leq 0\end{aligned}$$

which is always negative since we assumed $F_x < F_x^{crit}$ for the square roots to be real and forcing is assumed to be in the positive direction. □

Although the monotonicity of l is trivial, it would prove essential for the next series of reasoning. It is also easy to see that

$$\begin{aligned} F_x = 0 \quad &\text{then} \quad l = +\infty \\ F_x = F_x^{crit} \quad &\text{then} \quad l = 0. \end{aligned}$$

Since the derivative of l is negative this means that l would decrease from $+\infty$ to 0 as F_x increase from 0 to F_x^{crit}. This function being monotone means it would decrease to 0 without any oscillations. For short this means

$$l = \infty \downarrow 0 \quad \text{as} \quad F_x = 0 \uparrow F_x^{crit}.$$

But $\phi = \frac{1}{3}\tan^{-1}(l)$, with $\tan^{-1}$ is also monotone over $[-\infty, +\infty]$. So similarly we can also say

$$\phi = \frac{\pi}{6} \downarrow 0 \quad \text{as} \quad F_x = 0 \uparrow F_x^{crit}$$

monotonically for increasing F_x. This means that for $0 \leq F_x \leq F_x^{crit}$ we would have

$$0 \leq \phi \leq \frac{\pi}{6}.$$

We note the following values of the $\cos(\cdot)$ function.

$$\begin{aligned} \cos(0) = 1 \quad &\cos\left(\tfrac{\pi}{6}\right) = \tfrac{\sqrt{3}}{2} \\ \cos\left(\tfrac{2\pi}{3}\right) = \tfrac{-1}{2} \quad &\cos\left(\tfrac{2\pi}{3} + \tfrac{\pi}{6}\right) = \tfrac{-\sqrt{3}}{2} \\ \cos\left(\tfrac{4\pi}{3}\right) = \tfrac{-1}{2} \quad &\cos\left(\tfrac{4\pi}{3} + \tfrac{\pi}{6}\right) = 0. \end{aligned}$$

From this we can bound $\cos(\cdot)$ for the three values of k for $0 \leq F_x \leq F_x^{crit}$.

$$\begin{aligned} k=0 \quad \tfrac{\sqrt{3}}{2} &\leq \cos\left(\phi + k\tfrac{2\pi}{3}\right) \leq 1 \\ k=1 \quad \tfrac{-\sqrt{3}}{2} &\leq \cos\left(\phi + k\tfrac{2\pi}{3}\right) \leq \tfrac{-1}{2} \\ k=2 \quad \tfrac{-1}{2} &\leq \cos\left(\phi + k\tfrac{2\pi}{3}\right) \leq 0. \end{aligned}$$

We also note that the $\cos(\cdot)$ function is monotone on $[0,\pi]$ and $[\pi, 2\pi]$. But $\left(\phi + k\frac{2\pi}{3}\right)$ is in the intervals where $\cos(\cdot)$ is monotone, therefore we can say that the three critical points on the x-axis x_k are also monotone in F_x. We can now have bounds on the x_k for $0 \leq F_x \leq F_x^{crit}$.

$$\begin{aligned} \frac{-2}{\sqrt{3}}\sqrt{1+2a} \leq x_0 \leq -\sqrt{1+2a} \\ \frac{1}{\sqrt{3}}\sqrt{1+2a} \leq x_1 \leq \sqrt{1+2a} \\ 0 \leq x_2 \leq \frac{1}{\sqrt{3}}\sqrt{1+2a}. \end{aligned}$$

Using the monotonicity of x_k we get that

$$\begin{aligned} x_0: \quad -\sqrt{1+2a} &\longrightarrow \tfrac{-2}{\sqrt{3}}\sqrt{1+2a} \quad \text{as} \quad F_x \to F_x^{crit} \\ x_1: \quad \sqrt{1+2a} &\longrightarrow \tfrac{1}{\sqrt{3}}\sqrt{1+2a} \quad \text{as} \quad F_x \to F_x^{crit} \\ x_2: \quad 0 &\longrightarrow \tfrac{1}{\sqrt{3}}\sqrt{1+2a} \quad \text{as} \quad F_x \to F_x^{crit}. \end{aligned}$$

The monotonicity of x_k means the movements of the x_k are always in one direction and they will never oscillate. We obtain the following

Lemma 5.7. *Let $F_x > 0$, $F_y = 0$ and $b < \frac{1}{2}$. These three scenarios hold.*

If $F_x < F_x^{crit}$ we have 3 critical points on the x-axis: $x_0 < x_2 < x_1$.

If $F_x = F_x^{crit}$ we have 2 critical points on the x-axis: $x_0 < x_2 = x_1$.

If $F_x > F_x^{crit}$ we have 1 critical point on the x-axis: $x_0 < \frac{-2}{\sqrt{3}}\sqrt{1+2a}$ and $x_0 \to -\infty$ monotonically with increasing F_x.

Proof. The three x_k are solutions to a cubic equation. This cubic equation was derived assuming $y = 0$. The other critical points $(x_{saddle}, \pm y_{saddle})$ were derived assuming $y \neq 0$.

If $F_x < F_x^{crit}$ the discriminant of this cubic equation dictates that there should be three distinct real solution. The bounds on x_0, x_1 and x_2 show that $x_0 < x_2 < x_1$. If $F_x = F_x^{crit}$ the discriminant of this cubic equation dictates that there should be at least two repeated solution. It was shown that $x_1 = x_2 = \frac{1}{\sqrt{3}}\sqrt{1+2a}$ when $F_x = F_x^{crit}$. The bound on x_0, x_1 and x_2 shows that $x_0 < x_2 = x_1$. If $F_x > F_x^{crit}$ the discriminant of this cubic equation dictates that there should only be one real solution. If we can show that $x_0 < 0$ is real then we are done. For $F_x > F_x^{crit}$ the l function becomes

$$l = i\frac{\sqrt{27F_x^2 - 4(1+2a)^3}}{F_x\sqrt{27}}.$$

Differentiating the imaginary part gives

$$\begin{aligned}\frac{d}{dF_x}(-il) &= \frac{d}{dF_x}\left(\frac{\sqrt{27F_x^2 - 4(1+2a)^3}}{F_x\sqrt{27}}\right) \\ &= \frac{F_x\sqrt{27}}{\sqrt{27F_x^2 - 4(1+2a)^3}}\left(1 - \frac{27F_x^2 - 4(1+2a)^3}{27F_x}\right) \\ &> 0,\end{aligned}$$

since we have assumed $F_x > F_x^{crit}$. This also shows that the imaginary part of l is monotonically increasing as F_x increases. We also note that

$$0 \leq (-il) \leq 1$$

because $(-il) = 0$ when $F_x = F_x^{crit}$ and $(-il) = 1$ when $F_x = \infty$ and the increase in $(-il)$ is monotone. Using $\tan^{-1}(\cdot)$ defined for complex arguments gives

$$\begin{aligned}\phi &= \frac{1}{3}\tan^{-1}(l) \\ &= \frac{1}{3} \times \frac{1}{2} i\{\ln(1 - il) - \ln(1 + il)\} \\ &= \frac{1}{6} i\{\ln(1+\epsilon) - \ln(1-\epsilon)\} \quad \text{where} \quad 0 \leq \epsilon \leq 1\end{aligned}$$

after letting $\epsilon = -il$. Notice that ϕ now has zero real part, which means we can denote ϕ with a real γ by writing

$$\phi = i\gamma.$$

We also note that γ will be monotonically decreasing as F_x increases because

$$\begin{aligned}\frac{d}{d\epsilon}\left[\ln(1+\epsilon) - \ln(1-\epsilon)\right] &= \frac{1}{1+\epsilon} - \frac{1}{1-\epsilon} \\ &= \frac{-2\epsilon}{(1+\epsilon)(1-\epsilon)} \\ &< 0 \quad \text{for} \quad F_x > F_x^{crit}.\end{aligned}$$

Because it was shown that as F_x increases from $F_x = F_x^{crit}$ to $F_x = \infty$, ϵ would increase from $\epsilon = 0$ to $\epsilon = 1$, which means γ would monotonically decrease. So the critical point may now be written as

$$\begin{aligned}x_0 &= -\frac{2}{\sqrt{3}}\sqrt{1+2a}\cos(\phi) \\ &= -\frac{2}{\sqrt{3}}\sqrt{1+2a}\ \frac{e^{i(i\gamma)} + e^{-i(i\gamma)}}{2}\end{aligned}$$

$$= -\frac{2}{\sqrt{3}}\sqrt{1+2a}\ \cosh(\gamma)$$

which means $x_0 < \frac{-2}{\sqrt{3}}\sqrt{1+2a}$ and monotonically decreasing. Note that $\gamma = 0$ when $F_x = F_x^{crit}$. □

Now we consider a special case of the forcing when $F_x = F_x^{sad}$. This gives

$$\begin{aligned}(x_{saddle}, \pm y_{saddle}) &= (x_{saddle}, 0) \\ &= (\sqrt{1-2b}, 0) \\ \Rightarrow x_{saddle} &= \sqrt{1-2b}\end{aligned}$$

which seemingly adds a fourth critical point onto the x-axis. This brings us to the next Lemma.

Lemma 5.8. *$F_x^{sad} \leq F_x^{crit}$ holds.*

Proof. Proof by contradiction. Assume that $F_x^{sad} > F_x^{crit}$. Let $F_x = F_x^{sad}$ which means $(x_{saddle}, \pm y_{saddle}) = (x_{saddle}, 0)$ as a new critical point on the x-axis. Now we have to show that $(x_{saddle}, 0)$ is not one of $(x_0, 0)$, $(x_1, 0)$ or $(x_2, 0)$. The expressions for the critical points mean we would always have $x_{saddle} > 0$. But $F_x = F_x^{sad}$ also implies $F_x > F_x^{crit}$ which by Lemma 5.7 means the only critical point is $x_0 < 0$ which is a contradiction. □

The next Lemma will be useful in avoiding complicated manipulation of trigonometric identities when it comes to proving properties about the critical points.

Lemma 5.9. *Let $F_x > 0$, $F_y = 0$ and $b < \frac{1}{2}$. If $F_x = F_x^{sad}$ we must have either $x_{saddle} = x_1$ or $x_{saddle} = x_2$. If $F_x = F_x^{sad} = F_x^{crit}$ then $x_{saddle} = x_1 = x_2$.*

Proof. For strictly positive $F_x > 0$ some of the bounds on the critical points would have to be made strict inequalities. This means $x_0 < 0$, $x_1 > 0$, $x_2 > 0$ and $x_{saddle} > 0$. By the time $F_x = F_x^{sad}$, x_{saddle} would be a critical point on the x-axis. By Lemma 5.8 we must always have $F_x^{sad} \leq F_x^{crit}$. If $F_x^{sad} < F_x^{crit}$ then by Lemma 5.7 there must be three distinct critical points on the x-axis, so we must have either $x_{saddle} = x_1$ or $x_{saddle} = x_2$ (as $x_0 < 0$). If $F_x^{sad} = F_x^{crit}$ then again by Lemma 5.7 $x_1 = x_2$ and there can only be two critical points on the x-axis, therefore $x_{saddle} = x_1 = x_2$. □

Now we are ready for one of the main Theorems of this Chapter. The ultimate aim is to find the nature and position of all the critical points under different values of the forcing F_x.

Theorem 5.10. *Let $F_x > 0$, $F_y = 0$ and $b < \frac{1}{2}$. The positions and nature of the critical points are as follows*

$$\textit{For any} \quad F_x > 0 \quad \Big\{ \ (x_0, 0) \quad \textit{well}$$

$$\textit{and if} \quad F_x < F_x^{sad} \begin{cases} (x_1, 0) & \textit{well} \\ (x_2, 0) & \textit{hill} \\ (x_{saddle}, \pm y_{saddle}) & \textit{saddle} \end{cases}$$

$$\textit{or} \quad F_x^{sad} < F_x < F_x^{crit} \begin{cases} (x_1, 0) & \textit{saddle} & \textit{for} & \sqrt{1-2b} \in R_1 \\ (x_1, 0) & \textit{well} & \textit{for} & \sqrt{1-2b} \in R_2 \\ (x_2, 0) & \textit{hill} & \textit{for} & \sqrt{1-2b} \in R_1 \\ (x_2, 0) & \textit{saddle} & \textit{for} & \sqrt{1-2b} \in R_2 \\ (x_{saddle}, \pm y_{saddle}) & \textit{nonexistent} & & \end{cases}$$

$$\textit{or} \quad F_x = F_x^{sad} < F_x^{crit} \begin{cases} (x_1, 0) & \textit{unidentified} & \textit{for} & \sqrt{1-2b} \in R_1 \\ (x_1, 0) & \textit{well} & \textit{for} & \sqrt{1-2b} \in R_2 \\ (x_2, 0) & \textit{hill} & \textit{for} & \sqrt{1-2b} \in R_1 \\ (x_2, 0) & \textit{neither} & \textit{for} & \sqrt{1-2b} \in R_2 \\ (x_{saddle}, \pm y_{saddle}) & \textit{unidentified} & & \end{cases}$$

$$\textit{or} \quad F_x = F_x^{crit} \begin{cases} (x_1, 0) & \textit{unidentified} \\ (x_2, 0) & \textit{unidentified} \end{cases}$$

$$\textit{or} \quad F_x > F_x^{crit} \begin{cases} (x_1, 0) & \textit{nonexistent} \\ (x_2, 0) & \textit{nonexistent} \end{cases}$$

Proof. For the three x_k we note that for $0 < F_x \leq F_x^{crit}$ they are elements of the intervals

$$\begin{aligned} x_0 &\in \left[\frac{-2}{\sqrt{3}}\sqrt{1+2a}, -\sqrt{1+2a}\right) \\ x_1 &\in \left[\frac{1}{\sqrt{3}}\sqrt{1+2a}, \sqrt{1+2a}\right) \\ x_2 &\in \left(0, \frac{1}{\sqrt{3}}\sqrt{1+2a}\right]. \end{aligned}$$

Note also that at $F_x = F_x^{sad}$ the associated value of $x_{saddle}(a, b, F_x^{sad}) = \sqrt{1-2b}$ can only ever be elements of certain intervals.

$$\begin{aligned} \text{For} \quad F_x^{sad} < F_x^{crit} \quad \text{either} \quad & \sqrt{1-2b} \in R_1 := \left(\frac{1}{\sqrt{3}}\sqrt{1+2a}, \sqrt{1+2a}\right) \\ \text{or} \quad & \sqrt{1-2b} \in R_2 := \left(0, \frac{1}{\sqrt{3}}\sqrt{1+2a}\right) \\ \text{and if} \quad F_x^{sad} = F_x^{crit} \quad \text{then} \quad & \sqrt{1-2b} = \frac{1}{\sqrt{3}}\sqrt{1+2a} \end{aligned}$$

where R_1 and R_2 are defined as above. These statements above can be justified as follows. Lemma 5.8 says $F_x^{sad} \leq F_x^{crit}$. If $F_x = F_x^{sad} = F_x^{crit}$ then Lemma 5.9 says $x_{saddle} = x_1 =$

$x_2 = \frac{1}{\sqrt{3}}\sqrt{1+2a}$. Also, $\sqrt{1-2b}$ must always live in the regions specified, because we must always have $0 < \sqrt{1-2b} < \sqrt{1+2a}$ for $0 < b < \frac{1}{2}$. Or, to justify it in another way, if $\sqrt{1-2b} > 0$ live beyond the regions R_1 or R_2 then we would have four critical points on the x-axis which is not possible.

Now we see how the critical points collide. If $\sqrt{1-2b} \in R_1$ then we got to have $x_{saddle} = x_1$. If $\sqrt{1-2b} \in R_2$ then we got to have $x_{saddle} = x_2$. This is because for $F_x = F_x^{sad} < F_x^{crit}$ there have to be three distinct critical points on the x-axis as argued by Lemma 5.7 and Lemma 5.9. From this information new bounds on the x_k critical points may be derived. The bounds for the critical points on the x-axis written compactly, concisely and definitively for $F_x > 0$ are

$$F_x \leq F_x^{crit} \quad \left\{ \begin{array}{rcccl} \frac{-2}{\sqrt{3}}\sqrt{1+2a} & \leq & x_0 & < & -\sqrt{1+2a} \\ \frac{1}{\sqrt{3}}\sqrt{1+2a} & \leq & x_1 & < & \sqrt{1+2a} \\ 0 & < & x_2 & \leq & \frac{1}{\sqrt{3}}\sqrt{1+2a} \end{array} \right.$$

$$F_x < F_x^{sad} \quad \left\{ \begin{array}{l} x_1 > \sqrt{1-2b} \quad \text{for} \quad \sqrt{1-2b} \in R_1 \\ x_2 < \sqrt{1-2b} \quad \text{for} \quad \sqrt{1-2b} \in R_2 \end{array} \right.$$

$$F_x > F_x^{sad} \quad \left\{ \begin{array}{l} x_1 < \sqrt{1-2b} \quad \text{for} \quad \sqrt{1-2b} \in R_1 \\ x_2 > \sqrt{1-2b} \quad \text{for} \quad \sqrt{1-2b} \in R_2 \end{array} \right.$$

$$F_x = F_x^{sad} \quad \left\{ \begin{array}{l} x_1 = x_{saddle} = \sqrt{1-2b} \\ x_2 = x_{saddle} = \sqrt{1-2b} \end{array} \right. \quad \text{not necessarily } x_1 = x_2$$

$$F_x = F_x^{crit} \quad \left\{ \begin{array}{l} x_0 = \frac{-2}{\sqrt{3}}\sqrt{1+2a} \\ x_1 = x_2 = \frac{1}{\sqrt{3}}\sqrt{1+2a} \end{array} \right.$$

$$F_x > F_x^{crit} \quad \left\{ x_0 < \frac{-2}{\sqrt{3}}\sqrt{1+2a} \right.$$

$$F_x = F_x^{sad} = F_x^{crit} \quad \left\{ x_1 = x_2 = x_{saddle} = \sqrt{1-2b} = \frac{1}{\sqrt{3}}\sqrt{1+2a} \right.$$

Now that the bounds on the critical points for various forces are known, we can deduce their nature. The $(x_{saddle}, \pm y_{saddle})$ is the easiest to prove, taking into account of $x_{saddle}^2 + y_{saddle}^2 = (1-2b)$ we have for the determinant of the Hessian at $(x_{saddle}, \pm y_{saddle})$

$$\det H = -4y_{saddle}^2(a+b) < 0$$

which is definitely a saddle. The Hessian for the critical points $(x_k, 0)$ is already diagonal even with forcing. It is

$$H(x_k, 0) = \begin{pmatrix} \frac{\partial^2 V_F}{\partial x^2} & 0 \\ 0 & \frac{\partial^2 V_F}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 3x_k^2 - (1+2a) & 0 \\ 0 & x_k^2 - (1-2b) \end{pmatrix}$$

and so by using the bounds we derived, and by considering whether the eigenvalues are both positive (well), both negative (hill) or opposite signs (saddle) we can finally deduce the nature of all five critical points. □

The situation can be represented graphically as

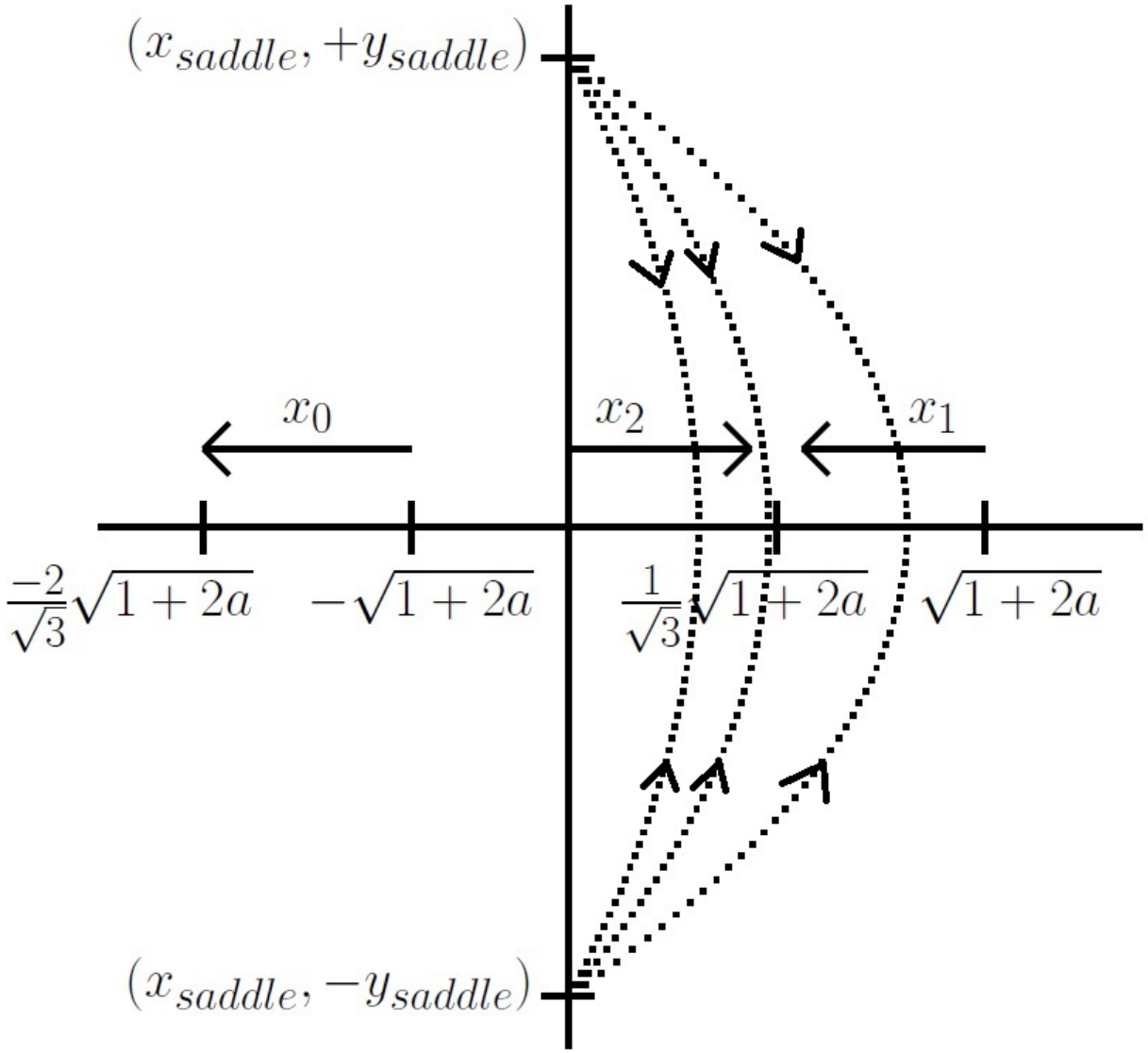

Figure 5.1: As F_x increases from 0 to F_x^{crit}, the $x_{1,2,3}$ move as shown in the diagram. As F_x increases from 0 to F_x^{sad} the $(x_{saddle}, \pm y_{saddle})$ meet each other on the x-axis. There are three possible paths for $(x_{saddle}, \pm y_{saddle})$. If $\sqrt{1-2b} \in R_1$, then the two $(x_{saddle}, \pm y_{saddle})$ would meet in the interval $\left(\frac{1}{\sqrt{3}}\sqrt{1+2a}, \sqrt{1+2a}\right)$ and collide into $(x_2, 0)$. If $\sqrt{1-2b} \in R_2$, then the two $(x_{saddle}, \pm y_{saddle})$ would meet in the interval $\left(0, \frac{1}{\sqrt{3}}\sqrt{1+2a}\right)$ and collide into $(x_1, 0)$. If $\sqrt{1-2b} = \frac{1}{\sqrt{3}}\sqrt{1+2a}$, which is also when $F_x^{sad} = F_x^{crit}$, then the two $(x_{saddle}, \pm y_{saddle})$ would meet at $x = \frac{1}{\sqrt{3}}\sqrt{1+2a}$ and collide simultaneously into $(x_1, 0)$ and $(x_2, 0)$.

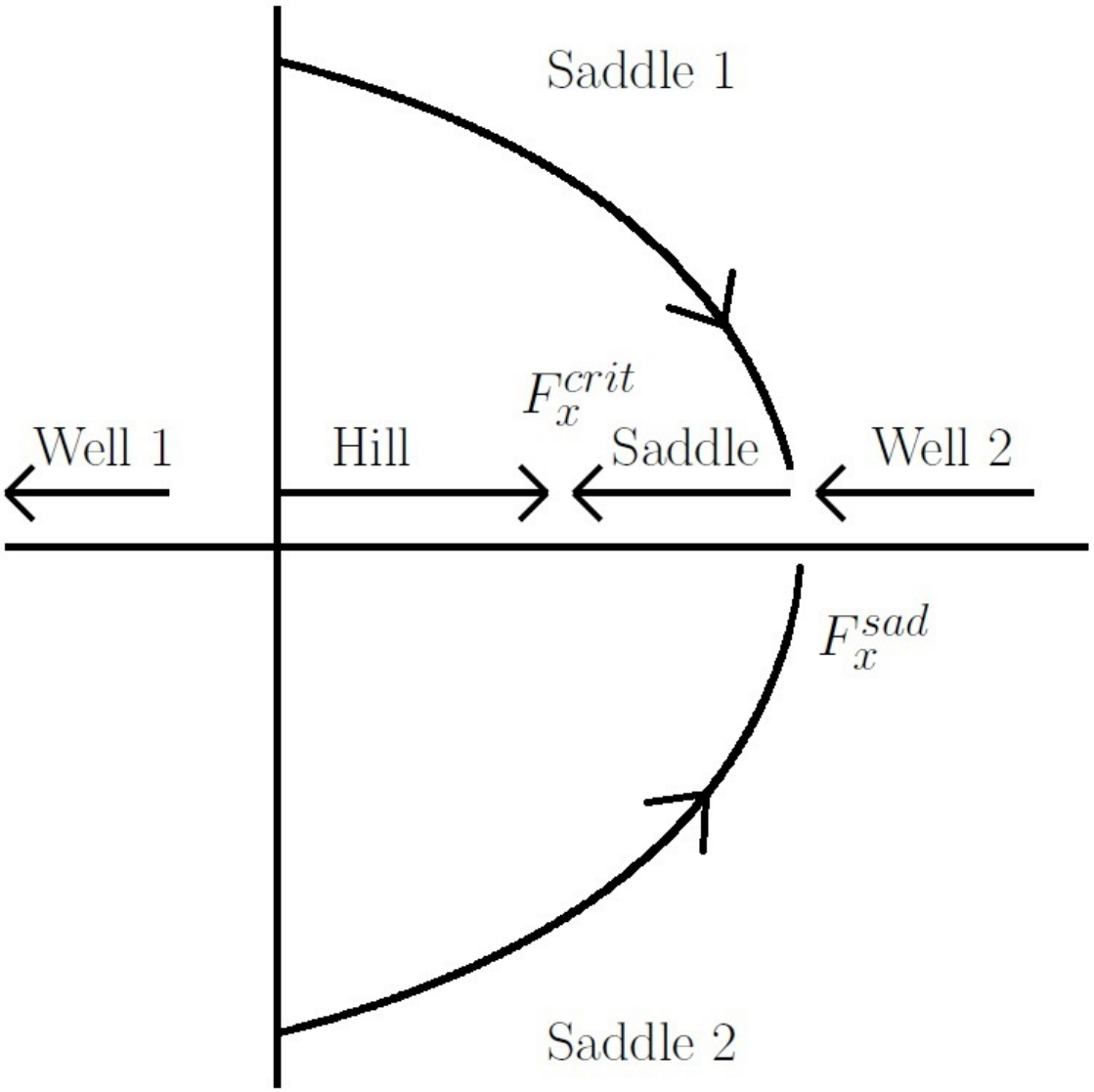

Figure 5.2: This is the case for when $\sqrt{1-2b} \in R_1$. When $F = F_x^{sad}$ the two saddles collide into the right well and turns into a new saddle. At F_x^{crit}, this newly created saddle collides into the hill and both disappears. When Well 2 turns into a Saddle here, it is like creating a new path for the particle to transit to Well 1.

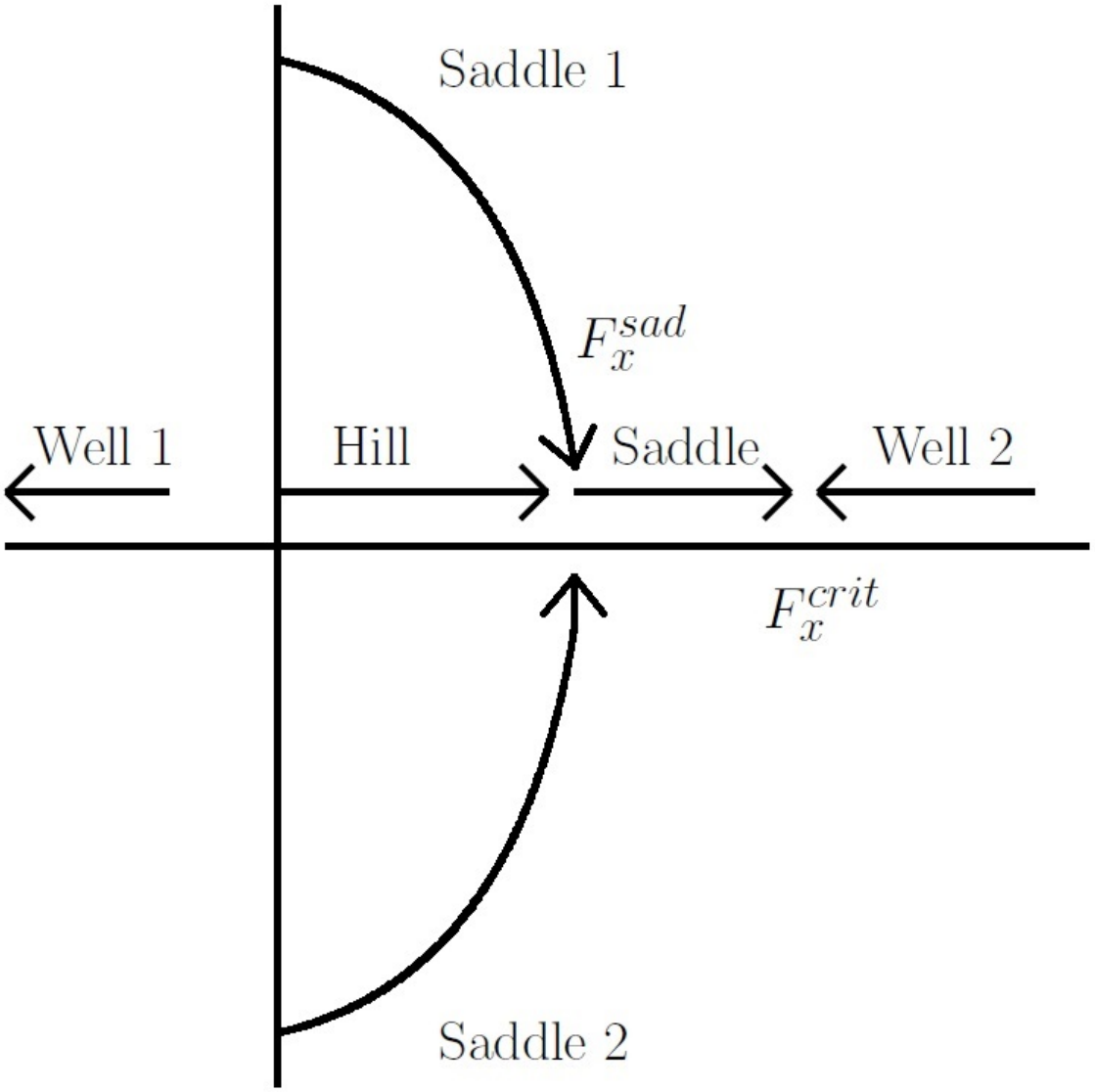

Figure 5.3: This is the case for $\sqrt{1-2b} \in R_2$. When $F = F_x^{sad}$ the two saddles collide into the hill and turns into a new saddle. At F_x^{crit}, this newly created saddle collides into the right well and both disappears. This system behaves in a similar way to a One Dimensional Potential.

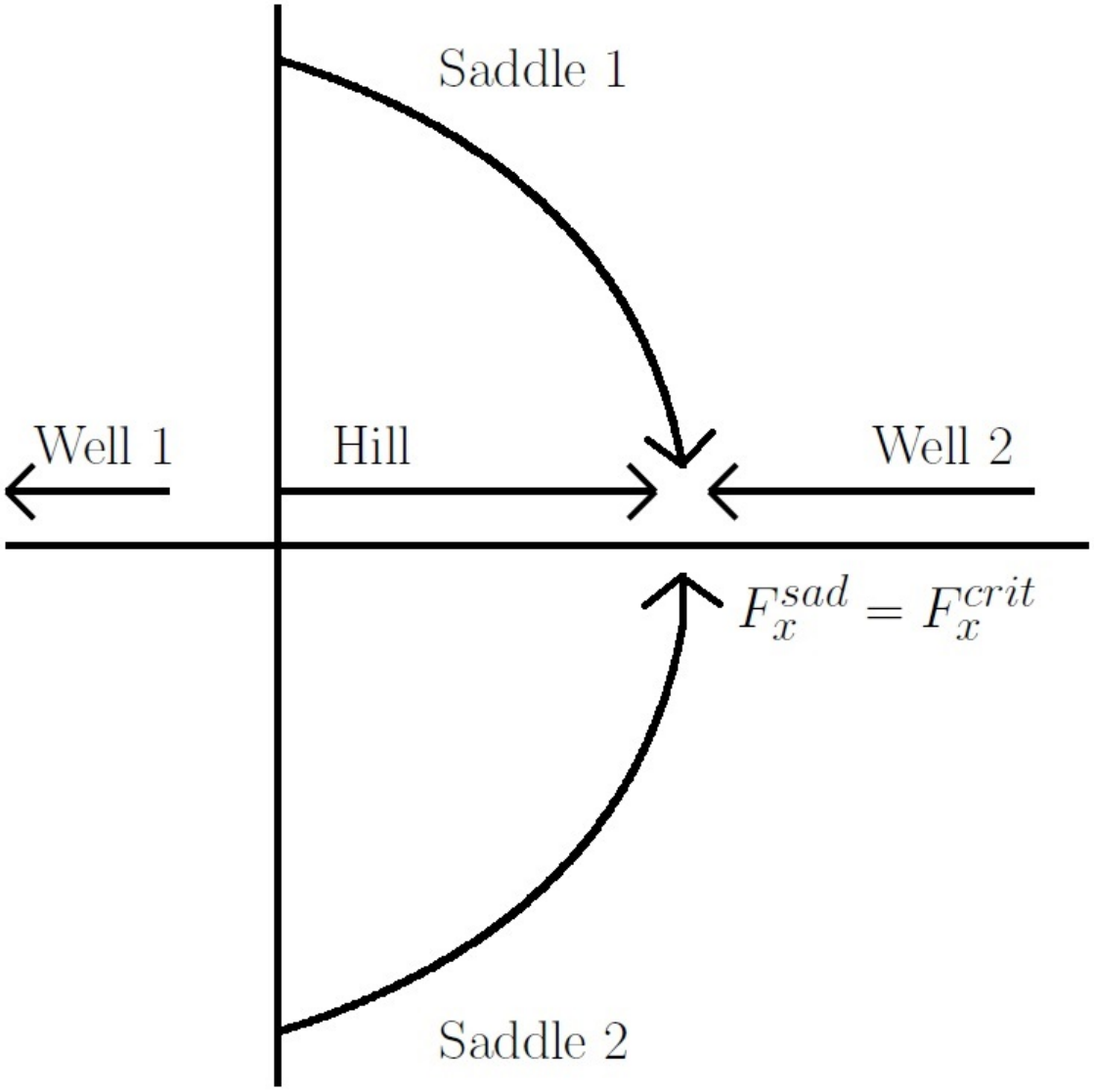

Figure 5.4: This is the case for when $\sqrt{1-2b} = \frac{1}{\sqrt{3}}\sqrt{1+2a}$, which is also when $F_x^{sad} = F_x^{crit}$. At $F = F_x^{sad} = F_x^{crit}$ the two saddles, hill and right well mutually collide at the same place and disappears.

5.2.2 Case $F_x > 0$, $F_y = 0$ and $b \geq \frac{1}{2}$

For $b \geq \frac{1}{2}$ the reasoning is similar to the $b < \frac{1}{2}$ case, but $(x_{saddle}, \pm y_{saddle})$ does not exist. We have the following Theorem.

Theorem 5.11. *Let $F_x > 0$, $F_y = 0$ and $b \geq \frac{1}{2}$. The positions and nature of the critical*

points are as follows

$$F_x < F_x^{crit} \begin{cases} (x_0, 0) & \textit{well} \\ (x_1, 0) & \textit{well} \\ (x_2, 0) & \textit{saddle} \end{cases}$$

$$F_x > F_x^{crit} \begin{cases} (x_0, 0) & \textit{well} \\ (x_1, 0) & \textit{nonexistent} \\ (x_2, 0) & \textit{nonexistent} \end{cases}$$

$$F_x = F_x^{crit} \begin{cases} (x_0, 0) & \textit{well} \\ (x_1, 0) & \textit{neither} \\ (x_2, 0) & \textit{neither} \end{cases}$$

This can be graphically conveyed as

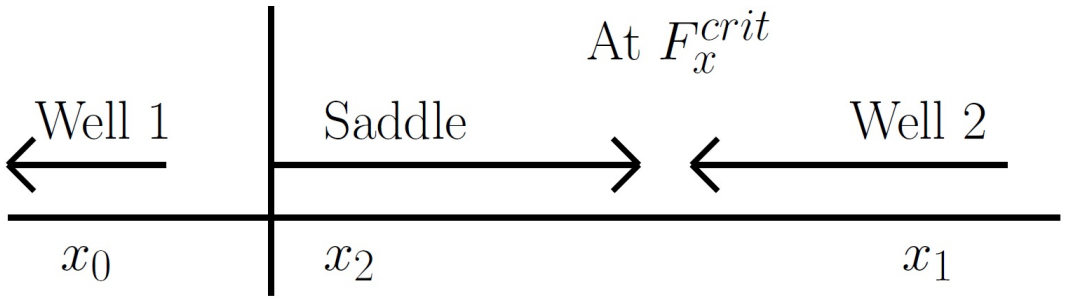

Figure 5.5: As F_x increases from 0 to F_x^{crit}, the saddle collides into the right well and disappears.

5.3 Case $F_x = 0$ and $F_y > 0$

Similarly when forcing is only in the y direction, the cases for $b < \frac{1}{2}$ and $b \geq \frac{1}{2}$ have to be considered separately. The case for $F_y \leq 0$ is similar.

5.3.1 Case $F_x = 0$, $F_y > 0$ and $b < \frac{1}{2}$

We have some Lemmas and Theorems which are almost analogous to the case for $F_x > 0$, $F_y = 0$ and $b < \frac{1}{2}$. Their proofs are very similar and are omitted.

Theorem 5.12. *Let $F_x = 0$, $F > 0$ and $b < \frac{1}{2}$. Let F_y be bounded by*

$$F_y \leq F_y^{sad} \quad \textit{and} \quad F_y \leq F_y^{crit}$$

where

$$F_y^{sad} = 2(a+b)\sqrt{1+2a} \quad \textit{and} \quad F_y^{crit} = \sqrt{\frac{4(1-2b)^3}{27}}$$

then there are five critical points given by

$$\begin{aligned} &(0, y_k) \quad k = 1, 2, 3 \\ &(\pm x_{well}, y_{well}) \end{aligned}$$

where

$$\begin{aligned} x_{well} &= \sqrt{(1+2a) - \left(\frac{F_y}{2(a+b)}\right)^2} \\ y_{well} &= \frac{-F_y}{2(a+b)} \\ y_k &= -\frac{2}{\sqrt{3}}\sqrt{1-2b}\,\cos\left(\psi + k\frac{2\pi}{3}\right) \\ \psi &= \frac{1}{3}\tan^{-1}(p) \\ p &= \frac{\sqrt{4(1-2b)^3 - 27F_y^2}}{F_y\sqrt{27}} \\ k &= 0, \quad k = 1, \quad k = 2. \end{aligned}$$

Lemma 5.13. *The function p (as in Theorem 5.12) is monotone in F_y for $F_y < F_y^{crit}$.*

Lemma 5.14. *Let $F_x = 0$, $F_y > 0$ and $b < \frac{1}{2}$. These three scenarios hold.*
If $F_y < F_y^{crit}$ we have 3 critical points on the x-axis: $y_0 < y_2 < y_1$.
If $F_y = F_y^{crit}$ we have 2 critical points on the y-axis: $y_0 < y_2 = y_1$.
If $F_y > F_y^{crit}$ we have 1 critical point on the y-axis: $y_0 < \frac{-2}{\sqrt{3}}\sqrt{1-2b}$ and $y_0 \to -\infty$ monotonically with increasing F_y.

Using the same reasoning as for the x-direction case we have bounds on the three y_k for $0 \le F_y \le F_y^{crit}$.

$$\begin{aligned} \frac{-2}{\sqrt{3}}\sqrt{1-2b} \le y_0 \le -\sqrt{1-2b} \\ \frac{1}{\sqrt{3}}\sqrt{1-2b} \le y_1 \le \sqrt{1-2b} \\ 0 \le y_2 \le \frac{1}{\sqrt{3}}\sqrt{1-2b}. \end{aligned}$$

The monotonicity of y_k means

$$\begin{array}{lrcll} y_0: & -\sqrt{1-2b} & \longrightarrow & \frac{-2}{\sqrt{3}}\sqrt{1-2b} & \text{as} \quad F_y \to F_y^{crit} \\ y_1: & \sqrt{1-2b} & \longrightarrow & \frac{1}{\sqrt{3}}\sqrt{1-2b} & \text{as} \quad F_y \to F_y^{crit} \\ y_2: & 0 & \longrightarrow & \frac{1}{\sqrt{3}}\sqrt{1-2b} & \text{as} \quad F_y \to F_y^{crit} \end{array}$$

without any oscillations. Now consider the special case when $F_y = F_y^{sad}$. This gives

$$\begin{aligned} (\pm x_{well}, y_{well}) &= (0, y_{well}) \\ &= (0, -\sqrt{1+2a}) \\ \Rightarrow y_{well} &= -\sqrt{1+2a} \end{aligned}$$

which definitely satisfies $-\sqrt{1+2a} < -\sqrt{1-2b}$ for $0 < b < \frac{1}{2}$. This seemingly adds a fourth critical point onto the y-axis. We have a Lemma whose method of proof is similar to Lemma 5.8. But now it does not take long to find numerical examples such that $F_y^{sad} < F_x^{crit}$, $F_y^{crit} < F_y^{sad}$ and $F_y^{sad} = F_y^{crit}$. These would form the separate sub-cases we would have to consider.

Lemma 5.15. *The following statements hold.*

1. *If* $F_y^{sad} < F_y^{crit}$, *then* $-\sqrt{1+2a} > \frac{-2}{\sqrt{3}}\sqrt{1-2b}$
2. *If* $F_y^{sad} > F_y^{crit}$, *then* $-\sqrt{1+2a} < \frac{-2}{\sqrt{3}}\sqrt{1-2b}$
3. *If* $F_y^{sad} = F_y^{crit}$, *then* $-\sqrt{1+2a} = \frac{-2}{\sqrt{3}}\sqrt{1-2b}$

Proof. If $F_y^{sad} < F_y^{crit}$, assume that $-\sqrt{1+2a} \leq \frac{-2}{\sqrt{3}}\sqrt{1-2b}$. Let $F_y = F_y^{sad}$. But this means $F_y < F_y^{crit}$ and by Lemma 5.14 there must be three critical points on the y-axis. But monotonicity implies $y_0 > \frac{-2}{\sqrt{3}}\sqrt{1-2b}$ for $F_y < F_y^{crit}$ meaning there would be 4 critical points on the y-axis, which is a contradiction.

If $F_y^{sad} > F_y^{crit}$ assume, that $-\sqrt{1+2a} \geq \frac{-2}{\sqrt{3}}\sqrt{1-2b}$. Let $F_y = F_y^{sad}$. But this means $F_y > F_y^{crit}$ and Lemma 5.14 implies that there should only be one critical point on the y-axis. We know that $y_0 = \frac{-2}{\sqrt{3}}\sqrt{1-2b}$ at $F_y = F_y^{crit}$ and yet the monotonicity of y_0 means $y_0 < \frac{-2}{\sqrt{3}}\sqrt{1-2b}$ for $F_y > F_y^{crit}$. This would mean 2 critical points on the y-axis which is a contradiction.[23]

If $F_y^{sad} = F_y^{crit}$ then let $F_y = F_y^{sad} = F_y^{crit}$. But Lemma 5.14 says there can only be 2 critical points on the y-axis. But $y_1 = y_2 > 0$ and $y_0 < 0$. This means y_{well} must collide into y_0, hence the statement of the Theorem. □

[23] Just like in the x-direction case it can be shown that for $F_y > F_y^{crit}$, $y_0 = \frac{-2}{\sqrt{3}}\sqrt{1-2b}\,\cosh(z)$ where z is a real number which monotonically decreases with increasing F_y, and yet $z = 0$ when $F_y = F_y^{crit}$, which justifies the idea of this proof.

Again we are ready for another main Theorem of this Chapter. It is finding the positions and nature of all the critical points for different values of F_y.

Theorem 5.16. *Let $F_x = 0$, $F_y > 0$ and $b < \frac{1}{2}$. The positions and nature of the critical points are as follows*

$$F_y < F_y^{crit} \quad \begin{cases} (0, y_1) & \textit{saddle} \\ (0, y_2) & \textit{hill} \end{cases}$$

$$F_y < F_y^{sad} \quad \begin{cases} (0, y_0) & \textit{saddle} \\ (\pm x_{well}, y_{well}) & \textit{well} \end{cases}$$

$$F_y > F_y^{crit} \quad \begin{cases} (0, y_1) & \textit{nonexistent} \\ (0, y_2) & \textit{nonexistent} \end{cases}$$

$$F_y > F_y^{sad} \quad \begin{cases} (0, y_0) & \textit{well} \\ (\pm x_{well}, y_{well}) & \textit{nonexistent} \end{cases}$$

$$F_y = F_y^{crit} \quad \begin{cases} (0, y_1) & \textit{unidentified} \\ (0, y_2) & \textit{unidentified} \end{cases}$$

$$F_y = F_y^{sad} \quad \begin{cases} (0, y_0) & \textit{unidentified} \\ (\pm x_{well}, y_{well}) & \textit{unidentified} \end{cases}$$

Proof. Just like in the x-direction case, monotonicity of p is essential in justifying the following bounds on the critical points for $F_y > 0$. Note that Lemma 5.15 is used to determine the bounds on the y_k critical points

$$F_y \leq F_y^{crit} \quad \begin{cases} \frac{-2}{\sqrt{3}}\sqrt{1-2b} \leq y_0 < -\sqrt{1-2b} \\ \frac{1}{\sqrt{3}}\sqrt{1-2b} \leq y_1 < \sqrt{1-2b} \\ 0 < y_2 \leq \frac{1}{\sqrt{3}}\sqrt{1-2b} \end{cases}$$

$$F_y < F_y^{sad} \quad \{ \; y_0 > -\sqrt{1+2a}$$

$$F_y > F_y^{sad} \quad \{ \; y_0 < -\sqrt{1+2a}$$

$$F_y > F_y^{crit} \quad \{ \; y_0 < \tfrac{-2}{\sqrt{3}}\sqrt{1-2b}$$

$$F_y = F_y^{sad} \quad \{ \; y_0 = y_{well} = -\sqrt{1+2a}$$

$$F_y = F_y^{crit} \quad \{ \; y_1 = y_2 = \tfrac{1}{\sqrt{3}}\sqrt{1-2b}$$

$$F_y = F_y^{sad} = F_y^{crit} \quad \{ \; y_0 = y_{well} = -\sqrt{1+2a} = \tfrac{-2}{\sqrt{3}}\sqrt{1-2b}$$

Now that the bounds on the critical points are found we can determine their nature. Similarly $(\pm x_{well}, y_{well})$ is the one whose nature is easiest to prove. After noting that $x_{well}^2 + y_{well}^2 = (1+2a)$, the determinant of the Hessian at $(\pm x_{well}, y_{well})$ gives

$$\det H = 4x_{well}^2(a+b) > 0$$

and the second partial derivative in x gives

$$\frac{\partial^2 V_F}{\partial x^2}(\pm x_{well}, y_{well}) = 2x_{well}^2 > 0$$

which is a well by Theorem 5.2. The Hessian matrix for the three critical points on the y-axis $(0, y_k)$ is already diagonal even with forcing

$$H(0, y_k) = \begin{pmatrix} \frac{\partial^2 V_F}{\partial x^2} & 0 \\ 0 & \frac{\partial^2 V_F}{\partial y^2} \end{pmatrix} = \begin{pmatrix} y_k^2 - (1 + 2a) & 0 \\ 0 & 3y_k^2 - (1 - 2b) \end{pmatrix}.$$

Lemma 5.15 has to be used in conjunction with the bounds on y_0, y_1, y_3 and y_{well} (as derived in the proof of this Theorem) to determine the nature of the critical points. □

This can be shown graphically.

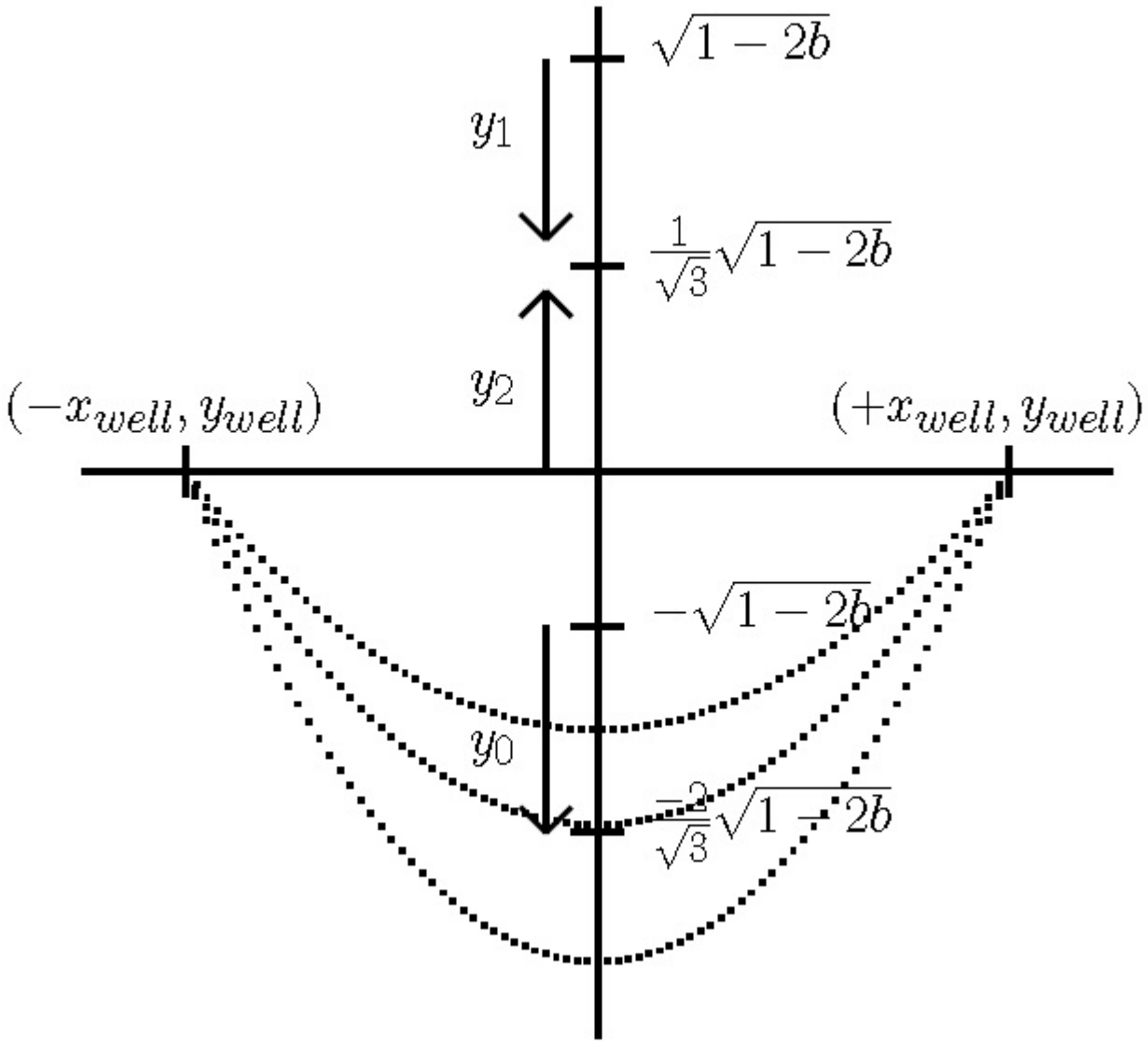

Figure 5.6: As F_y increases from 0 to F_y^{crit} the y_0, y_1 and y_2 move as shown in the diagram. As F_y increases from 0 to F_y^{sad}, the two $(\pm x_{well}, y_{saddle})$ meet each other on the y-axis. There are three possible paths for $(\pm x_{well}, y_{well})$. If $F_y^{sad} < F_y^{crit}$, then the two $(\pm x_{well}, y_{well})$ meet between in the interval $\left(-\frac{2}{\sqrt{3}}\sqrt{1-2b}, -\sqrt{1-2b}\right)$. If $F_y^{sad} = F_y^{crit}$, then the two $(\pm x_{well}, y_{well})$ meet at $y = -\frac{2}{\sqrt{3}}\sqrt{1-2b}$. If $F_y^{sad} > F_y^{crit}$ then the two $(\pm x_{well}, y_{well})$ meet in the interval $\left(-\infty, -\frac{2}{\sqrt{3}}\sqrt{1-2b}\right)$.

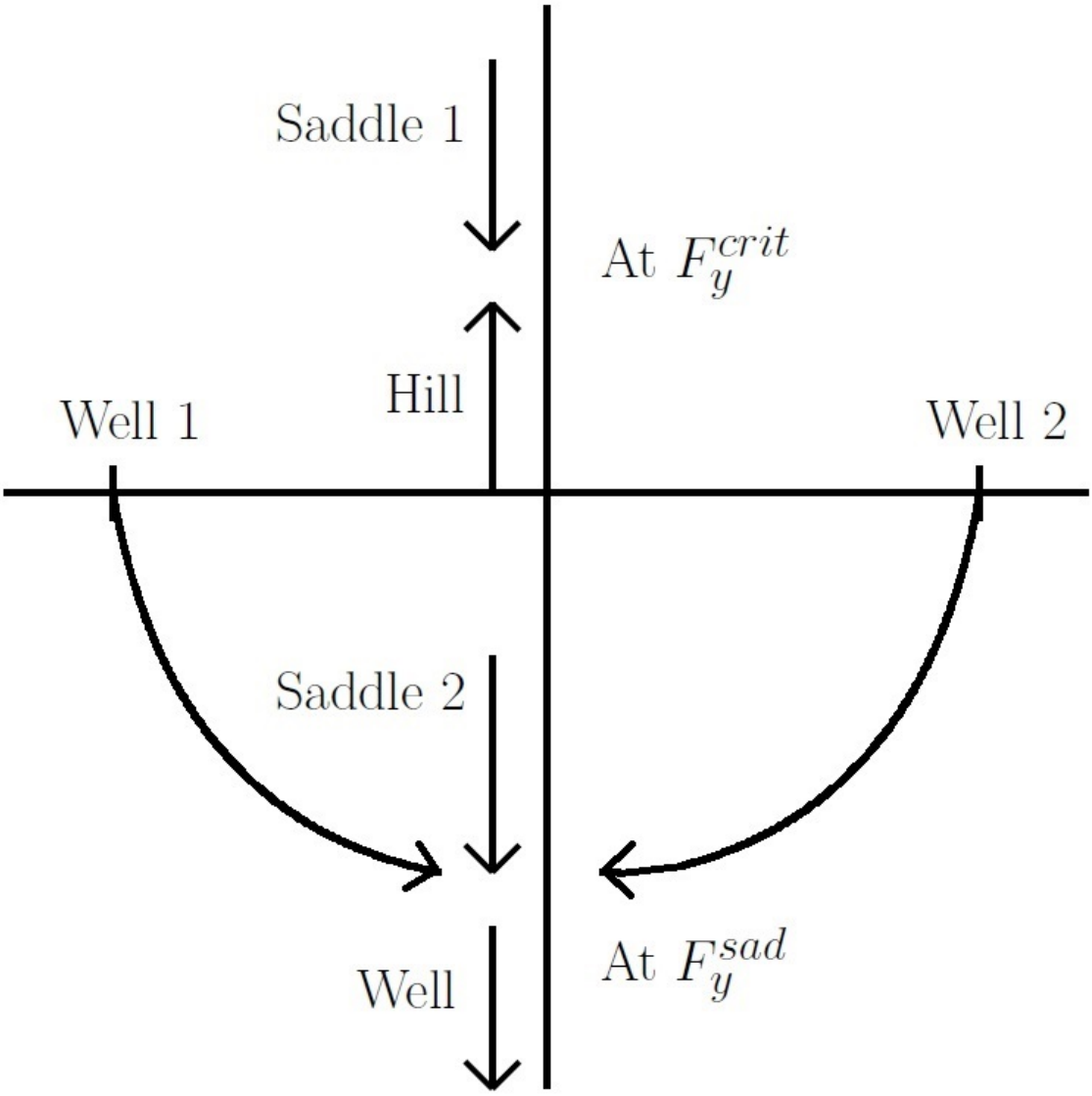

Figure 5.7: At $F = F_y^{crit}$ the top saddle collides into the hill and both then disappears. At $F = F_y^{sad}$ the bottom saddle collides with the two wells and turns into a new well. These two collisions can occur simultaneously or occur one after the other, depending on whether we have $F_y^{sad} < F_y^{crit}$, $F_y^{sad} = F_y^{crit}$ or $F_y^{sad} > F_y^{crit}$.

5.3.2 Case $F_x = 0$, $F_y > 0$ and $b \geq \frac{1}{2}$

The case for $F_y > 0$, $b \geq \frac{1}{2}$ is slightly different in the sense that we have to consider the discriminant of the cubic equation with one real solution for the potential. We have the last main Theorem in this Chapter.

Theorem 5.17. *Let $F_x = 0$, $F_y > 0$ and $b \geq \frac{1}{2}$. The positions and nature of the critical points are as follows*

$$
\begin{aligned}
F_y < F_y^{sad} \quad & \begin{cases} (0, y_0) & saddle \\ (\pm x_{well}, y_{well}) & well \end{cases} \\
F_y = F_y^{sad} \quad & \{ \ (0, y_0) = (\pm x_{well}, y_{well}) \quad unidenitified \\
F_y > F_y^{sad} \quad & \begin{cases} (0, y_0) & well \\ (\pm x_{well}, y_{well}) & nonexistent \end{cases}
\end{aligned}
$$

Proof. The simultaneous equations we have to solve are

$$\frac{\partial V_F}{\partial x} = x(x^2 + y^2) - (1 + 2a)x = 0 \tag{5.3}$$

$$\frac{\partial V_F}{\partial y} = y(x^2 + y^2) - (1 - 2b)y + F_y = 0. \tag{5.4}$$

Equation 5.3 holds if either $x = 0$ or $(x^2+y^2)-(1+2a) = 0$. The case for $(x^2+y^2)-(1+2a) = 0$ gives $(\pm x_{well}, y_{well})$ as a critical point in similar way as before. The case for $x = 0$ reduces Equation 5.4 to

$$y^3 - (1 - 2b)y + F_y = 0.$$

It is this resulting cubic equation which forms the next series of discussions. The required expressions in solving this cubic equation are

$$\Delta_0 = 3(1 - 2b), \quad \Delta_1 = 27F_y, \quad \Delta = 4(1 - 2b)^3 - 27F_y^2.$$

Since $b \geq \frac{1}{2}$ the discriminant of this cubic equation is strictly negative meaning $\Delta < 0$, which means there can only be one real solution. All the solutions whether complex or real are given by

$$
\begin{aligned}
y_k &= -\frac{1}{3}\left(e^{i\psi_k} C + e^{-i\psi_k}\frac{\Delta_0}{C}\right) \quad \text{where} \\
\psi_0 &= 0, \quad \psi_1 = 2\pi/3, \quad \psi_2 = 4\pi/3 \\
C &= \sqrt[3]{\frac{\Delta_1 + \sqrt{-27\Delta}}{2}}.
\end{aligned}
$$

Notice that the $\Delta < 0$, C and Δ_0 are all real numbers. Since there can only be one real y_k solution this has to be y_0, as y_0 would just be a sum of real numbers. This then means y_1 and y_2 would be complex conjugate solutions. Now written explicitly we have

$$C = \sqrt[3]{\frac{27F_y + \sqrt{(27F_y^2 - 4(1 - 2b)^3) \times 27}}{2}}$$

which for $b \geq \frac{1}{2}$ is clearly monotonically increasing in F_y for $F_y > 0$. This is because the $\sqrt{\cdot}$ and $\sqrt[3]{\cdot}$ are both monotone with respect to their own argument. This means we can say

$$\begin{aligned}
\frac{dC}{dF_y} &\geq 0 \\
\frac{dy_0}{dF_y} &= -\frac{1}{3}\left(1 - \frac{\Delta_0}{C^2}\right)\frac{dC}{dF_y} \\
&\leq 0
\end{aligned}$$

because $(-\Delta_0) > 0$ for $b \geq \frac{1}{2}$. This means y_0 would always monotonically decrease with increasing F_y. We also know that at $F_y = F_y^{sad}$ the two wells on the sides become

$$(\pm x_{well}, y_{well}) = (0, -\sqrt{1+2a}).$$

Note that earlier when $x = 0$ was imposed in Equation 5.3 and 5.4 we were reduced with a cubic equation that only admits one real solution in y. This means there can only be one critical point on the y-axis so we must have

$$y_0 = -\sqrt{1+2a}$$

at $F_y = F_y^{sad}$. This justifies the following bounds

$$\begin{aligned}
F_y < F_y^{sad} \quad &\left\{ \; y_0 > -\sqrt{1+2a} \right. \\
F_y = F_y^{sad} \quad &\left\{ \; y_0 = -\sqrt{1+2a} \right. \\
F_y > F_y^{sad} \quad &\left\{ \; y_0 < -\sqrt{1+2a} \right.
\end{aligned}$$

which can be justified by the monotonicity of y_0. This together with the Hessian can allow us to identify the nature of the critical points. □

Again the situation can be represented graphically.

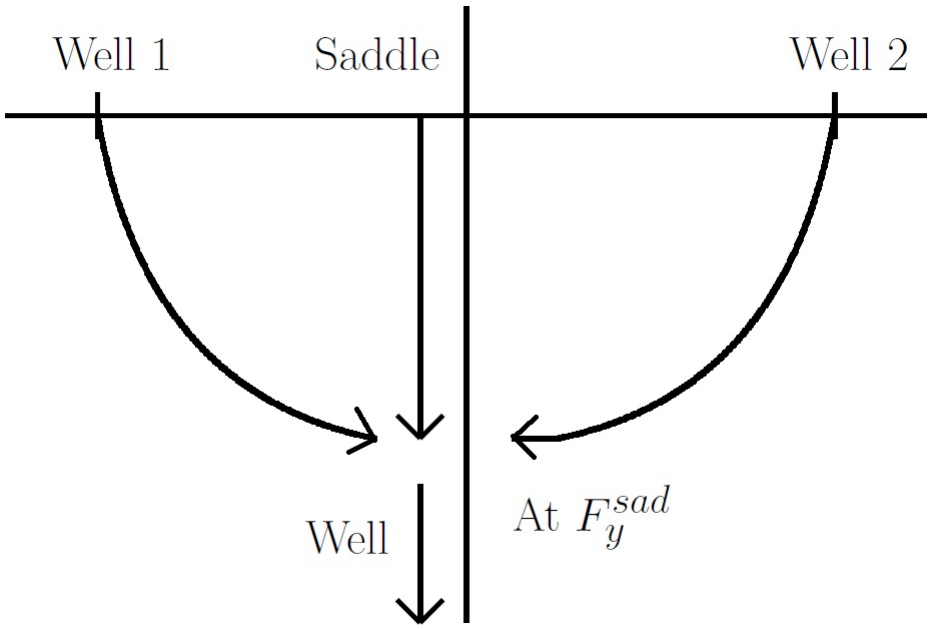

Figure 5.8: At $F_y = F^{sad}$ the saddle collides with the two wells and turns into a new well.

5.4 Case $F_x \neq 0$ and $F_y \neq 0$

Consider the forcing

$$\mathbf{F} = \begin{pmatrix} F_x \\ F_y \end{pmatrix} = \begin{pmatrix} F\cos\phi \\ F\sin\phi \end{pmatrix} \quad \text{where} \quad F = \sqrt{F_x^2 + F_y^2}$$

so far we have only studied the case when $\phi = 0°, 90°, 180°, 360°$. Now we consider the case for forcing in a general direction. The critical points are given by solutions to

$$\frac{\partial V_F}{\partial x} = x(x^2+y^2) - (1+2a)x + F_x = 0 \tag{5.5}$$

$$\frac{\partial V_F}{\partial y} = y(x^2+y^2) - (1-2b)y + F_y = 0. \tag{5.6}$$

The arguments presented next can actually apply for any a, b, F_x and F_y regardless of whether they are positive or negative. Notice that if $\phi \notin \{0°, 90°, 180°, 360°\}$ then $F_x \neq 0$ and $F_y \neq 0$. This means $x = 0$ or $y = 0$ must not appear in any solution. This is the assumption used. Solving Equation 5.6 for x^2 gives

$$y(x^2+y^2) - (1-2b)y + F_y = 0$$

$$(x^2+y^2) - (1-2b) + \frac{F_y}{y} = 0$$

$$(x^2+y^2) = (1-2b) - \frac{F_y}{y}$$
$$x^2 = (1-2b) - \frac{F_y}{y} - y^2 \tag{5.7}$$

and substituting Equation 5.7 into Equation 5.5 gives

$$\sqrt{(1-2b)-\frac{F_y}{y}-y^2}\left((1-2b)-\frac{F_y}{y}\right) - (1+2a)\sqrt{(1-2b)-\frac{F_y}{y}-y^2} + F_x = 0$$

$$\sqrt{(1-2b)-\frac{F_y}{y}-y^2}\left((1-2b)-\frac{F_y}{y}-(1+2a)\right) + F_x = 0$$

$$\left[(1-2b)-\frac{F_y}{y}-y^2\right]\left[-2b-\frac{F_y}{y}-2a\right]^2 - F_x^2 = 0$$

$$\left[(1-2b)-\frac{F_y}{y}-y^2\right]\left[2(a+b)+\frac{F_y}{y}\right]^2 - F_x^2 = 0$$

$$\left[y(1-2b)-F_y-y^3\right]\left[2(a+b)+\frac{F_y}{y}\right]^2 - F_x^2 y = 0$$

$$\left[y(1-2b)-F_y-y^3\right]\left[4(a+b)^2+4(a+b)\frac{F_y}{y}+\frac{F_y^2}{y^2}\right] - F_x^2 y = 0$$

$$\left[y(1-2b)-F_y-y^3\right]\left[4(a+b)^2y^2+4y(a+b)F_y+F_y^2\right] - F_x^2 y^3 = 0$$

which rearranges into a fifth degree polynomial in y

$$\begin{aligned}
&y^5[-4(a+b)^2] \\
&+ y^4[-4(a+b)F_y] \\
&+ y^3[(1-2b)4(a+b)^2 - F_y^2 - F_x^2] \\
&+ y^2[(1-2b)4(a+b)F_y - F_y 4(a+b)^2] \\
&+ y^1[(1-2b)F_y^2 - F_y 4(a+b)F_y] \\
&+ y^0[-F_y^3] \\
&= 0
\end{aligned} \tag{5.8}$$

which can only be solved numerically. The five solutions for Equation 5.8 are denoted by

$$y_i \quad \text{where} \quad i = 1, 2, 3, 4, 5.$$

Now consider Equation 5.5

$$\begin{aligned} x(x^2+y^2)-(1+2a)x+F_x &= 0 \\ x &= \frac{-F_x}{x^2+y^2-(1+2a)} \\ &= \frac{-F_x}{(1-2b)-\frac{F_y}{y}-(1+2a)} \end{aligned}$$

where we have used Equation 5.7 for the last line. This means the x part of the final set of solutions is

$$x_i = \frac{-F_x}{(1-2b)-\frac{F_y}{y_i}-(1+2a)} \quad \text{where} \quad i=1,2,3,4,5.$$

Notice how the calculation is not straightforward if one feeds y_i into Equation 5.7 by means of

$$x_i \neq \sqrt{(1-2b)-\frac{F_y}{y_i}-y_i^2}.$$

We do not know whether to take the positive or negative solution. Any easy counting shows that one obtains too many solutions. These five solutions (x_i, y_i) are sorted into wells, hills and saddles. Although an explicit value for the critical forcing cannot be given analytically, an educated guess can be made

$$F^{crit} = \min\left\{F_x^{sad}, F_x^{crit}, F_y^{sad}, F_y^{crit}\right\} \tag{5.9}$$

that is because a critical force in a general direction must encompass all the other directions. Here are some real examples of the critical points, as the force is being changed during half a period of an oscillatory potential.

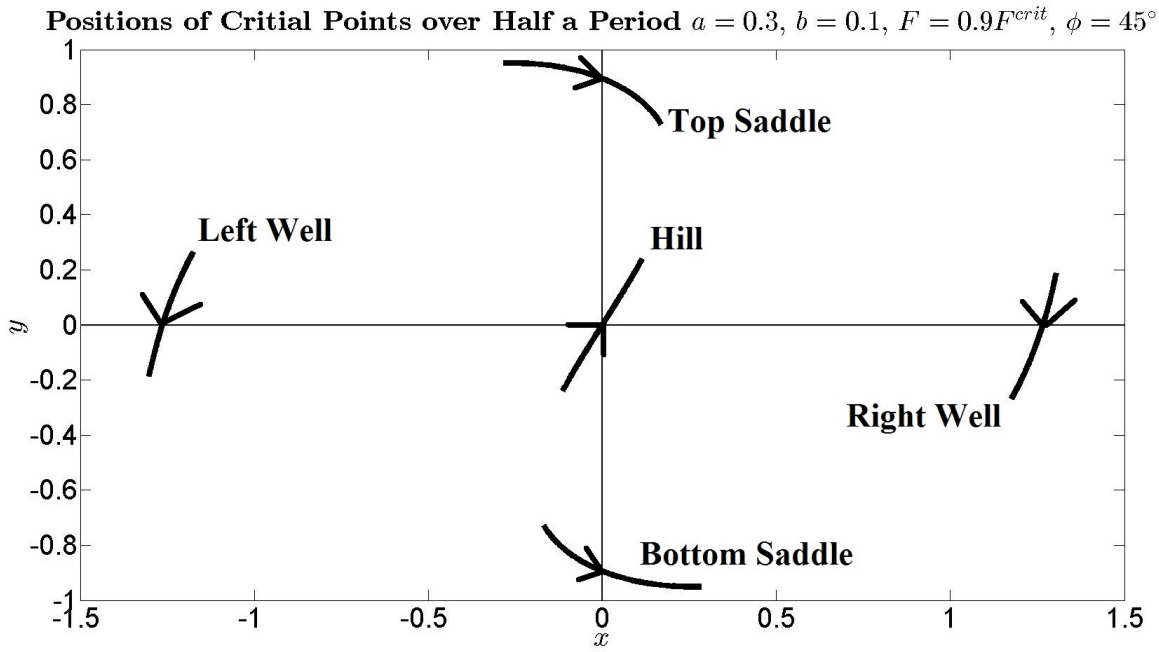

Figure 5.9: The critical points move very little here.

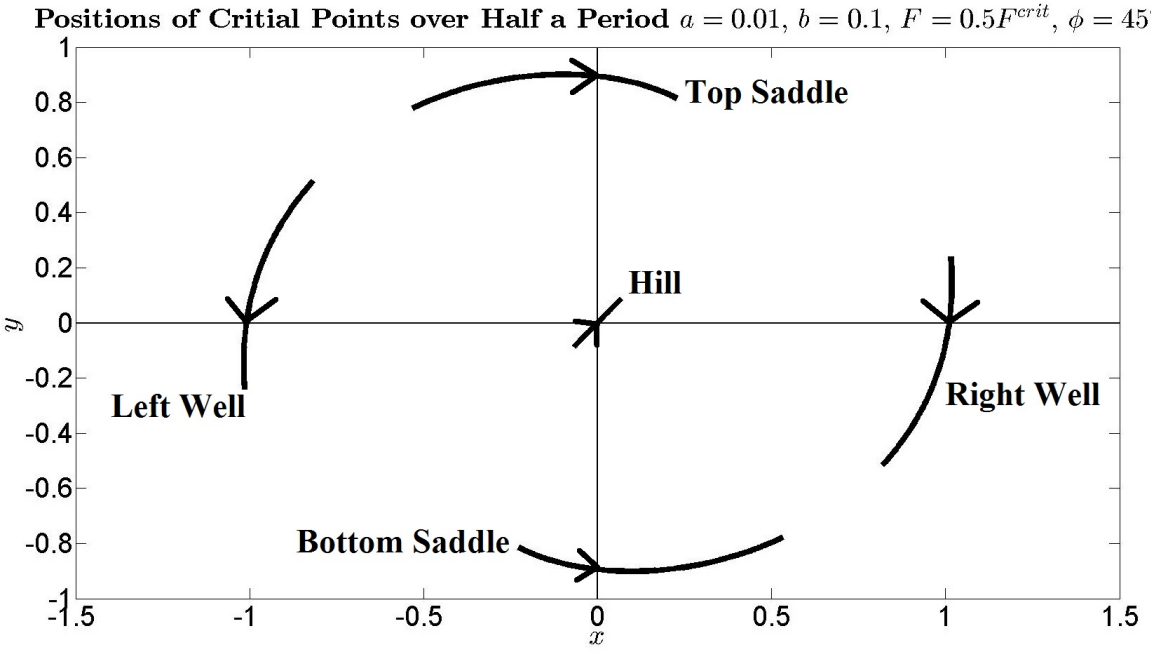

Figure 5.10: The critical points have a more extreme trajectory here.

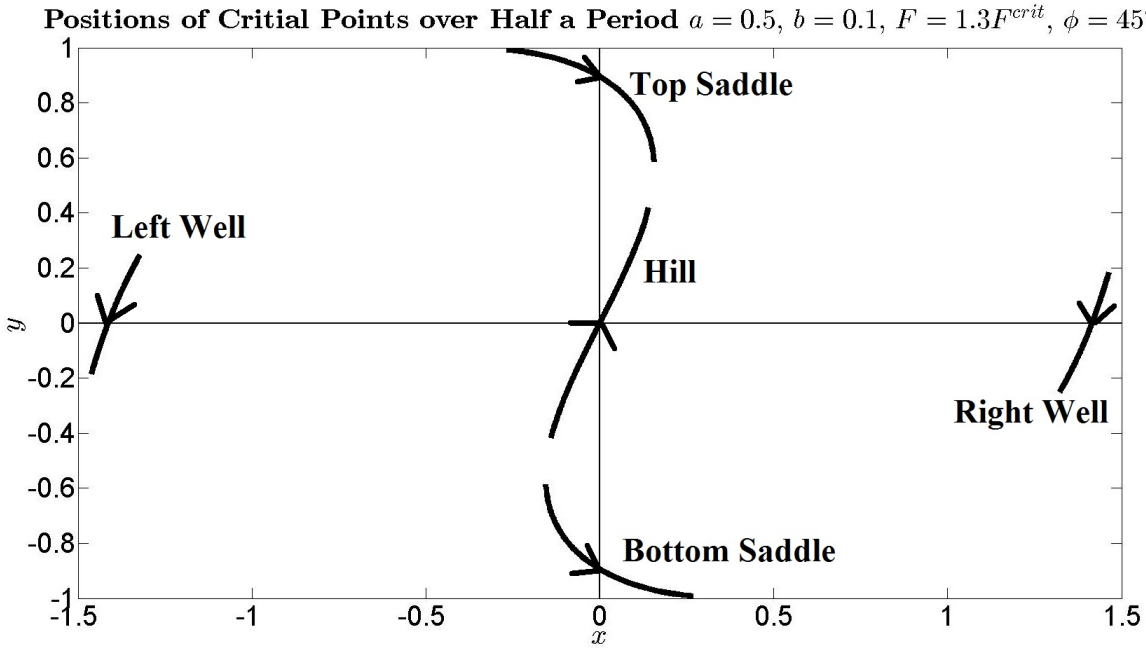

Figure 5.11: Notice that the use of F^{crit} as a critical force is just an educated guess. Here the system is so close to criticality the saddle is almost colliding with the hill.

5.5 Remarks on Mexican Hat

We give an example of how the Mexican Hat Toy Model look like.

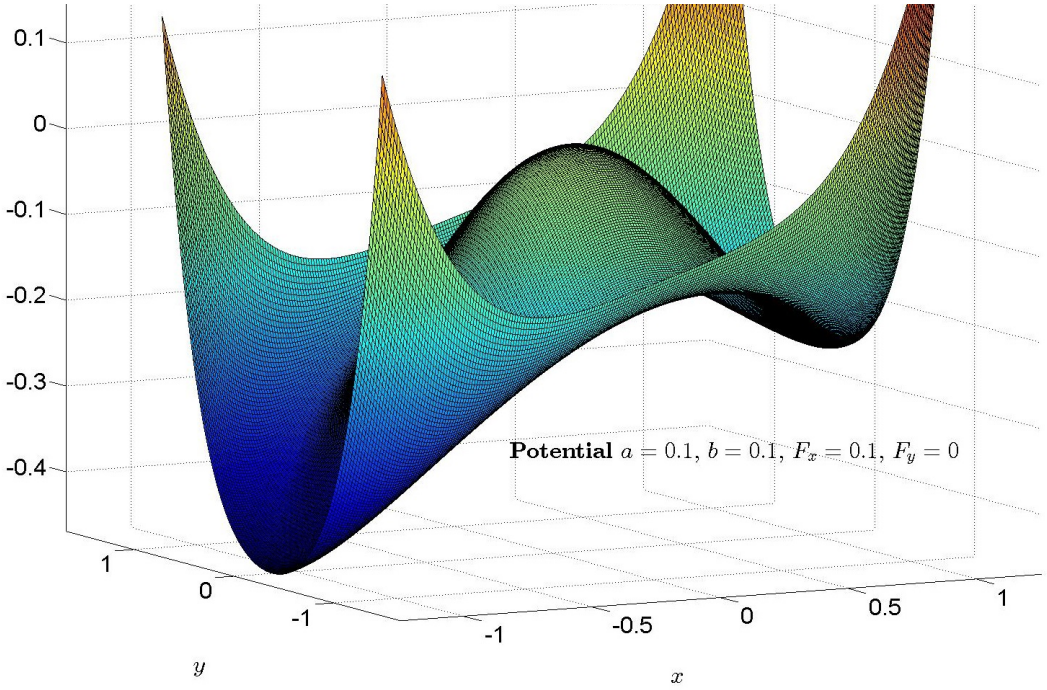

Figure 5.12: An example of the potential $V_F(x, y) = \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 + F_x x + F_y y$ where $r = \sqrt{x^2 + y^2}$. Here $a = 0.1$, $b = 0.1$, $F_x = 0.1$ and $F_y = 0$. Notice there are two saddles just ahead of the hill. The well on the right is higher than the well on the left.

5.5.1 Beyond Criticality

At first glance since the critical points can all be found numerically, one may ask why studied the cubic formula. This actually provided exact analytic information about the system near and beyond criticality and in the extremal cases as well. Besides, stochastic resonance is studied when the forcing is small enough such that the topology of the potential does not change significantly. This is because if the forcing is too large (beyond criticality) then transitions are almost certain, and there is little point to consider stochastic resonance in this case.

5.5.2 Numerical Problems

When the critical points are numerically found, they were fed back into Equations 5.5 and 5.6 and were correct to 10^{-9}. But a few problems remain. In simulations when the angle of the forcing

$$\phi = \tan^{-1}\left(\frac{F_y}{F_x}\right)$$

is changed from $\phi = 0°$ to $\phi = 90°$ the potential was continuously changing from a system needing to solve third order roots to a system needing to solve fifth order roots. This

means the numerical algorithms for solving the quintic polynomial (Equation 5.8) became very unstable when the system is close to solving a cubic equation.[24] No further numerical investigation is necessary as analytic results are available to interpolate the correct solution.

5.5.3 Comparison with One Dimensional Case

The one dimensional potential is

$$V_F = \frac{x^4}{4} - a\frac{x^2}{2} + Fx$$

and their critical points are given by solutions to the equation

$$\frac{\partial V_F}{\partial x} = x^3 - ax + F = 0$$

which when compared to the solutions in the Mexican Hat yields the critical points as

$$x_k = -\frac{2}{3}\sqrt{a}\cos\left\{\frac{1}{3}\tan^{-1}\left(\frac{\sqrt{4a^3 - 27F^2}}{F\sqrt{27}} + \frac{2\pi}{3}k\right)\right\}.$$

Similarly if we want three critical points, the $\tan^{-1}(\cdot)$ must take real arguments, which means

$$F < F^{crit} = \sqrt{\frac{4a^3}{27}}$$

and the nature of the critical points are

$$F < F^{crit} \begin{cases} x_0 & \text{well} \\ x_1 & \text{well} \\ x_2 & \text{hill} \end{cases}$$

$$F > F^{crit} \begin{cases} x_0 & \text{well} \\ x_1 & \text{nonexistent} \\ x_2 & \text{nonexistent} \end{cases}$$

$$F = F^{crit} \begin{cases} x_0 & \text{well} \\ x_1 & \text{unidentified} \\ x_2 & \text{unidentified} \end{cases}$$

which is similar to the $F_x > 0$, $F_y = 0$ and $b \geq \frac{1}{2}$ case. These critical points are bounded by

$$x_0 < x_2 < x_1.$$

[24] Algorithms used in the roots(·) function in MatLab.

This was a calculation not studied in the paper by Benzi et al [4] and other literature. In the paper [4] it was assumed that the forcing is so small the hill is very near to $x_2 = 0$. In [4], the escape times were defined by Equations 1.5 and 1.6, where having reached the hill (which is assumed to be at $x_2 = 0$) is sufficient for escape. Our calculations show that the hill actually moves as the potential oscillate. Taking this into account may give a better approximation of the exit times than in [4].

Chapter 6

Numerical Methods

Let us remind ourselves of the unperturbed potential of the Mexican Hat Toy Model.

$$V_0(x,y) = \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 \quad \text{where} \quad r = \sqrt{x^2+y^2}.$$

The SDE we want to study and simulate is

$$\dot{X}_t^\epsilon = -\nabla V_0 + F\cos\Omega t + \epsilon\,\dot{W}_t$$

where W_t is a two dimensional Wiener process. When this SDE is expressed for the separate x and y components we have

$$dx = \left[-\frac{\partial V_0}{\partial x} + F_x\cos\Omega t\right]dt + \epsilon\; dw_x$$
$$dy = \left[-\frac{\partial V_0}{\partial y} + F_y\cos\Omega t\right]dt + \epsilon\; dw_y$$

where ϵ is the noise level and w_x and w_y are two independent Wiener processes. When this SDE is numerically approximated by the Euler method we have

$$t_n = t_{n-1} + t_{step}$$
$$x_n = x_{n-1} + \left[-\frac{\partial V_0}{\partial x}(x_{n-1},y_{n-1}) + F_x\cos(\Omega t_{n-1})\right]t_{step} + \epsilon\sqrt{t_{step}}\;\xi_x$$
$$y_n = y_{n-1} + \left[-\frac{\partial V_0}{\partial y}(x_{n-1},y_{n-1}) + F_y\cos(\Omega t_{n-1})\right]t_{step} + \epsilon\sqrt{t_{step}}\;\xi_y$$

for the iterative scheme, where ξ_x and ξ_y are two independent normal random variables. The level of precision for the numerics are estimated assuming the Euler method is being used in simulations. More on numerical solutions to SDEs can be found in [61].

When the escape times are measured for an oscillatory potential, a fixed radius R is defined around the well which moves with the well. The two parameters we need to consider are

$$t_{step} \quad \text{and} \quad R$$

and derive appropriate values for them. The particle is defined as having entered a well when it enters the area covered by the radius R around the well. The time difference between entering the first well and the second is measured as the escape time from the first well. The idea behind the considerations is to identify possible sources of error for t_{step} and R and eliminate them. The consequence is that t_{step} is bounded by six bounds on the time step

$$t_{step} \leq \min\{t_1, t_2, t_3, t_4, t_5, t_6\}$$

where the time step t_{step} has to meet six different conditions. Similarly the radius is bounded in the following way

$$R \leq \min\{R_1, R_2\}.$$

Although these theories are not rigorous, it gives a fair idea of the level of precision that is needed. It is assumed in the following that the same precision needed for measuring escape times is also precise enough for studying the six measures of resonance. Although the Mexican Hat is a two dimensional system, part of the numerical theory is derived on a general number of dimensions $\mathbb{R}^r$.

6.1 Basic Conditions - Estimating $t_{step} \leq t_1$, $t_{step} \leq t_2$ and $t_{step} \leq t_3$

When the potential oscillates, one period T is achieved when

$$\Omega T = 2\pi \quad \Rightarrow \quad T = \frac{2\pi}{\Omega}$$

and the time step t_{step} has to be precise enough such that the potential is well presented. Thus it is reasonable to take as the first bound

$$t_1 = \frac{2\pi}{\Omega N_1}$$

where N_1 is an appropriately large number say $N_1 = 1000$. Denote by t_{end} for the end time. This is the time the simulation is being run for. At least 1000 transitions need to be detected, which makes the following a reasonable choice

$$t_{end} = 1000 \times \left(\max_{\substack{w_j(t) \\ 0 \leq t \leq T}} \tau + \min_{\substack{w_j(t) \\ 0 \leq t \leq T}} \tau \right)$$

where τ is the predicted escape time as given by Kramers' formula, $w_j(t)$ where $j = 1, 2, \ldots$ are the positions of the wells at time t. Thus the t_{end} is 1000 times the minimum and

maximum predicted escape times over all wells over one period. Trivially, the time step has to be smaller than the shortest predicted escape time so the second bound is

$$t_2 = \frac{1}{N_2} \min_{\substack{w_j(t) \\ 0 \le t \le T}} \tau$$

where we use $N_2 = 1000$. Most of the time, the particle is near the bottom of the well, then one can approximate the iteration scheme to

$$\begin{aligned} t_n &= t_{n-1} + t_{step} \\ x_n &= x_{n-1} + \epsilon\sqrt{t_{step}}\ \xi_x \\ y_n &= y_{n-1} + \epsilon\sqrt{t_{step}}\ \xi_y \end{aligned}$$

which allows the distance travelled by the particle in one increment, in the time of one time step t_{step} to be given as

$$\Delta z = \sqrt{(x_n - x_{n-1})^2 + (y_n - y_{n-1})^2}.$$

Let the critical points be given by

$$c_1(t),\ c_2(t),\ \ldots$$

If the particle starts at the well and ever reaches a hill or saddle, then transition to the other well is almost certain. Thus travelling from the well to another critical point should be almost impossible in a single time step t_{step}, that is one increment Δz. The length of this forbidden jump is

$$l_1 = \min_{\substack{w_i(t) \neq c_j(t) \\ 0 \le t \le T}} |w_i(t) - c_j(t)|$$

which is the minimal distance from the wells to any other critical points (which are not wells) over one period. But the normal random variables ξ_i are normal distributed in $N(0, 1)$. This means

$$P(\xi_x > r) = \int_r^\infty \frac{e^{-x^2/2}}{\sqrt{2\pi}}\ dx$$

and the joint distribution is given by

$$P\left(\xi_x > r_1,\ \xi_y > r_2\right) = \int_{x=r_1}^\infty \int_{y=r_2}^\infty n(x, y)\, dx\, dy$$

where

$$n(x, y) = \frac{1}{2\pi} \exp\left\{-\frac{1}{2}(x^2 + y^2)\right\}.$$

This gives

$$\begin{aligned} P(\Delta z > l_1) &= P\left(\sqrt{\xi_x^2 + \xi_y^2} > \frac{l_1}{\epsilon\sqrt{t_{step}}}\right) \\ &= \iint_{\left\{(x,y):\sqrt{x^2+y^2} > \frac{l_1}{\epsilon\sqrt{t_{step}}}\right\}} n(x,y)\ dx\ dy \\ &= \int_{\theta=0}^{\theta=2\pi} \int_{\frac{l_1}{\epsilon\sqrt{t_{step}}}}^{\infty} \frac{1}{2\pi} e^{-r^2/2}\ r\ dr\ d\theta \\ &= \exp\left\{-\frac{1}{2}\left(\frac{l_1}{\epsilon\sqrt{t_{step}}}\right)^2\right\} \end{aligned}$$

and the number of increments achieving such a direct jump must be almost zero in one session of the simulation. So

$$\frac{t_{end}}{t_{step}} P(\Delta z > l_1) < N_3$$

where N_3 needs to be smaller than one for example $N_3 = 0.1$. Rearranging the expression to

$$\frac{t_{end}}{t_{step}} \exp\left\{-\frac{1}{2}\left(\frac{l_1}{\epsilon\sqrt{t_{step}}}\right)^2\right\} = N_3$$

and solving for the time step t_{step} we get

$$t_{step} = J(\epsilon, l_1, N_3, t_{end})$$

where J is a function which numerically inverts the expressions to give the time step required. This gives the third bound as

$$t_3 = J(\epsilon, l_1, N_3, t_{end}).$$

6.2 Increment Conditions

The bound t_3 on the time step hinges on finding bounds on the length of a single increment

$$\Delta z = \sqrt{(x_n - x_{n-1})^2 + (y_n - y_{n-1})^2}$$

that is the distance travelled in the time of one time step t_{step}. This is now considered again but in a more general setting in $\mathbb{R}^r$ a general number of dimensions.

6.2.1 Increment Theory - Developing $W(\mathcal{S}, l)$

Let $V_0 : \mathbb{R}^r \longrightarrow \mathbb{R}$ be a real function from $\mathbb{R}^r$ to $\mathbb{R}$. Its gradient with a periodic forcing and noise gives rise to the SDE

$$\dot{X}_t = -\nabla V_0 + F\cos(\Omega t) + \epsilon \dot{w}_t$$

where X_t is a trajectory in $\mathbb{R}^r$, F is the force in $\mathbb{R}^r$ and w_t is a vector of r independent Wiener processes. Thus

$$\begin{aligned} X_t &= (x_1(t), x_2(t), \ldots, x_r(t)) \\ F &= (F_1, F_2, \ldots, F_r) \\ w_t &= (w_1(t), w_2(t), \ldots, w_r(t)). \end{aligned}$$

This trajectory can be numerically approximated with the Euler scheme

$$\begin{aligned} t_{n+1} &= t_n + t_{step} \\ x_i^{n+1} &= x_i^n + \left[-\frac{\partial V_0}{\partial x_i}(x_1^n, x_2^n, \ldots, x_r^n) + F_i\cos(\Omega t_n) \right] t_{step} + \epsilon\sqrt{t_{step}}\ \xi_i \end{aligned} \tag{6.1}$$

where the partial derivative is evaluated at the previous iteration step $(x_1^n, x_2^n, \ldots, x_r^n)$ and ξ_i is a normal random variable. Rearranging Equation 6.1 gives

$$\left(x_i^{n+1} - x_i^n\right) \leq \left[-\frac{\partial V_0}{\partial x_i}(x_1^n, x_2^n, \ldots, x_r^n) + F_i\cos(\Omega t_n) \right] t_{step} + \epsilon\sqrt{t_{step}}\ \xi_i. \tag{6.2}$$

Notice that we can make the following bound

$$\begin{aligned} \left[-\frac{\partial V_0}{\partial x_i}(x_1^n, x_2^n, \ldots, x_r^n) + F_i\cos(\Omega t_n) \right] t_{step} &\leq \left| -\frac{\partial V_0}{\partial x_i}(x_1^n, x_2^n, \ldots, x_r^n) + F_i\cos(\Omega t_n) \right| t_{step} \\ &\leq \left[\left| -\frac{\partial V_0}{\partial x_i}(x_1^n, x_2^n, \ldots, x_r^n) \right| + |F_i\cos(\Omega t_n)| \right] t_{step} \\ &\leq \left[\left| -\frac{\partial V_0}{\partial x_i}(x_1^n, x_2^n, \ldots, x_r^n) \right| + |F_i| \right] t_{step} \\ &\leq \left[\max_{\mathcal{S}} \left| \frac{\partial V_0}{\partial x_i} \right| + |F_i| \right] t_{step} \end{aligned} \tag{6.3}$$

where $\mathcal{S}$ is a set that is large enough such that

$$(x_1^n, x_2^n, \ldots, x_r^n) \in \mathcal{S} \subset \mathbb{R}^n.$$

We define

$$\Delta x_i = \left[\max_{\mathcal{S}} \left| \frac{\partial V_0}{\partial x_i} \right| + |F_i| \right] t_{step} + \epsilon\sqrt{t_{step}}\ \xi_i = A_i + B\xi_i. \tag{6.4}$$

By using the triangle inequality we can bound Equation 6.2 and 6.4 in the following way

$$\left|x_i^{n+1} - x_i^n\right| \leq \left|-\frac{\partial V_0}{\partial x_i}(x_1^n, x_2^n, \ldots, x_r^n) + F_i \cos(\Omega t_n)\right| t_{step} + \left|\epsilon\sqrt{t_{step}}\ \xi_i\right|$$

$$|\Delta x_i| \leq \left[\max_{\mathcal{S}} \left|\frac{\partial V_0}{\partial x_i}\right| + |F_i|\right] t_{step} + \left|\epsilon\sqrt{t_{step}}\ \xi_i\right|$$

and by using 6.3 we know that

$$\left|x_i^{n+1} - x_i^n\right| \leq |\Delta x_i|\,.$$

This means the total increment in the time of one iteration is bounded by

$$\sqrt{\left|x_1^{n+1} - x_1^n\right|^2 + \left|x_2^{n+1} - x_2^n\right|^2 + \ldots + \left|x_r^{n+1} - x_r^n\right|^2} \leq \sqrt{\Delta x_1^2 + \Delta x_2^2 + \ldots + \Delta x_r^2}$$

leading us to define

$$\Delta z = \sqrt{\Delta x_1^2 + \Delta x_2^2 + \ldots + \Delta x_r^2}.$$

Now introduce a new variable

$$\eta_i = \frac{\Delta x_i}{B} \sim N\left(\frac{A_i}{B}, 1\right)$$

which is a normal random variable with mean A_i/B and variance one. Let

$$\eta = \eta_1^2 + \eta_2^2 + \ldots + \eta_r^2 \quad \text{and} \quad \lambda = \frac{1}{B^2}\left(A_1^2 + A_2^2 + \ldots + A_r^2\right)$$

which means η is a sum of the squares of r normal random variables with variance one and λ is the sum of their means. This means η is noncentral chi-squared distributed with r degrees of freedom. Its CDF is

$$P(\eta \leq x) = 1 - Q_{\frac{r}{2}}(\sqrt{\lambda}, \sqrt{x})$$

where Q_m is the Marcum Q-function. Thus the total increment is distributed by

$$\begin{aligned} P(\Delta z > l) &= P(\Delta z^2 > l^2) \\ &= P(\eta B^2 > l^2) \\ &= P\left(\eta > \frac{l^2}{B^2}\right) \\ &= Q_{\frac{r}{2}}\left(\sqrt{\lambda}, \frac{l}{\epsilon\sqrt{t_{step}}}\right) \end{aligned}$$

where l would be the length of a forbidden increment. The number of such forbidden jumps must stay below an appropriate number N_3, that is

$$\frac{t_{end}}{t_{step}} P(\Delta z > l) = N_3 \tag{6.5}$$

and Equation 6.5 has to be numerically inverted to give the required time step t_{step}

$$t_{step} = W(\mathcal{S}, l)$$

where a different region $\mathcal{S}$ and length l fulfilling different criteria is used to calculate a different bound on the time step. Note that we choose $N_3 = 0.1$.

6.2.2 Absence of Large Jumps - Estimating $t_{step} \leq t_4$ and $t_{step} \leq t_5$

Let the set $\mathcal{S}_1$ be given by

$$\mathcal{S}_1 = \left\{ \left(-\sqrt{1+2a}, 0\right), \left(+\sqrt{1+2a}, 0\right), \left(0, -\sqrt{1-2b}\right), \left(0, +\sqrt{1-2b}\right), (0,0) \right\}$$

which is the positions of all the critical points as they would be when the forcing is zero $F = 0$. This means

$$\max_{\mathcal{S}_1} \left| \frac{\partial V_0}{\partial x} \right| = 0 \quad \text{and} \quad \max_{\mathcal{S}_1} \left| \frac{\partial V_0}{\partial y} \right| = 0.$$

The set $\mathcal{S}_1$ can be used as an approximation for small forcing. Equation 6.4 now becomes

$$\begin{aligned} \Delta x &= |F_x|\, t_{step} + \epsilon \sqrt{t_{step}}\, \xi_x \\ \Delta y &= |F_y|\, t_{step} + \epsilon \sqrt{t_{step}}\, \xi_y \end{aligned}$$

and we want a time step t_{step} small enough such that almost every value of $\Delta z = \sqrt{\Delta x^2 + \Delta y^2}$ is bounded by

$$\Delta z = \sqrt{\Delta x^2 + \Delta y^2} \leq l_1$$

and this time step is given by t_4 below

$$t_4 = W(\mathcal{S}_1, l_1).$$

Now we remind ourselves that if ζ is an exponentially distributed random variable, its PDF, CDF, mean and variance are given by

$$\begin{aligned} P(\zeta \in A) &= \int_A \lambda e^{-\lambda x}\, dx \\ P(\zeta \leq x) &= \int_{-\infty}^{x} \lambda e^{-\lambda x}\, dx = 1 - e^{-\lambda x} = F(x) \\ \langle \zeta \rangle &= \frac{1}{\lambda} \\ \text{var}(\zeta) &= \frac{1}{\lambda^2} \end{aligned}$$

where λ is the parameter associated with the exponential distribution. Now define the height

$$h_0 = \max_{\substack{w_j(t) \\ 0 \le t \le T}} V_t(w_j(t))$$

which is the maximum height any well can ever reach. Now define the expression

$$\Delta V_h = h_1 - h_0$$

where h_1 is chosen so high that the particle will probably never reach there. Even if it starts from the highest possible well the chances are still very slim. Now we try to estimate what this height h_1 may be. From Freidlin-Wentzell we know that the escape time from $V = h_0$ to $V = h_1$ is roughly

$$\tau_h \approx e^{2\Delta V_h/\epsilon^2}$$

and we want the time it takes to reach $V = h_1$ to be significantly more than the duration of the simulation

$$\tau_h \gg t_{end}$$

which means a reasonable estimate would be

$$\begin{aligned} N_2 t_{end} &= e^{2\Delta V_h/\epsilon^2} \\ 2\Delta V_h &= \epsilon^2 \ln(N_2 t_{end}) \\ \Rightarrow h_1 &= \frac{1}{2}\epsilon^2 \ln(N_2 t_{end}) + h_0 \end{aligned}$$

where as before we choose $N_2 = 1000$. The escape times leaving $V = h_0$ and arriving at $V = h_1$ is exponentially distributed. The average of them would be

$$\langle \tau_h \rangle = N_2 t_{end} = e^{2\Delta V_h/\epsilon^2}$$

and so by using the CDF of the exponential distribution we can say

$$\begin{aligned} P(\tau_h < t_{end}) &= 1 - e^{-1/N_2} \\ &= 0.632 \quad \text{for} \quad N_2 = 1 \\ &= 0.095 \quad \text{for} \quad N_2 = 10 \\ &= 0.01 \quad \text{for} \quad N_2 = 100 \\ &= 0.001 \quad \text{for} \quad N_2 = 1000. \end{aligned}$$

So if we choose $N_2 \geq 1000$ the chances of reaching $V = h_1$ are less then one in a thousand. Define the set

$$\mathcal{S}_2 = \{x \in \mathbb{R}^r : [V_0(x) - Fx\cos(\Omega t)] \leq h_1 : 0 \leq t \leq T\}$$

$$= \{x \in \mathbb{R}^r : V_t \leq h_1 : 0 \leq t \leq T\}$$

which is the set of all the points below $V_t \leq h_1$ over the time of one period. The forbidden increment is taken as the same as last time

$$l_2 = l_1$$

so the second bound is

$$t_5 = W(\mathcal{S}_2, l_2).$$

6.3 Stability and Radius Conditions

Stability of the trajectory in the context of this thesis is for the time step t_{step} to be small enough such that the simulated discrete trajectory is a good enough approximation of a physical continuous trajectory. For example consider the following trajectories for a particle falling down to the well of the Mexican Hat starting at $(x_{start}, y_{start}) = (-0.75, -0.75)$.

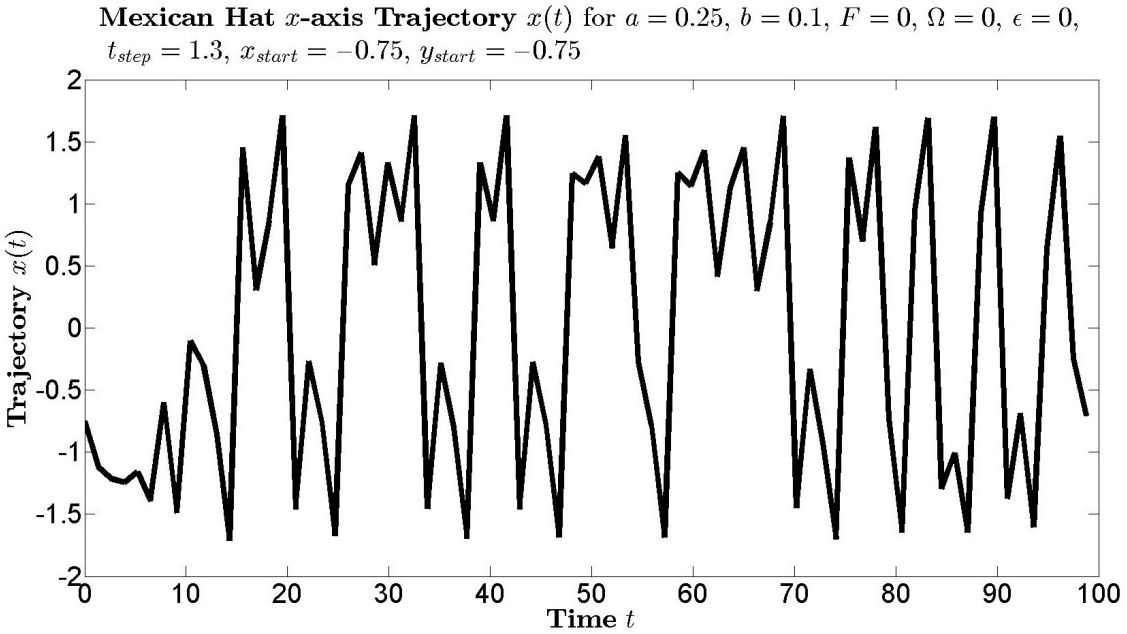

Figure 6.1: The trajectory is so unstable the particle even transits to the other well.

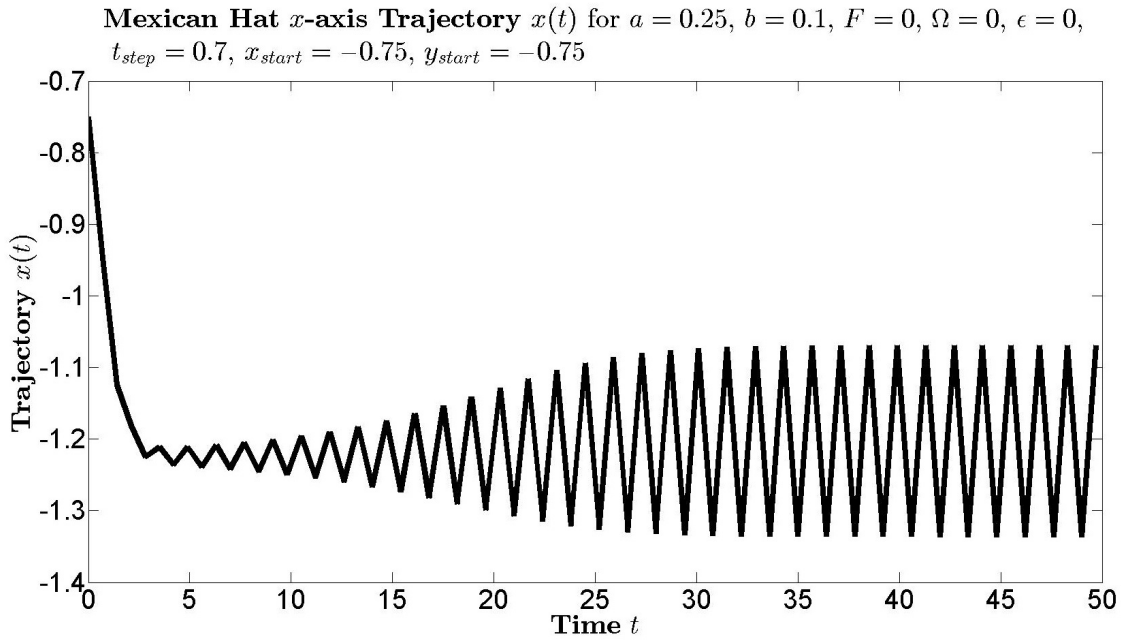

Figure 6.2: The trajectory is more stable but the particle now oscillates near the well.

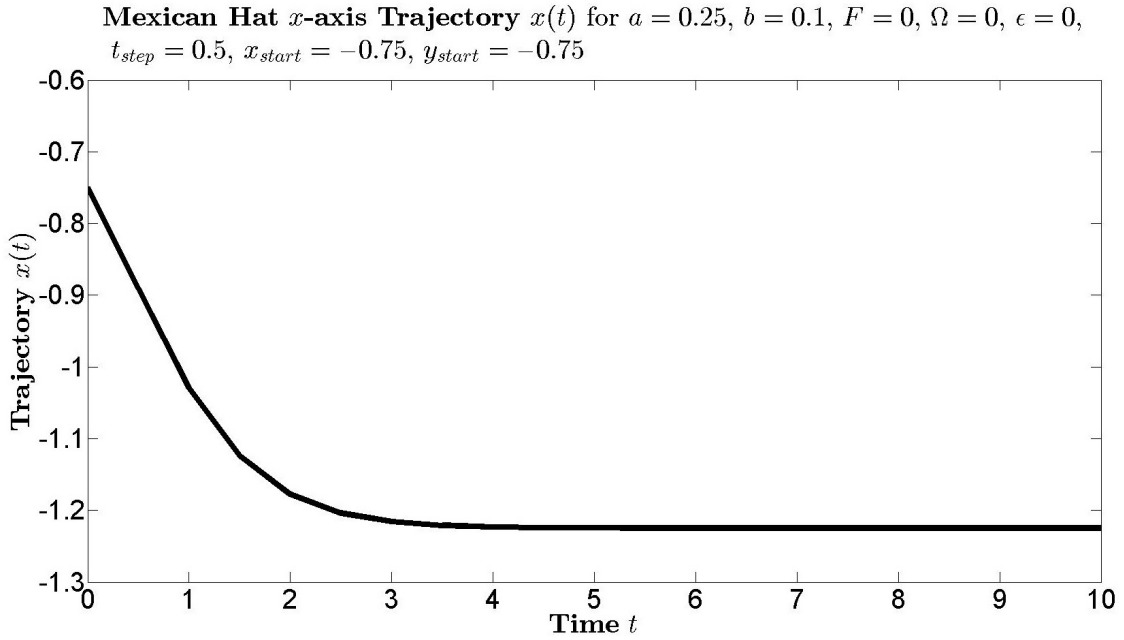

Figure 6.3: The trajectory is sufficiently stable here.

It is the aim of this section to study these stability problems. The Euler method is effectively a discrete iterative map with some operations. We have the following.

Lemma 6.1. *Let $(X, \|\cdot\|)$ be a vector space endowed with the norm $\|\cdot\|$ over the complex scalar field $\mathbb{F}$. Let M be an operator $M : X \longrightarrow X$ with the property $\|Mx\| \leq \|M\|\|x\|$ and $\|M\| \geq 0$. Let $x \in X$ be part of an iterative scheme*

$$x_n = Mx_{n-1} + \epsilon\sqrt{t}\,\xi_{n-1}$$

where $\xi_{n-1} \in X$ is a term which depends on the iterative step n. The entire term x_n is then bounded by

$$\|x_n\| \leq \|M\|^n\|x_0\| + \epsilon\sqrt{t}\sum_{i=0}^{n-1}\|M\|^i\ \|\xi_{n-1-i}\|$$

where x_0 and ξ_0 are the starting (first) steps.

Proof. Rewrite the iterative scheme with a new operator

$$\begin{aligned} x_n &= Mx_{n-1} + \epsilon\sqrt{t}\,\xi_{n-1} \\ &= M'_n x_{n-1} \end{aligned}$$

where M'_n is the total operator which depends on the nth step. In general M'_n is not commutative so we write

$$x_n = M'_n M'_{n-1} \dots M'_1 x_0$$

and consider just one operation on M'_n

$$\begin{aligned} M'_n x &= Mx + \epsilon\sqrt{t}\,\xi_{n-1} \\ \|M'_n x\| &= \left\| Mx + \epsilon\sqrt{t}\,\xi_{n-1} \right\| \\ &\leq \|Mx\| + \epsilon\sqrt{t}\ \|\xi_{n-1}\| \\ &\leq \|M\|\,\|x\| + \epsilon\sqrt{t}\ \|\xi_{n-1}\|\,. \end{aligned}$$

So we have an iterative expression

$$\|M'_n x\| \leq \|M\|\,\|x\| + \epsilon\sqrt{t}\ \|\xi_{n-1}\| \tag{6.6}$$

and iterating Equation 6.6 gives

$$\begin{aligned} \|x_n\| &= \left\| M'_n M'_{n-1} \dots M'_1 x_0 \right\| \\ &\leq \|M\|^n\|x_0\| + \epsilon\sqrt{t}\sum_{i=0}^{n-1}\|M\|^i\ \|\xi_{n-1-i}\| \end{aligned}$$

which completes the proof. □

Remark 6.2. *There are a lot of remarks to say about this simple Lemma.*

1. *Nowhere was it assumed M is bounded, linear or commutative.*

2. *Notice that $\|M\|$ was not defined. It could be the usual operator norm or something else.*

3. *Notice that we are only truncating the bound as in Equation 6.6 and NOT truncating the original iterative Euler scheme.*

4. *Notice that $M0 = 0$ for the zero vector. Notice also that $M(\lambda x) = \lambda M(x)$ where λ is a scalar was not assumed.*

5. *Notice that a norm on $\|M_n'\|$ was not needed. This is because such a norm would have to satisfy $\|M_n' x\| \leq \|M_n'\| \, \|x\|$ which implies $M_n' 0 = 0$. But with the random vector ξ_{n-1} in M_n' this cannot be achieved.*

Consider the two dimensional Mexican Hat system with $z = (x, y) \in \mathbb{R}^2$. Let the potential be stationary by freezing it as it would be at a certain fixed point in time $t = t_{fix}$. The gradient then becomes

$$\nabla V_{t=t_{fix}}(z) = \begin{pmatrix} \frac{\partial V_0}{\partial x} - F_x \cos \Omega t_{fix} \\ \frac{\partial V_0}{\partial y} - F_y \cos \Omega t_{fix} \end{pmatrix}$$

where t_{fix} is a constant point in time. The Euler method can now be rewritten as

$$\begin{aligned} t_n &= t_{n-1} + t_{step} \\ x_n &= x_{n-1} + \left[-\frac{\partial V_0}{\partial x} + F_x \cos \Omega t_{fix} \right] t_{step} + \epsilon \sqrt{t_{step}} \, \xi_x \\ y_n &= y_{n-1} + \left[-\frac{\partial V_0}{\partial y} + F_y \cos \Omega t_{fix} \right] t_{step} + \epsilon \sqrt{t_{step}} \, \xi_y \end{aligned}$$

and we recast it into vector notation by writing

$$\begin{aligned} t_n &= t_{n-1} + t_{step} \\ z_n &= z_{n-1} - t_{step} \nabla V_{t=t_{fix}}(z_{n-1}) + \epsilon \sqrt{t_{step}} \, \xi_{n-1} \end{aligned} \tag{6.7}$$

where

$$z_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}.$$

Now we apply Lemma 6.1 to Equation 6.7. Clearly the operator as mentioned in Theorem 6.1 is

$$M(z) = z - t_{step} \nabla V_{t=t_{fix}}(z)$$

and this operator has to satisfy $\|Mz\| \leq \|M\| \, \|x\|$. The usual operator norm would suffice with some restrictions

$$\|M\|_{\mathcal{S}} = \sup_{\substack{z \in \mathcal{S} \\ z \neq 0}} \frac{\|Mz\|}{\|z\|}$$

where $\mathcal{S} \subset \mathbb{R}^2$ is a strict and suitable subset of the whole space. This is because in general, the expression $\|Mz\|$ is unbounded on $\mathbb{R}^2$. For example in our case ∇V is unbounded. This would give

$$\|Mz\| = \|z\| \frac{\|Mz\|}{\|z\|} \leq \|z\| \sup_{\substack{z \in \mathcal{S} \\ z \neq 0}} \frac{\|Mz\|}{\|z\|} \leq \|M\|_{\mathcal{S}} \, \|z\| \, .$$

But we also need the condition $M0 = 0$ for the zero vector (see Remark 6.2). This means we have to shift the coordinates to

$$z_{new} = z_{old} - z_{well}$$

where the well is now the new origin. Now we can apply Lemma 6.1 to Equation 6.7 and get

$$\|z_n\| \leq \|M\|_{\mathcal{S}}^n \, \|z_0\| + \epsilon \sqrt{t_{step}} \sum_{i=0}^{n-1} \|M\|_{\mathcal{S}}^i \, \|\xi_{n-1-i}\|$$

where the starting position is

$$z_0 = (x_{start}, y_{start}) - (x_{well}(t_{fix}), y_{well}(t_{fix}))$$

where (x_{start}, y_{start}) is the starting position and $(x_{well}(t_{fix}), y_{well}(t_{fix}))$ is the position of one well at time $t = t_{fix}$

6.3.1 Stability Problems

Lemma 6.1 has allowed us to rewrite the Euler method, expressed in Equation 6.7, as

$$\begin{aligned} |z_n| &= \left| M(z_{n-1}) + \epsilon \sqrt{t_{step}} \, \xi_{n-1} \right| \\ &\leq \|M\|_{\mathcal{S}}^n \, \|z_0\| + \epsilon \sqrt{t_{step}} \sum_{i=0}^{n-1} \|M\|_{\mathcal{S}}^i \, \|\xi_{n-1-i}\| \end{aligned} \tag{6.8}$$

where if $\epsilon = 0$ we would reduce back to the deterministic system

$$|z_n| \leq \|M\|_{\mathcal{S}}^n \, \|z_0\| \, .$$

It is known that if the time step t_{step} is too large then even the deterministic trajectory is unstable. A stable time step t_{step} is one which gives

$$\|M\|_{\mathcal{S}} \leq (1 - \delta) \tag{6.9}$$

where $0 < \delta < 1$ and the particle would settle at the bottom of the well as $n \longrightarrow \infty$. There are two approaches to the problem here.

1. Fix the set $\mathcal{S}$ and solve for the time step t_{step} such that $\|M\|_{\mathcal{S}} \leq (1-\delta)$ holds. This is solving for time.

2. Fix the time step t_{step} and solve for the set $\mathcal{S}$ such that $\|M\|_{\mathcal{S}} \leq (1-\delta)$ holds. This is solving for space.

and each approach hinges on the assumption that the $\mathcal{S}$ and t_{step} we fixed to begin with is a good and stable choice. We can only solve for time or space but not both. This is also complicated by the fact that the following estimate is too rough

$$\frac{\|Mz\|}{\|z\|} \leq \frac{\left\|z - t_{step}\nabla V_{t=t_{fix}}(z)\right\|}{\|z\|} \leq \frac{\|z\| + \left\|t_{step}\nabla V_{t=t_{fix}}(z)\right\|}{\|z\|} \leq 1+\delta' \tag{6.10}$$

where $\delta' > 0$.

6.3.2 Estimating $R \leq R_1$ and $R \leq R_2$

Now two bounds on the radius R is derived. The first bound is

$$R_1 = \min_{\substack{w_j(t), c_i(t) \\ w_j \neq c_i \\ 0 \leq t \leq T}} \left\{ \frac{1}{2} \left| w_j(t) - c_i(t) \right| \right\}$$

which is half the distance from the wells to all critical points, for all wells, over one period. This is such that there is always exactly one critical point inside the region covered by the radius R. The second bound comes from considering Equation 6.8

$$z_n \leq \|M\|_{\mathcal{S}}^n \ \|z_0\| + \epsilon\sqrt{t_{step}} \sum_{i=0}^{n-1} \|M\|_{\mathcal{S}}^i \ \ \|\xi_{n-1-i}\|$$

and seek a bound on the variance of the random part. The variance of the random part is

$$\text{var}\left(\epsilon\sqrt{t_{step}} \sum_{i=0}^{n-1} \|M\|_{\mathcal{S}}^i \ \ \|\xi_{n-1-i}\| \right) = \epsilon^2 t_{step} \sum_{i=0}^{n-1} \|M\|_{\mathcal{S}}^{2i} \ \ \text{var}\left(\|\xi_{n-1-i}\| \right).$$

Notice how each of the $\|\xi_{n-1-i}\|$ is χ distributed with $r = 2$ degrees of freedom. Its variance is [25]

$$\begin{aligned} \sigma^2 &= \text{var}\left(\|\xi_{n-1-i}\| \right) \\ &= r - \left(\sqrt{2} \frac{\Gamma\left(\frac{r+1}{2}\right)}{\Gamma\left(\frac{r}{2}\right)} \right)^2 \end{aligned}$$

[25] Let ξ_x and ξ_y be independently and normally distributed in $N(0,1)$. Let $\zeta = \sqrt{\xi_x^2 + \xi_y^2}$ and $\eta = \xi_x^2 + \xi_y^2$, then ζ is χ distributed and η is χ^2 distributed. Notice that we want the variance of ζ and NOT the variance of η, therefore only the χ distribution is needed.

$$= 0.4292 \quad \text{for} \quad r = 2$$

which means the variance of the random part is bounded by

$$\begin{aligned}
\text{var}\left(\epsilon\sqrt{t_{step}}\sum_{i=0}^{n-1}\|M\|_{\mathcal{S}}^{i}\ \|\xi_{n-1-i}\|\right) &= \epsilon^2 t_{step}\sum_{i=0}^{n-1}\|M\|_{\mathcal{S}}^{2i}\ \sigma^2 \\
&\leq \epsilon^2\sigma^2 t_{step}\sum_{i=0}^{\infty}\|M\|_{\mathcal{S}}^{2i} \\
&= \epsilon^2\sigma^2 t_{step}\,\frac{1}{1-\|M\|_{\mathcal{S}}^2}
\end{aligned}$$

which means the random part has bounded variance, even if one considers infinite time. The error is constantly cancelling out with itself, and the Euler method can run for a very long time and still be stable. The radius must be similar to the size of this variance. Define

$$t_{step}^{osc} = \min\left\{t_1, t_2, t_3, t_4, t_5\right\}$$

which is the smallest of all the time steps we have derived so far. Assume that the operator can indeed be bounded by $\|M\|_{\mathcal{S}} \leq (1-\delta)$, then we can define a second bound R_2 on the radius as

$$\begin{aligned}
R_2^2 &= \text{var}\left(\epsilon\sqrt{t_{step}}\sum_{i=0}^{n-1}\|M\|_{\mathcal{S}}^{i}\ \|\xi_{n-1-i}\|\right) \\
&\leq \epsilon^2\sigma^2 t_{step}\,\frac{1}{1-\|M\|_{\mathcal{S}}^2} \\
&= \epsilon^2\sigma^2 t_{step}\,\frac{1}{1-(1-\delta)^2}
\end{aligned}$$

where we have used $\|M\|_{\mathcal{S}} \leq (1-\delta)$. This gives R_2 as

$$R_2 = \sqrt{\frac{\epsilon^2\sigma^2 t_{step}^{osc}}{2\delta-\delta^2}}$$

after choosing a suitable value for δ say $\delta = 0.01$, then a reasonable choice on the radius R could be

$$R = \min\left\{R_1, R_2\right\}.$$

6.3.3 Estimating $t_{step} \leq t_6$

Define the set

$$\mathcal{S}_3 = \left\{x \in \mathbb{R}^r : |w_j(t) - x| \leq R : 0 \leq t \leq T,\ \forall w_j(t)\right\}$$

which is the set of points where the distance to a well is less than the radius over all times and all wells. The new jump size we do not want to see during our simulation is $l_3 = R$

$$l_3 = R$$

which gives another condition on the time step as

$$t_6 = W(\mathcal{S}_3, l_3).$$

The idea behind this final condition is so that the time step t_{step} is small and precise enough such that the region around the well can capture it.

6.4 Selection of Parameters

Six conditions on the time step and two bounds on the radius were developed. These give the recommended values for R and t_{step} as

$$\begin{aligned} R &= \min\{R_1, R_2\} \\ t_{step} &= \min\{t_1, t_2, t_3, t_4, t_5, t_6\}. \end{aligned}$$

These are not rigorous estimates and hence can only be used as a guideline. We performed several checks of consistency of our simulations at a different level of precision before comparing them with the theoretical result. Notice t_{step} and R are just some of the considerations we have to make when choosing a set of parameters to use for the simulations. More details are given below.

6.4.1 Selection of Parameters - Simulation

In Chapter 7 we will introduce the simulations which we are going to do. Notice that the SDE we want to simulate can be rewritten as

$$\begin{aligned} dx &= \left[-\frac{\partial V_0}{\partial x} + F\cos\phi\cos\Omega t \right] dt + \epsilon\, dw_x \\ dy &= \left[-\frac{\partial V_0}{\partial y} + F\sin\phi\cos\Omega t \right] dt + \epsilon\, dw_y. \end{aligned}$$

The following parameters will be fixed with the values

$$a = 0.15 \quad b = 0.1 \quad F = 0.7F^{crit} \quad \Omega = 0.001$$

and ϵ and ϕ will systemically vary by going through all possible combinations of

$$\epsilon = 0.15, 0.16, \ldots, 0.30 \quad \text{and} \quad \phi = 0°, 75°, 78°, 81°, 84°, 87°, 90°$$

and the following value of the time step and radius is used

$$t_{step} = 0.014 \quad R = 0.19.$$

We give some reasons as to why these parameters were chosen.

6.4.2 Selection of Parameters - Validity of Kramers' Formula

There is the validity of Kramers' formula. Consider the graphs below.

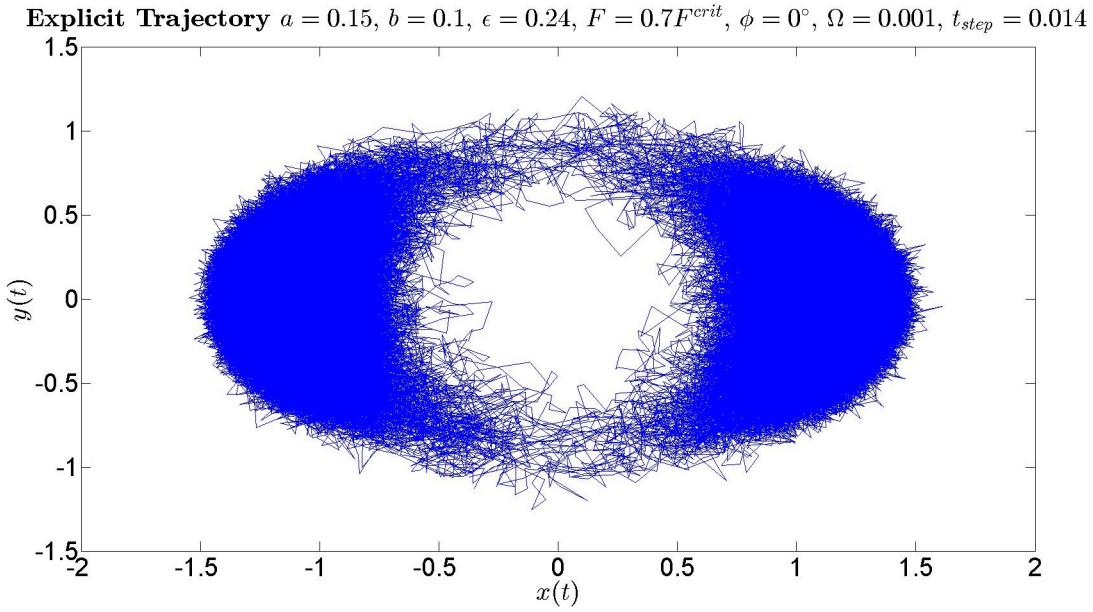

Figure 6.4: Notice that transitions tend to occur near the saddles. This is when Kramers' formula gives a good approximation for the escape rates and escape times.

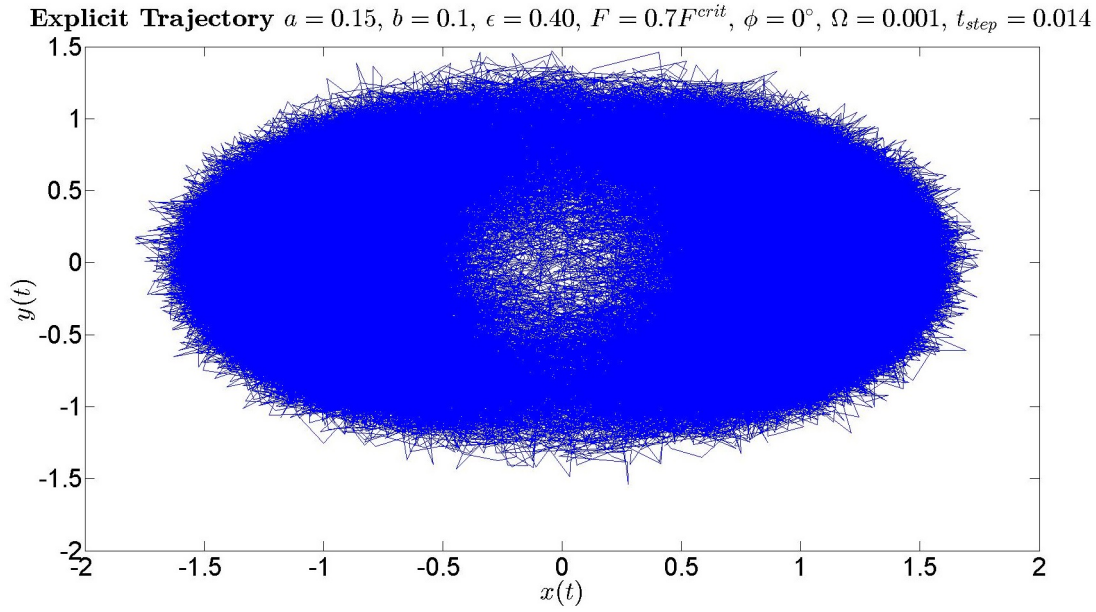

Figure 6.5: For higher noise levels transitions would occur near the hill, which is close to the origin. Kramers' formula is not a good approximation here.

Note that in Figures 6.4 and 6.5 the position of the hill is near the origin and the positions of the saddles are near the y-axis. For very high noise levels Kramers' formula would start to fail as an approximation to the escape times and rates. This is when the particle tend to transit through both the saddles and the hill. A level of subjective judgement is required to gauge how good an approximation Kramers' formula is. Nevertheless it has been checked that at $\epsilon = 0.30$, that Kramers' formula is good enough an approximation for all angles. This checking was also done for an unperturbed static potential, a static potential with maximal forcing and an oscillating potential.[26]

6.4.3 Selection of Parameters - Adiabatic Approximation

There is a reason why we need Kramers' formula to be valid, that is a good approximation to the escape times. Recall that in Chapter 3 theories about escape times from an oscillatory potential was developed. One of the main results were the continuous time invariant measures for an oscillatory potential (see Corollary 3.9) and the PDF for the escape times (see Theorem 3.13). The invariant measures are

$$\overline{\nu}_-(t) = \frac{\int_0^t p(s)g(s)\,ds}{g(t)} + \frac{\int_0^T p(s)g(s)\,ds}{g(t)\,(g(T)-1)}$$

$$\overline{\nu}_+(t) = \frac{\int_0^t q(s)g(s)\,ds}{g(t)} + \frac{\int_0^T q(s)g(s)\,ds}{g(t)\,(g(T)-1)}$$

where

$$g(t) = \exp\left\{\int_0^t p(u) + q(u)\,du\right\}$$

and the PDFs are

$$p_-(t,u) = R_{-1+1}(t)\exp\left\{-\int_u^t R_{-1+1}(s)\,ds\right\}$$

$$p_+(t,u) = R_{+1-1}(t)\exp\left\{-\int_u^t R_{+1-1}(s)\,ds\right\}$$

where we will make the approximation

$$p_{tot}(t) = \frac{1}{2}\int_0^T p_-(t+u,u)m_-(u) + p_+(t+u,u)m_+(u)\,du \approx p_+(t,0)$$

[26]This footnote also applies for graphs later in the thesis. Note that Figures 6.4, 6.5, 7.9, 7.10 and 7.11 have use the following parameters in the simulations.

$$t_{start} = 0 \quad t_{step} = 0.014 \quad t_{end} = 100000$$

although a time step size of $t_{step} = 0.014$ was used, only the data for every ten time steps were shown. This was done to avoid handling a very large graph in MatLab.

because we do not have expressions for $m_-(u)$ and $m_+(u)$. Notice one subtlety about all the theories developed in Chapter 3. It was assumed that the probabilities for transit, that is $p(t)$ and $q(t)$, were known for whatever the driving frequency Ω may be. In the PDFs it was also assumed that the escape rates $R_{-1+1}(t)$ and $R_{+1-1}(t)$ were known for however fast or slow the driving frequency Ω may be.

But such ideal expressions for $p(t)$, $q(t)$, $R_{-1+1}(t)$ and $R_{+1-1}(t)$ are not known. Note that the escape rates as given by Kramers' formula in Chapter 2.2 are only valid for small noise for a static potential. When we do analysis in Chapter 7 the rates $R_{-1+1}(t)$ and $R_{+1-1}(t)$ are calculated by Kramers' formula as though it is a static potential. This means an oscillatory potential is being approximated by a static potential. This is the adiabatic approximation. The following conditions are proposed to decide if the adiabatic approximation is valid

$$\min_{t\in[0,T]} \left\{\tau^{kram}_{-1+1}(t), \tau^{kram}_{+1-1}(t)\right\} \leq \frac{2\pi}{\Omega} \tag{6.11}$$

$$\max_{t\in[0,T]} \left\{\tau^{kram}_{-1+1}(t), \tau^{kram}_{+1-1}(t)\right\} \leq \frac{2\pi}{\Omega} \tag{6.12}$$

that is we consider the minimum and maximum escape times over one period as given by Kramers' formula for a static potential. If this is less than the period of the driving frequency $T = 2\pi/\Omega$, then the adiabatic approximation may be valid. This was checked for all the parameters and Equations 6.11 and 6.12 only hold for the following range of parameters

$$\phi \geq 75^\circ \quad \text{and} \quad \epsilon \geq 0.28.$$

This is a compromise we made. Nevertheless Equations 6.11 and 6.12 do not define the adiabatic approximation, but give an idea of what range the parameters need to be in.

6.4.4 Selection of Parameters - Stability of Deterministic Trajectory

There is also the stability of the deterministic trajectory (when $\epsilon = 0$) to be concerned about. The following starting positions were chosen for four particles

$$\begin{aligned} x_{start} &= +2 \quad y_{start} = +2 \\ x_{start} &= +2 \quad y_{start} = -2 \\ x_{start} &= -2 \quad y_{start} = +2 \\ x_{start} &= -2 \quad y_{start} = -2 \end{aligned}$$

and their trajectories falling through an unperturbed static potential, a static potential with maximal forcing and an oscillating potential were all shown to be stable for all angles. Note that these values of x_{start} and y_{start} were chosen because they are at a place where

the potential is so high the particle will probably never go there. Notice how in Figures 6.4 and 6.5 the trajectory almost never reaches any of the four corners $(2,2)$, $(2,-2)$, $(-2,2)$ and $(-2,-2)$. This is the reason for checking the trajectories there.

6.4.5 Selection of Parameters - Random Number Generator

A few words may also be said about the random number generator we are using. Note that these are pseudo random numbers. They are deterministic sequences of numbers with a very long period. We use the randn() function in MatLab. It is very random and will almost certainly not repeat itself for many years. This is because the period is 2^{1492}. Even with the computer generating 60 million random numbers per second it would still take 10^{434} years to reach the end of the cycle [62]. The function rng('shuffle') was also used, which picks a seed for the random number generator according to the time of the computer clock. When the Parallel ToolBox is used in MatLab, each worker randomly picks a seed for itself.

6.4.6 Selection of Parameters - Calculating Positions of Critical Points

As mentioned in Chapter 5.5.2, the numerical algorithms can be very unstable for calculating the positions of the critical points. In the simulations which we are going to conduct, a table of the positions of all the critical points within one period are calculated first, then stored in the temporary memory of the computer, and looked up every time the position of a critical point is needed. This table is calculated in the following way. Define what we call the pseudo parameters to be

$$u_{start} = 0 \quad u_{step} = 0.001 \quad u_{end} = 2\pi \quad \Omega = 1$$

and then numerically find the positions of the critical points of the equation

$$V_t = V_0 - F_x \cos \Omega t - F_y \cos \Omega t$$

where

$$t = 0, \quad t = u_{step}, \quad 2u_{step}, \quad 3u_{step}, \quad \ldots \quad t \approx 2\pi.$$

Due to the way a matrix is define in MatLab, the last value of t is not exactly at $t = 2\pi$. This table was checked for all angles, and the pseudo parameters we have chosen are stable.

6.4.7 Selection of Parameters - Higher Precision Numerics

We also have some remarks about the time step we have chosen. As we shall see in Chapter 7, one of the main effects which we have observed in this thesis is what we call the Single,

Intermediate and Double Frequency in the histograms of escape times. These effects were first observed for the values of $a = 0.15$, $b = 0.1$, $F = 0.7F^{crit}$ and $\Omega = 0.001$ when

$$t_{step} \geq 0.0286 \quad R \geq 0.3218$$

and was observed again when $t_{step} = 0.014$ and $R = 0.19$. Note that $t_{step} = 0.014$ and $R = 0.19$ was used for the results of this thesis. Thus we have confidence in believing that the data we have collected is reliable. Consider the graphs below. They are histograms of escape times from both the left and right wells combined. They are also normalised to give an empirical PDF.

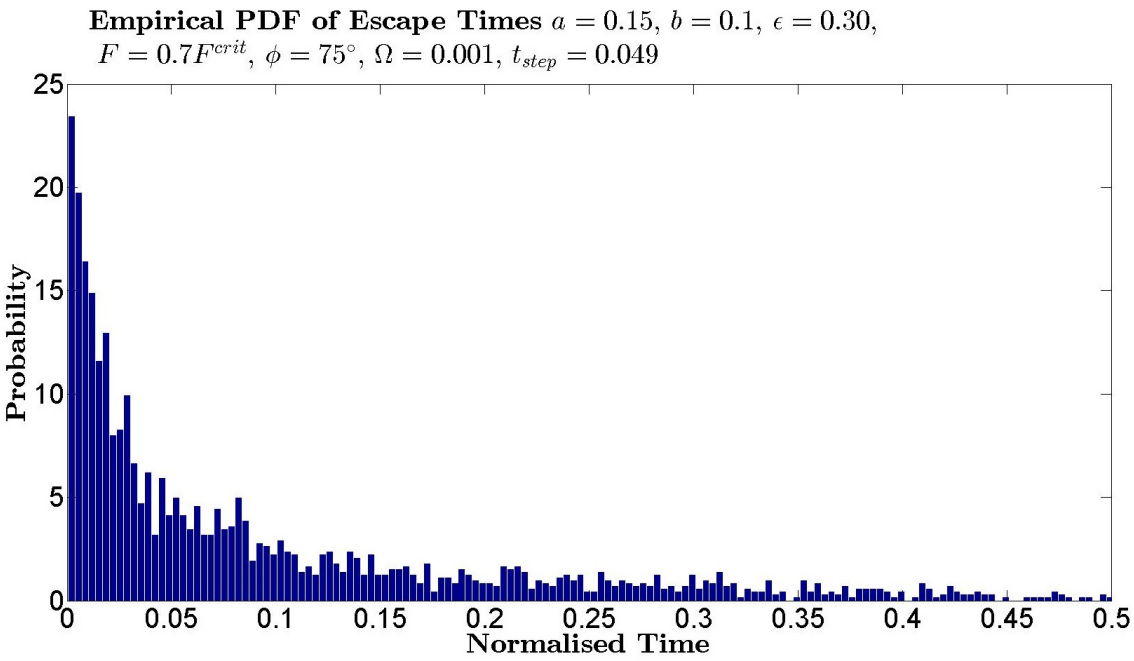

Figure 6.6: Here 2239 transitions were used. The averaged measured escape time is $0.0977T$. The radius used was $R = 0.5386$.

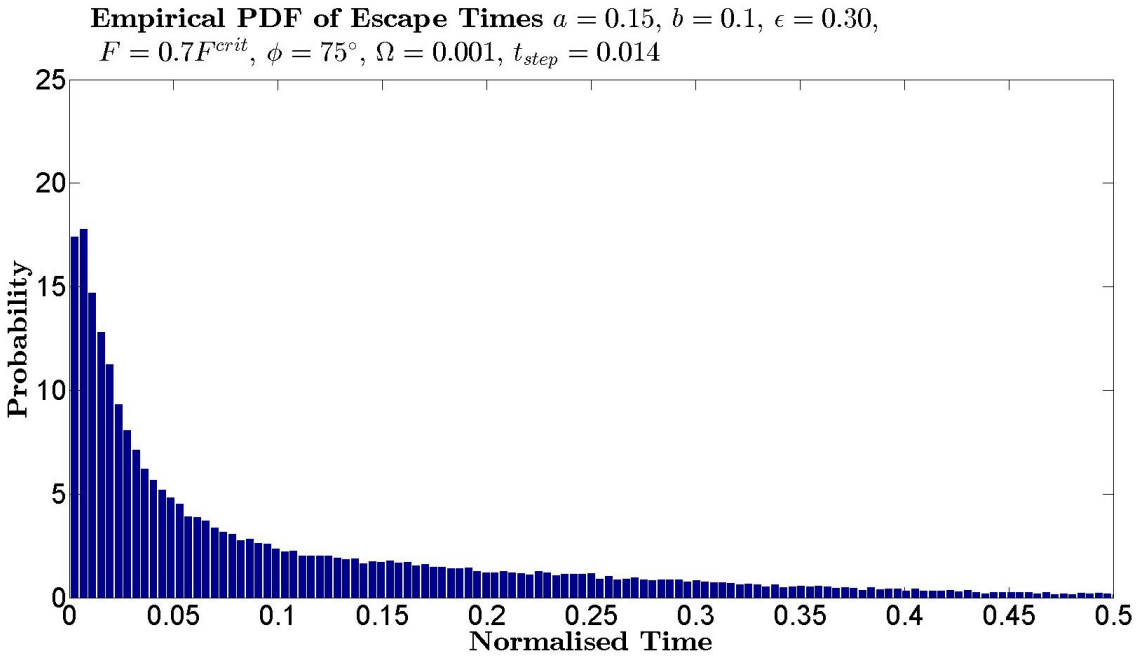

Figure 6.7: Here 56244 transitions were used. The averaged measured escape time is $0.1064T$. The radius used was $R = 0.19$.

The higher the noise level ϵ is the more susceptible to errors would the Euler method be. This is why the highest level of noise $\epsilon = 0.30$ is chosen for these examples. One may argue that having more transitions would give a better measurement of the escape times in Figure 6.7. But the difference in real time between the measured averaged escape times is only 54 seconds out of a period of $T = 2\pi/\Omega = 6283$ seconds.

Chapter 7

Simulations, Results and Analysis

This Chapter presents the main results of this thesis. The six measures M_1, M_2, M_3, M_4, M_5, M_6, the distributions of escape times and the newly developed conditional Kolmogorov-Smirnov test are used to analyse simulations of the SDE with the Mexican Hat Toy Model used as the potential. The six measures are shown to be insensitive to the saddles changing from alternating to synchronised. This is shown to be due to the fact that the invariant measure is constant for synchronised saddles. The distribution of escape times shows new signatures as the saddles change from alternating to synchronised and the conditional Kolmogorov-Smirnov test is demonstrated to be an appropriate way to analyse the escape times collected from many transitions.

We simulate a series of stochastic trajectories for the Mexican Hate Toy Model and analyse them. We remind ourselves of the unperturbed Mexican Hat potential

$$V_0(x,y) = \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 \quad \text{where} \quad r = \sqrt{x^2+y^2}$$

and the SDE we want to simulate is

$$\begin{aligned} dx &= \left[-\frac{\partial V_0}{\partial x} + F_x \cos \Omega t \right] dt + \epsilon \, dw_x \\ dy &= \left[-\frac{\partial V_0}{\partial y} + F_y \cos \Omega t \right] dt + \epsilon \, dw_y \end{aligned}$$

where F_x and F_y are the x and y components of the forcing, Ω is the forcing frequency, ϵ is the noise level and w_x and w_y are two independent Wiener processes. We can define the magnitude and angle of the forcing by

$$F = \sqrt{F_x^2 + F_y^2} \quad \text{and} \quad \phi = \tan^{-1}\left(\frac{F_y}{F_x}\right).$$

This means the SDE can be written alternatively as

$$dx = \left[-\frac{\partial V_0}{\partial x} + F \cos\phi \cos \Omega t \right] dt + \epsilon \, dw_x$$

$$dy = \left[-\frac{\partial V_0}{\partial y} + F \sin\phi \cos\Omega t \right] dt + \epsilon \, dw_y.$$

The critical forcing is defined by (see Equation 5.9)

$$F^{crit} = \min\left\{ F_x^{sad}, F_x^{crit}, F_y^{sad}, F_y^{crit} \right\}.$$

7.1 Details of the Simulations

The Euler method was used to simulate this SDE with the following parameters being fixed at the following values

$$a = 0.15 \quad b = 0.1 \quad F = 0.7F^{crit} \quad \Omega = 0.001.$$

The angle of the forcing ϕ and the noise level ϵ were varied. The values used were

$$\epsilon = 0.15, 0.16, \ldots, 0.30 \quad \text{and} \quad \phi = 0^\circ, 75^\circ, 78^\circ, 81^\circ, 84^\circ, 87^\circ, 90^\circ.$$

The averaged diffusion trajectories $\langle x_t \rangle$ and $\langle y_t \rangle$ were collected. The averaged Markov Chain $\langle Y_t^\epsilon \rangle$ and the averaged Out-of-Phase Markov Chain $\langle \overline{Y}_t^\epsilon \rangle$ were collected as well. This would allow for the calculation of the invariant measures $\overline{\nu}_-(\cdot)$ and $\overline{\nu}_+(\cdot)$. The time coordinates of the entrance and exit to and from the left and right wells were also collected. This would allow for the calculation of the escape times. We use the following values for the time step and the radius around the wells

$$t_{step} = 0.014 \quad \text{and} \quad R = 0.19.$$

Note that the period of the forcing is denoted by

$$T = \frac{2\pi}{\Omega}.$$

The averaged trajectories were simulated by taking the averaged of 200 realisations. Each realisation was 30 periods long, that is a trajectory over the interval $[0, 30T]$. The initial value of the state probabilities were set at

$$\nu_-(0) = \nu_+(0) = \frac{1}{2}$$

which assists in giving a faster convergence to the invariant measures (see Theorems 3.1, 3.2, and 3.8). We should also stress that a lot of the data and results presented in this Chapter is just a selection out of a much wider range of results. All 112 combinations of the parameters were simulated and analysed.

7.2 Six Measures Analysis

The six measures are calculated for the diffusion and Markov Chain case for all angles of the forcing ϕ and all noise levels ϵ used in the simulations. When $\phi = 90°$ the wells were moving up and down but they were always at the same height as each other. The distance from either wells to the saddles, which is a gateway for escape, is the same in both wells at all times. This means the $\phi = 90°$ can be modelled by a synchronised Markov Chain with $p = q$. The invariant measures for the $\phi = 90°$ case as predicted by Corollary 3.4 and 3.10 is $\overline{\nu}_- = \overline{\nu}_+ = \frac{1}{2}$. The Fourier Transform of the averaged Markov Chain is predicted to be zero by Corollary 3.6 and 3.12. This predicts the six measures at $\phi = 90°$ to be

$$
\begin{aligned}
M_1 &= 0 \\
M_2 &= 0 \\
M_3 &= \int_0^T \langle Y_t^\epsilon \rangle^2 \, dt = \int_0^T \left(\nu_+(t) - \nu_-(t)\right)^2 \, dt = 0 \\
M_4 &= \int_0^T \left\langle \overline{Y}_t^\epsilon \right\rangle dt \\
&= \int_0^{T/2} 0 \times \nu_-(t) + 1 \times \nu_+(t) \, dt + \int_{T/2}^T 1 \times \nu_-(t) + 0 \times \nu_+(t) \, dt \\
&= \frac{1}{2} T \\
M_5 &= \int_0^T \phi^-(t) \ln\left(\frac{\phi^-(t)}{\overline{\nu}_-(t)}\right) + \phi^+(t) \ln\left(\frac{\phi^+(t)}{\overline{\nu}_+(t)}\right) dt \\
&= \int_0^{T/2} \ln\left(\frac{1}{\overline{\nu}_-(t)}\right) \, dt + \int_{T/2}^T \ln\left(\frac{1}{\overline{\nu}_+(t)}\right) \, dt \\
&= +T \ln(2) \\
M_6 &= \int_0^T -\overline{\nu}_-(t) \ln \overline{\nu}_-(t) - \overline{\nu}_+(t) \ln \overline{\nu}_+(t) \, dt \\
&= +T \ln(2).
\end{aligned}
$$

Note that $\ln(2) = 0.6931 \approx 0.7$. Notice that for very low noise level $\epsilon \approx 0$ the probabilities of escape from either well is so small it may be approximately modelled by a synchronised Markov Chain with $p \approx q$. The results below confirm the predictions for the case of $\phi = 90°$.

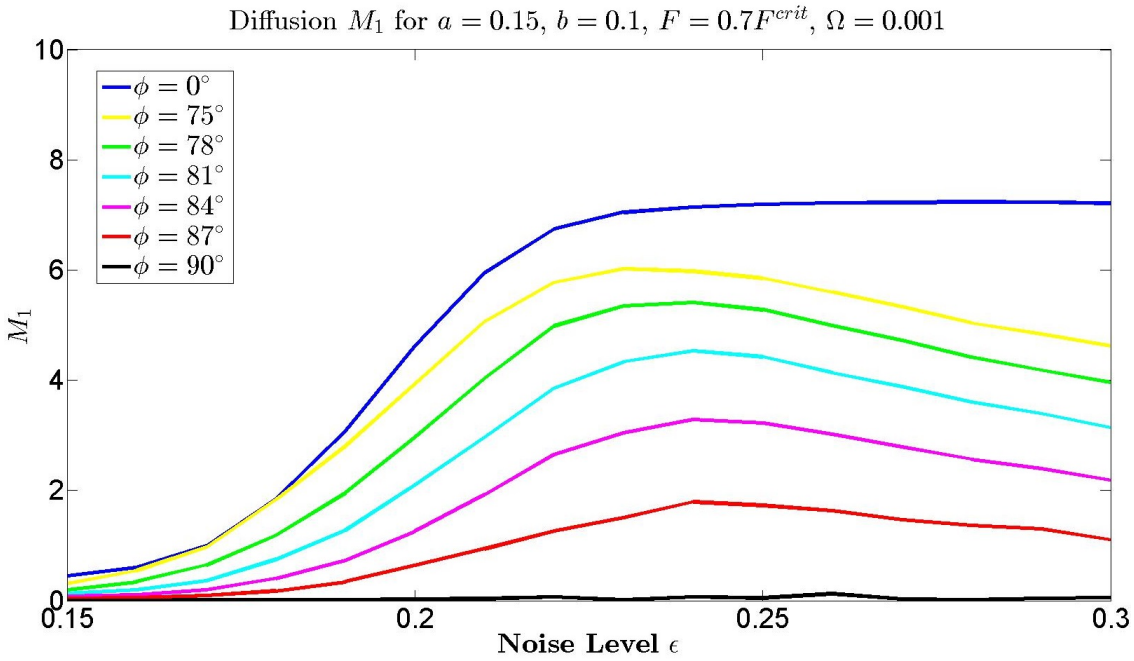

Figure 7.1: The measure M_1 for the diffusion case for various angles and noise levels.

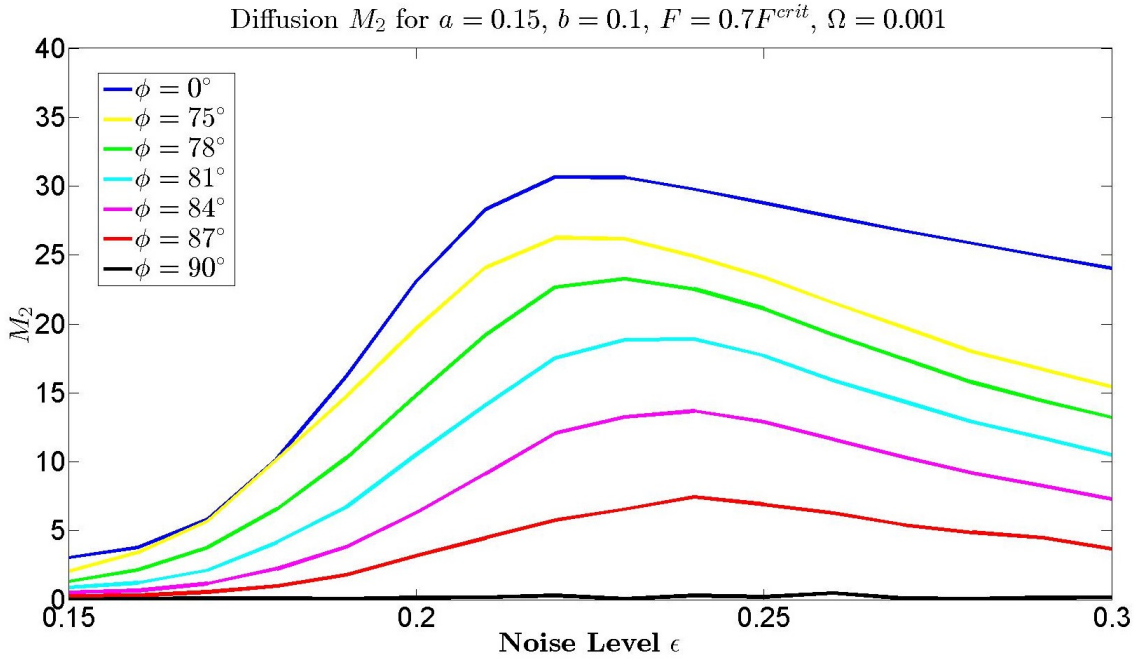

Figure 7.2: The measure M_2 for the diffusion case for various angles and noise levels.

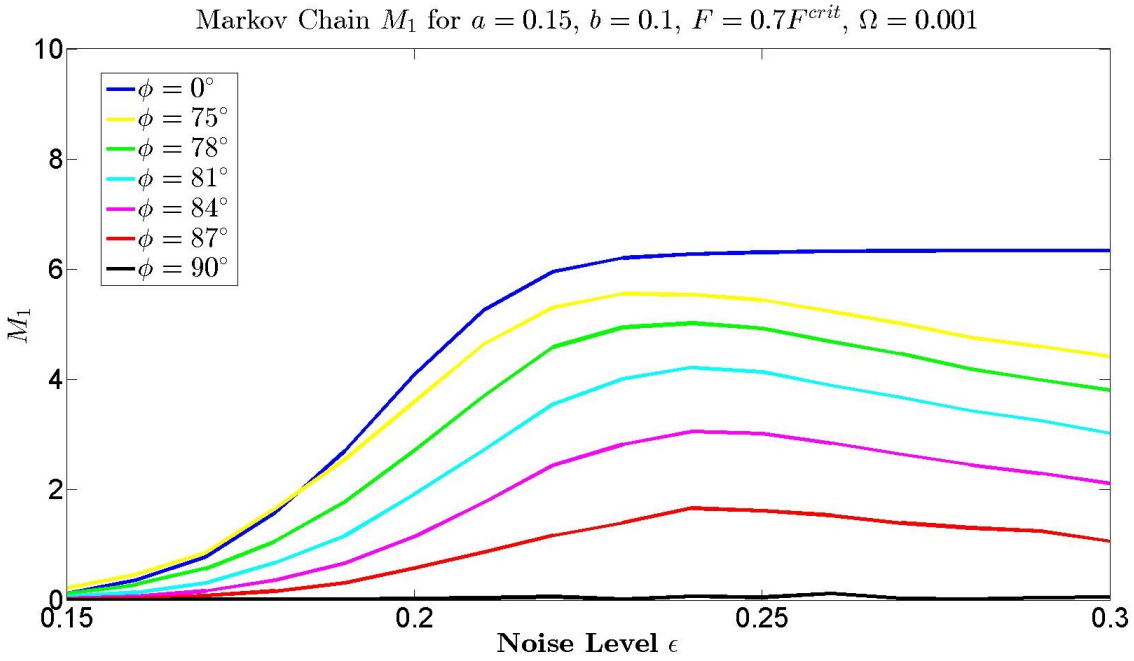

Figure 7.3: The measure M_1 for the Markov Chain for various angles and noise levels.

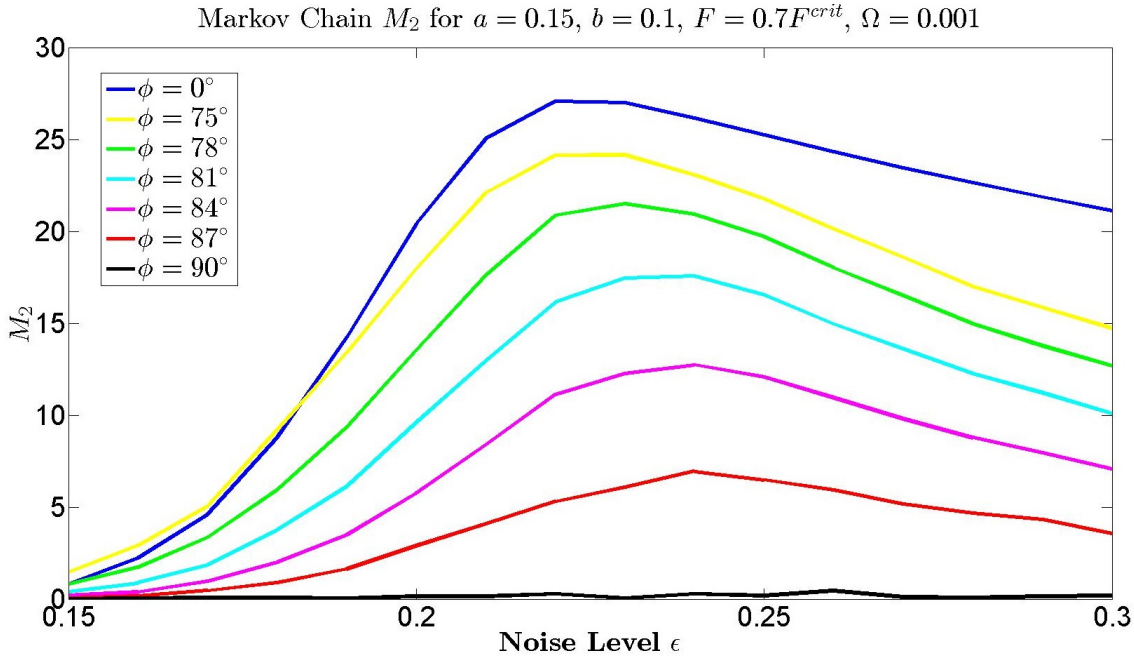

Figure 7.4: The measure M_2 for the Markov Chain for various angles and noise levels.

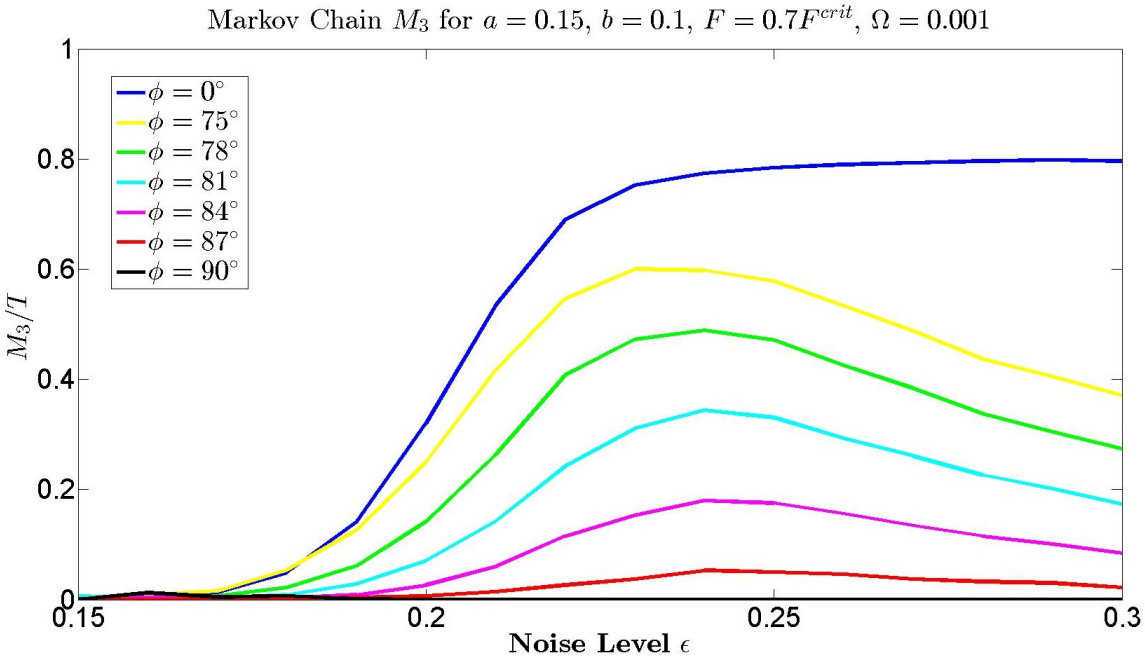

Figure 7.5: The measure M_3 for the Markov Chain for various angles and noise levels.

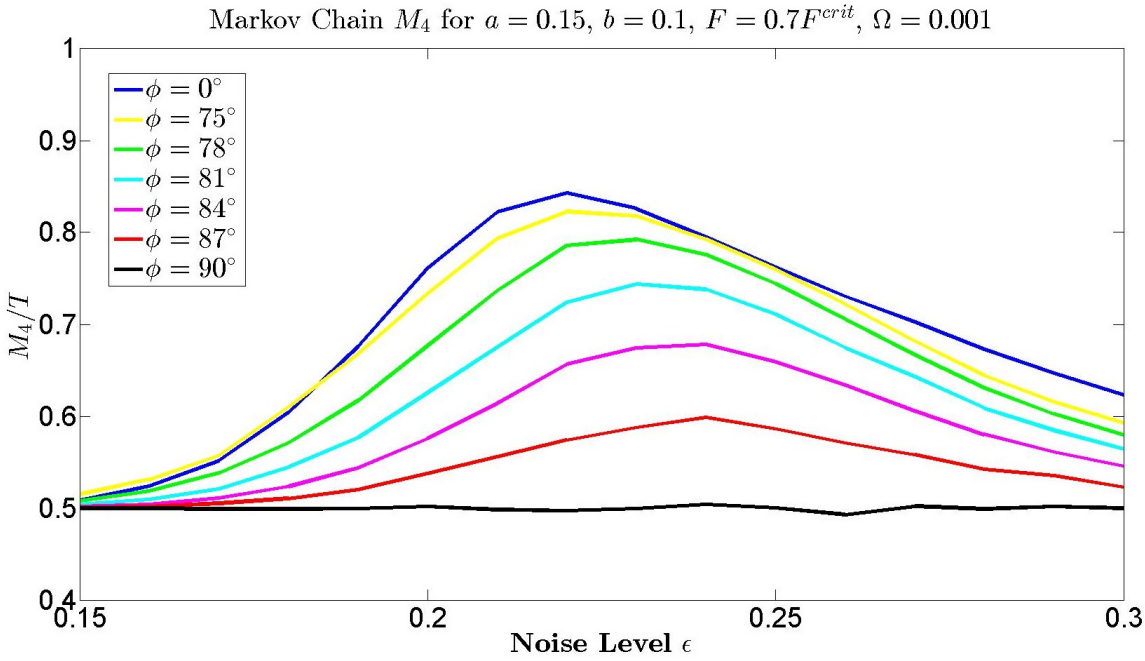

Figure 7.6: The measure M_4 for the Markov Chain for various angles and noise levels.

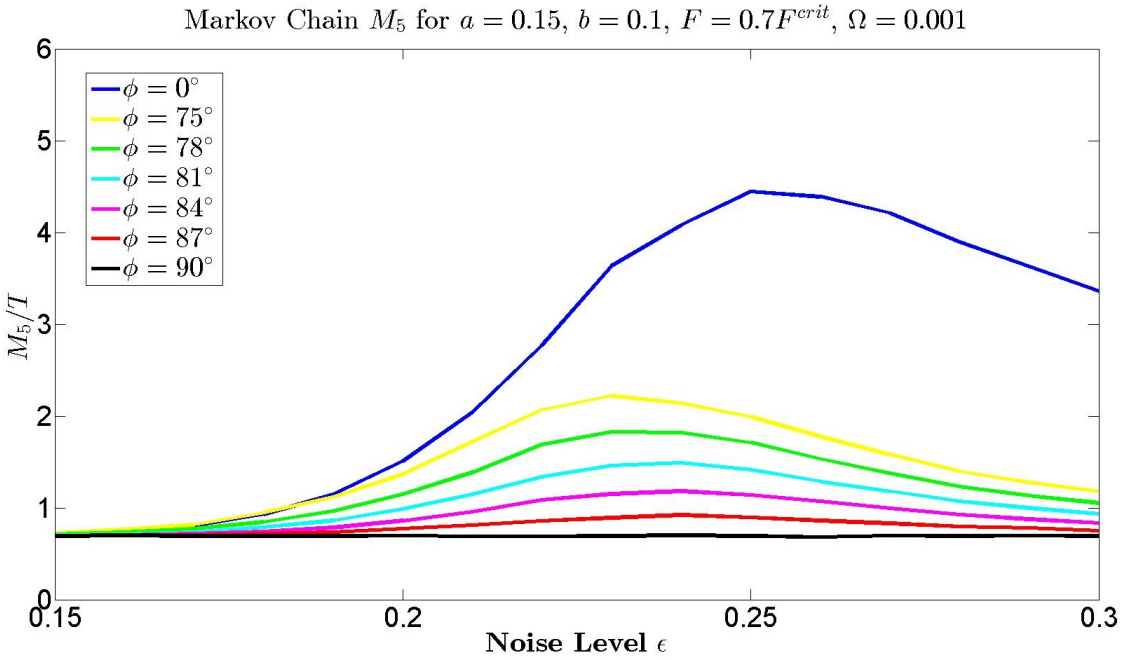

Figure 7.7: The measure M_5 for the Markov Chain for various angles and noise levels.

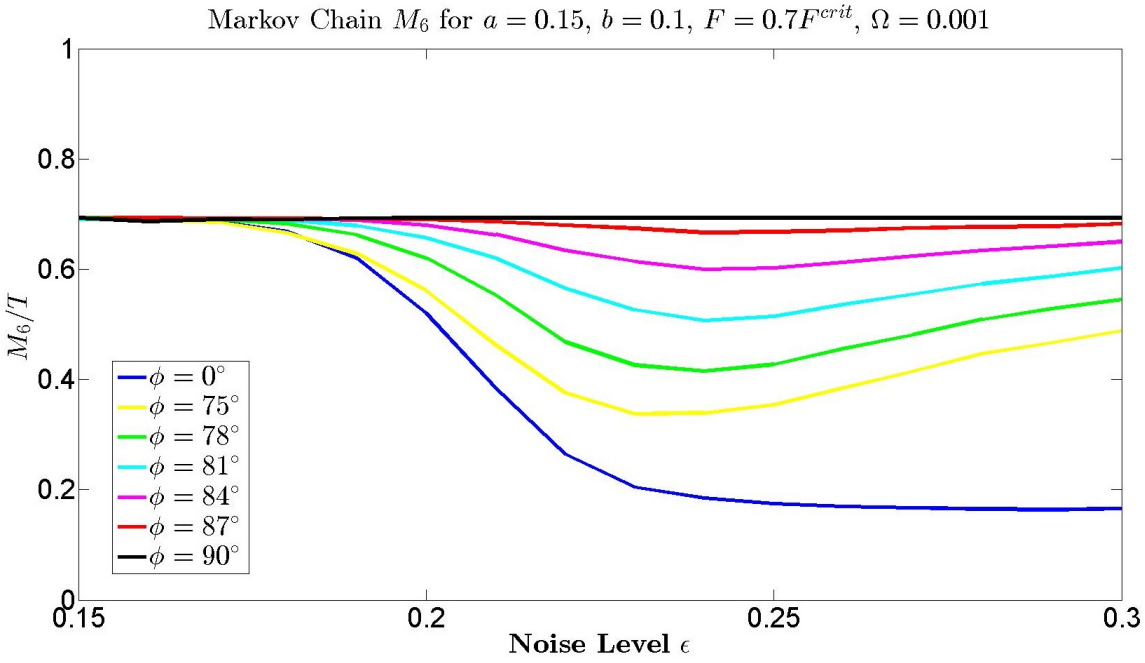

Figure 7.8: The measure M_6 for the Markov Chain for various angles and noise levels.

7.2.1 Interpretation of the Six Measures Analysis

The six measures M_1, M_2, M_3, M_4, M_5 and M_6 were plotted as a function of the noise level ϵ. The six measures show a regular systematic behaviour in the angle ϕ. The shape of the graphs of the six measures were very similar for all the angles. As the angle increased from $\phi = 0°$ to $\phi = 90°$ the six measures gradually tended to being nearly constant in ϵ.

This effect can be explained with the invariant measures. When $\phi = 0°$ the probabilities for escaping from left to right p_{-1+1} was different to the probabilities for escaping from right to left p_{+1-1}. But in the $\phi = 90°$ case they are the same, that is

$$\begin{aligned} \phi = 0° \quad & p_{-1+1} \neq p_{+1-1} \\ \phi = 90° \quad & p_{-1+1} = p_{+1-1}. \end{aligned}$$

The is can be understood geometrically. For $\phi = 0°$ we have $F_x \neq 0$ and $F_y = 0$. The two wells in the Mexican Hat potential move up and down and are alternating with each other. When one well is high the other is low. For $\phi = 90°$ we have $F_x = 0$ and $F_y \neq 0$. The two wells are always at the same height as each other and the distance to the saddles (which is a gateway to escape) is also the same in both wells.

Recall our discussions on the Markov Chain in Chapter 3. The p is related to left to right escape p_{-1+1} and q was related to right to left escape p_{+1-1}. For $\phi = 0°$ the Markov Chain can be modelled with $p \neq q$ and for $\phi = 90°$ the Markov Chain can be modelled with $p = q$. In the case of $\phi = 0°$ the invariant measure was cyclically changing in time. In the case of $\phi = 90°$ the invariant measure was constant at $\overline{\nu}_-(\cdot) = \overline{\nu}_+(\cdot) = \frac{1}{2}$. This explains why the six measures M_1, M_2, M_3, M_4, M_5 and M_6 were nearly constant for angle $\phi = 90°$. As ϕ changed from $\phi = 0°$ to $\phi = 90°$, the Markov Chain changed from being modelled by $p \neq q$ to being modelled by $p = q$. This explains the change in the six measures tending to being constant in ϵ as ϕ was varied. The six measures can be thought of as a way of measuring how far away the invariant measures are from being constant. If the invariant measures are constant then the six measures will also be constant.[27]

For fixed ϕ near $\phi = 90°$ there is no pronounced maximum of any measure for varying ϵ. Hence the six measures indicate the absence of a pronounced stochastic resonance near $\phi = 90°$. But consider the trajectories at a range of angles.

[27] See Appendix B.3 for how M_5 and M_6 were numerically calculated. The ideas were not that trivial.

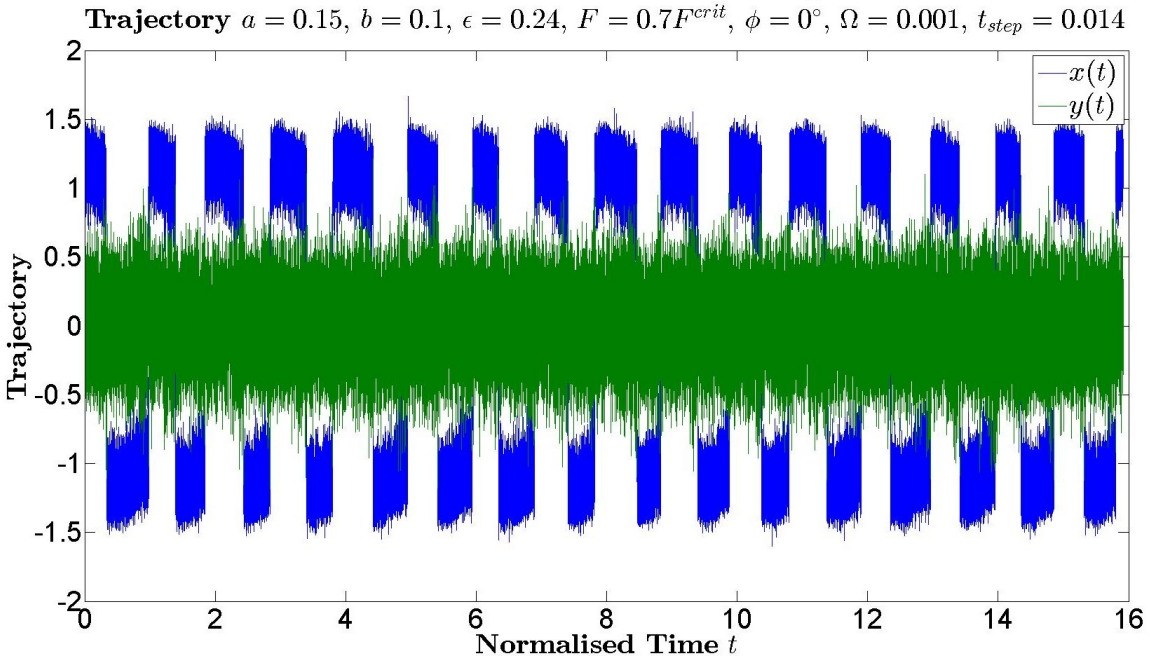

Figure 7.9: The blue trajectory is $x(t)$ and the green trajectory is $y(t)$.

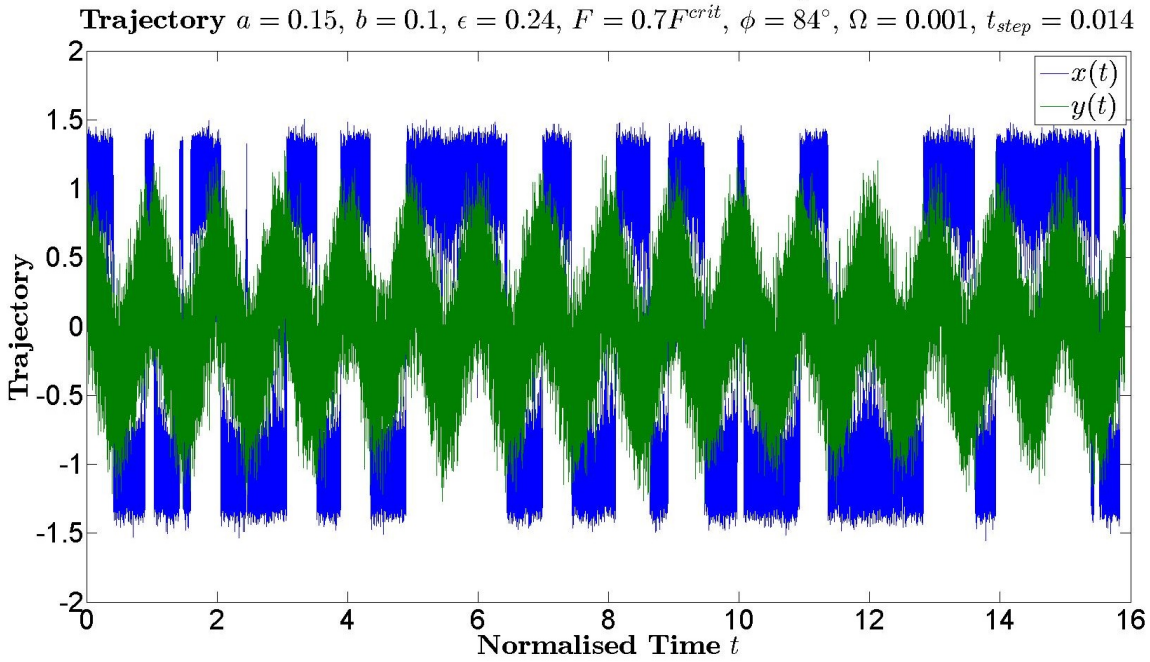

Figure 7.10: The blue trajectory is $x(t)$ and the green trajectory is $y(t)$.

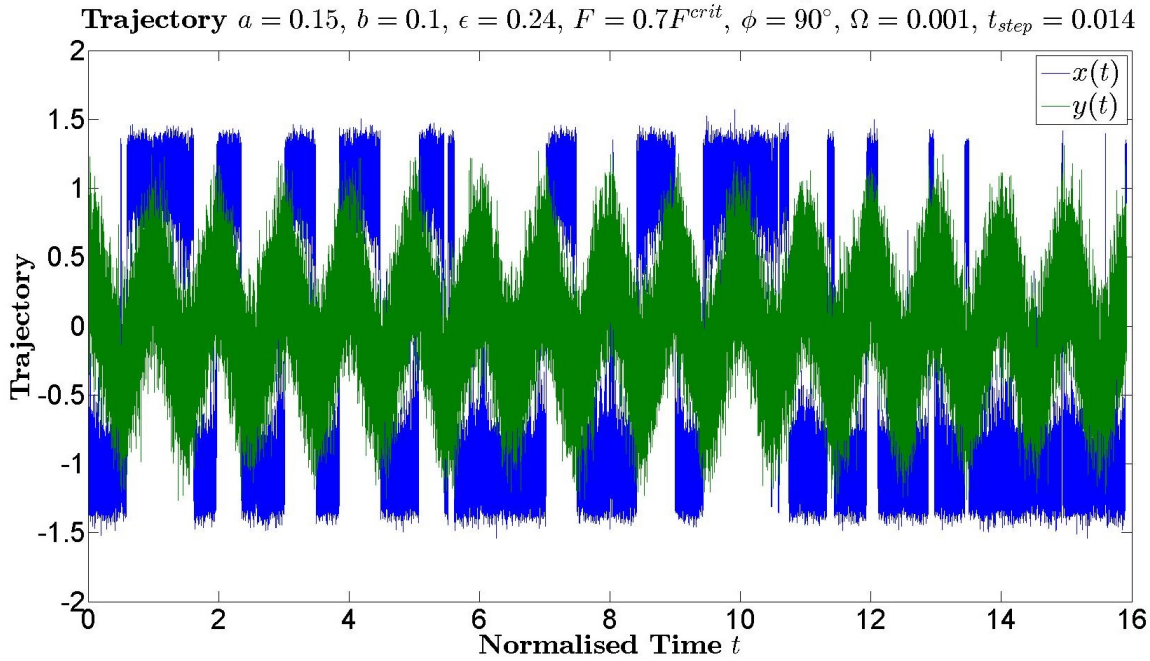

Figure 7.11: The blue trajectory is $x(t)$ and the green trajectory is $y(t)$.

When $\phi = 0°$ the $x(t)$ shows quasi-deterministic behaviour. The transitions are very regular and $y(t)$ fluctuates around zero. As the angle varies the transitions become less regular and $y(t)$ starts to oscillate. This suggests that there is some regularity in the behaviour of the trajectories but the six measures are not detecting it. Further studies with the escape times would tell us more.

7.3 Escape Time and Conditional KS Test Analysis

We remind ourselves of the PDF of escape times and the way the conditional KS test can be applied in our context. The conditional PDF of the escape times are

$$p_-(t,u) = R_{-1+1}(t) \exp\left\{-\int_u^t R_{-1+1}(s)\, ds\right\}$$

$$p_+(t,u) = R_{+1-1}(t) \exp\left\{-\int_u^t R_{+1-1}(s)\, ds\right\}$$

where R_{-1+1} and R_{+1-1} are the Kramers' escape rate from left to right and right to left. In the case of $p_-(t,u)$, t is the time coordinate of escape from the left well and u is the time coordinate of entrance into the left well. In the case of $p_+(t,u)$, t is the time coordinate of escape from the right well and u is the time coordinate of entrance into the right well. If we do not differentiate between escaping from the left or right then the PDF for an escape

time t is (note that t here is an escape time as it is and not a time coordinate)

$$p_{tot}(t) = \frac{1}{2}\int_0^T p_-(t+u,u)m_-(u) + p_+(t+u,u)m_+(u)\,du$$

where $m_-(\cdot)$ and $m_+(\cdot)$ are PDFs of the time of entrance into the left and right well respectively. We do not have explicit expressions for $m_-(\cdot)$ and $m_+(\cdot)$. The $p_{tot}(t)$ is approximated by

$$p_{tot}(t) \approx p_+(t,0).$$

The times it took to escape from both the left or right wells are plotted in histograms. This is an empirical approximation to the PDF $p_{tot}(t) \approx p_+(t,0)$. A selection of some of the results are given below for various angles of the forcing ϕ and noise level ϵ. They are examples of the Singles, Intermediate and Double Frequencies which we will explain later. Note that the escape times are given in units of normalised time, which is in the number of periods T.

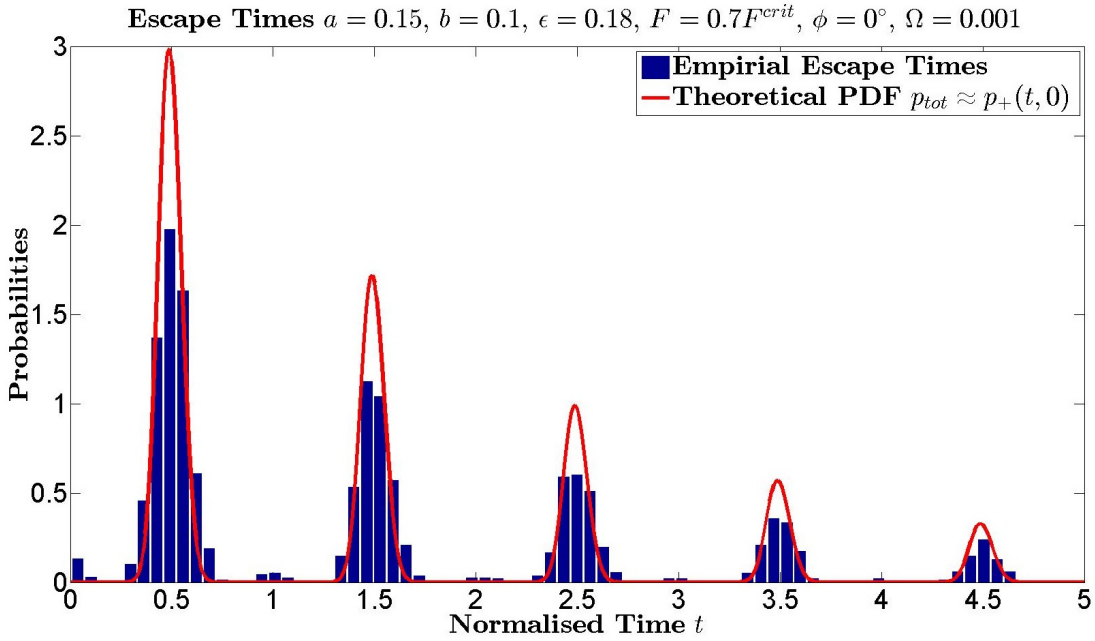

Figure 7.12: This is an example of the Single Frequency.

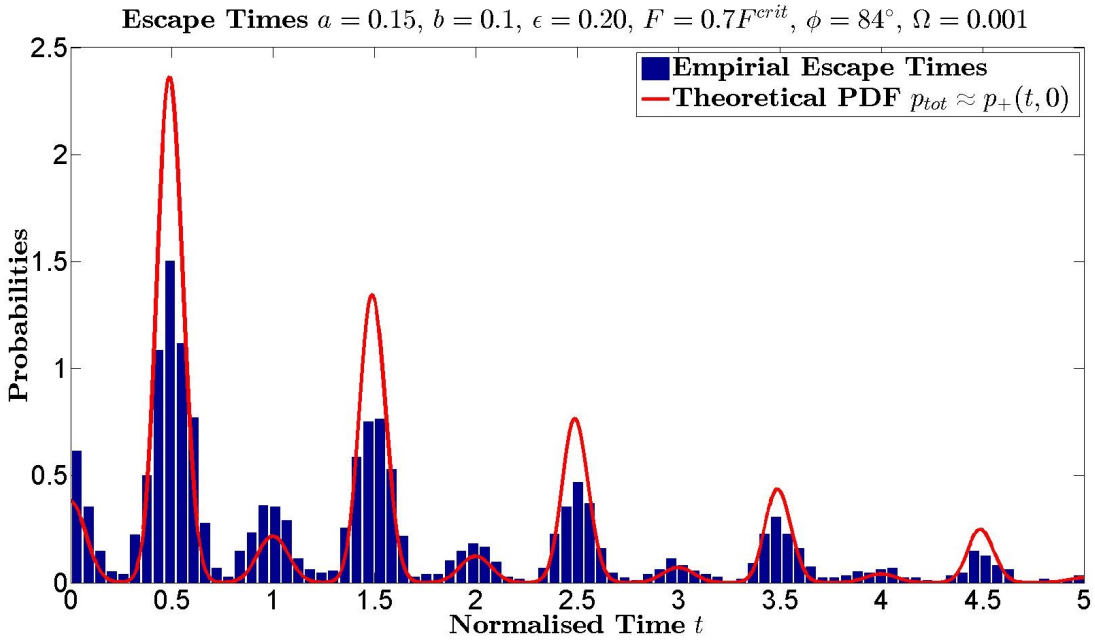

Figure 7.13: This is an example of the Intermediate Frequency.

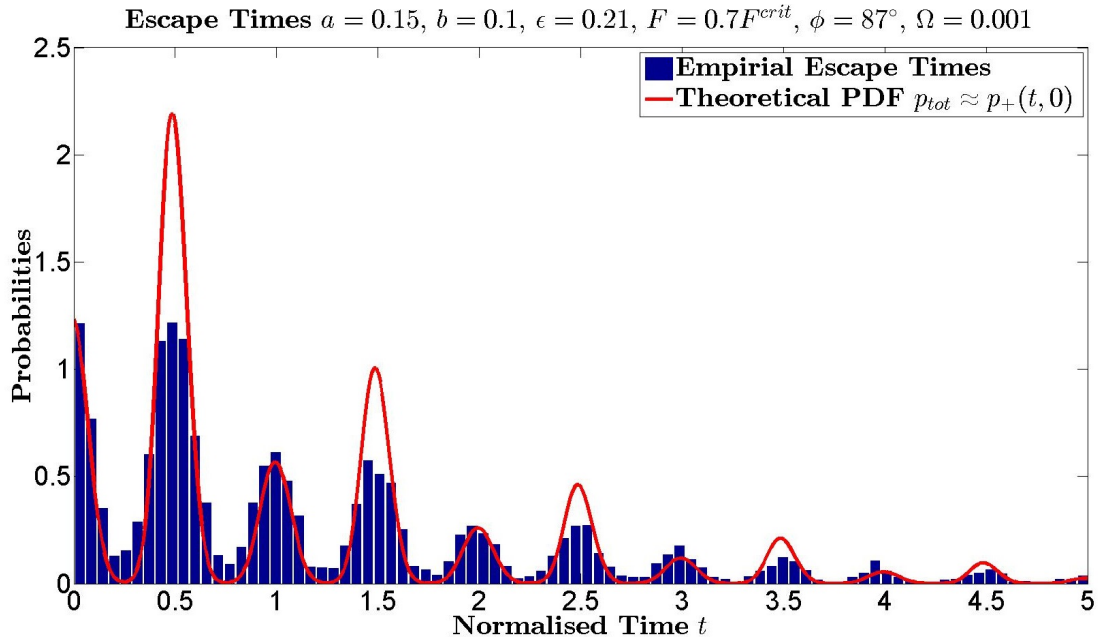

Figure 7.14: This is an example of the Intermediate Frequency tending closer to the Double Frequency.

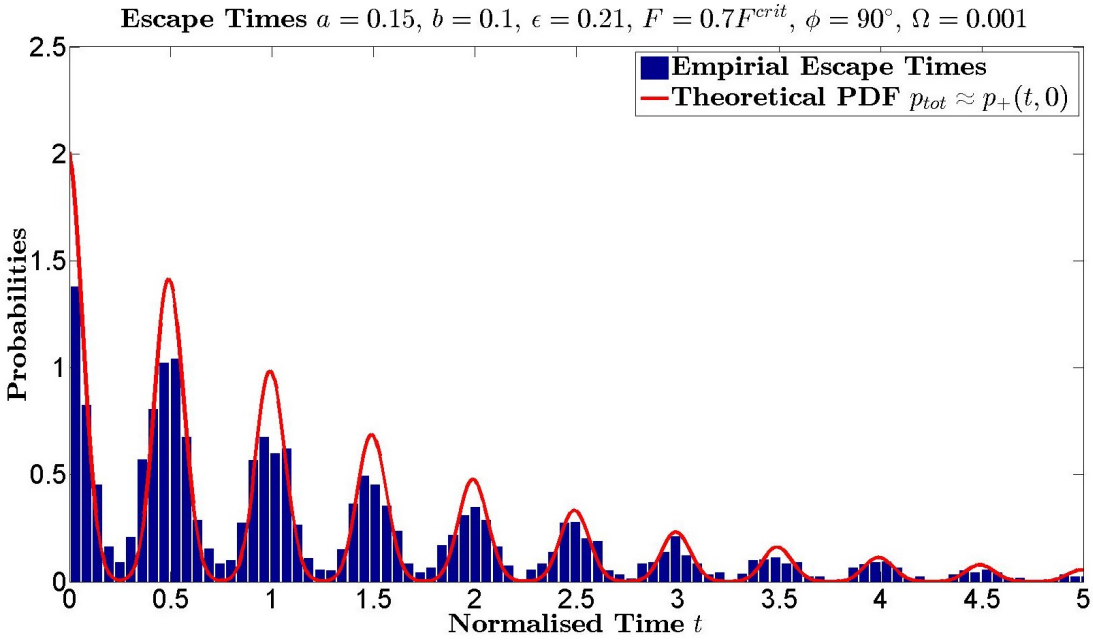

Figure 7.15: This is an example of the Double frequency.

It is important to note that Figures 7.12, 7.13, 7.14 and 7.15 are histograms of the actual times it took to escape from either wells without differentiation between wells on the left or right. The times of entrance into the wells are not shown. The PDF used is $p_{tot}(\cdot)$ which is being approximated by $p_{tot}(t) \approx p_+(t, 0)$.

These escape times can be analysed in a different way. Let u be the time of entrance into a well and t the time of exit from a well. Figures 7.12, 7.13, 7.14 and 7.15 are therefore histograms of the $(t-u)$ for both left and right escapes combined. Thus $0 \leq mod(u, T) \leq 1$ is the phase of entrance into a well and $mod(t-u, T)$ is the escape time itself in normalised time. Such an analysis is done for the times in Figures 7.12, 7.13, and 7.15 for both the left and right wells respectively.

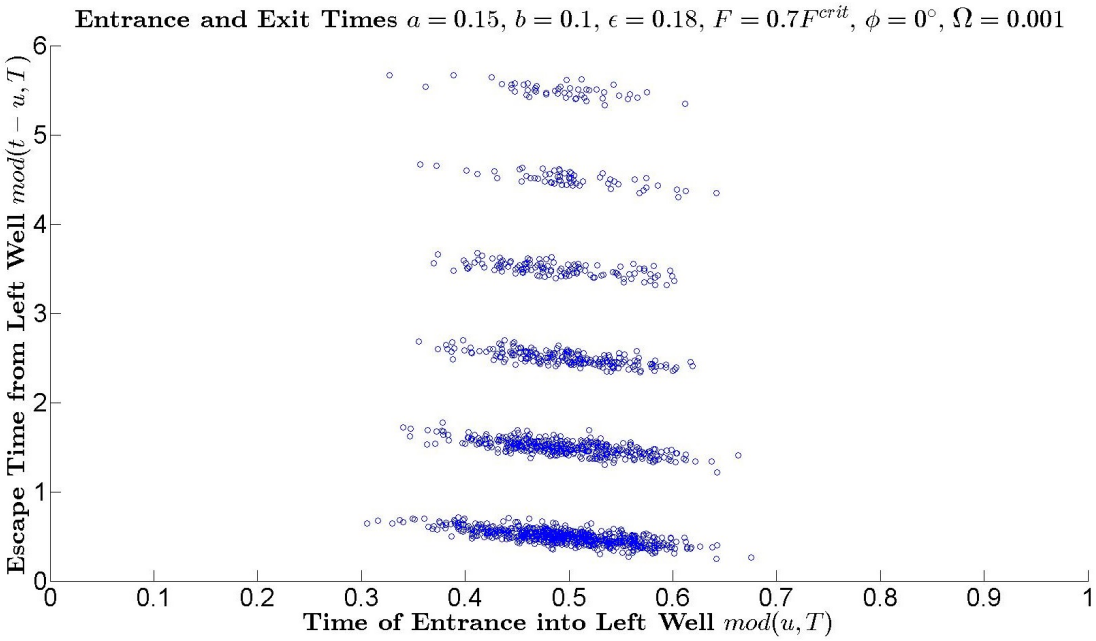

Figure 7.16: The u is the time of entrance into the well and t is the time of exit from the well.

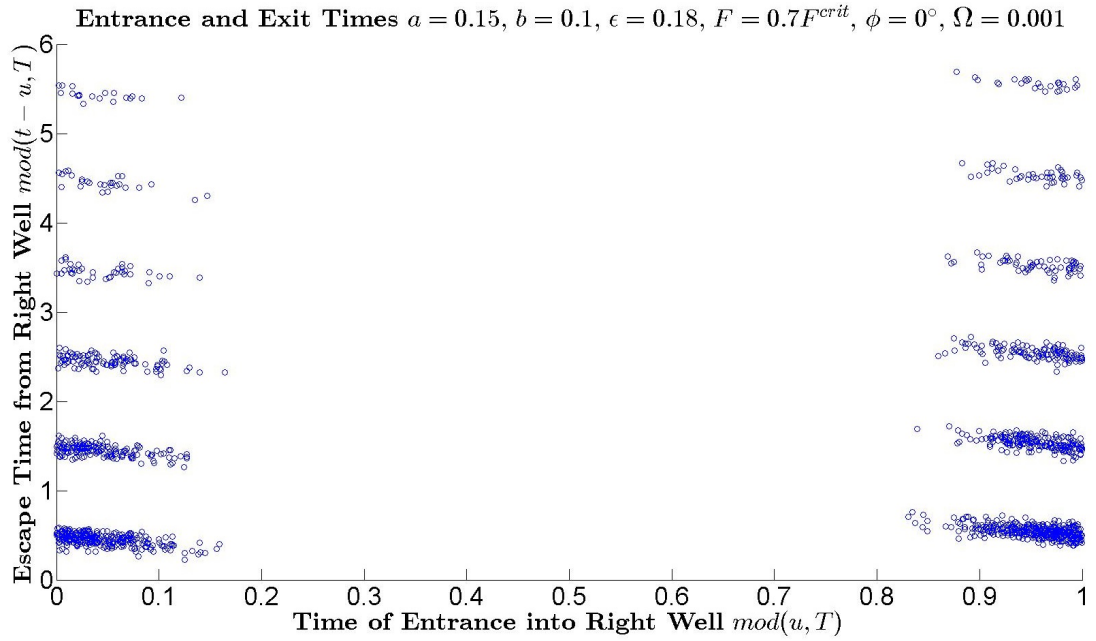

Figure 7.17: The u is the time of entrance into the well and t is the time of exit from the well.

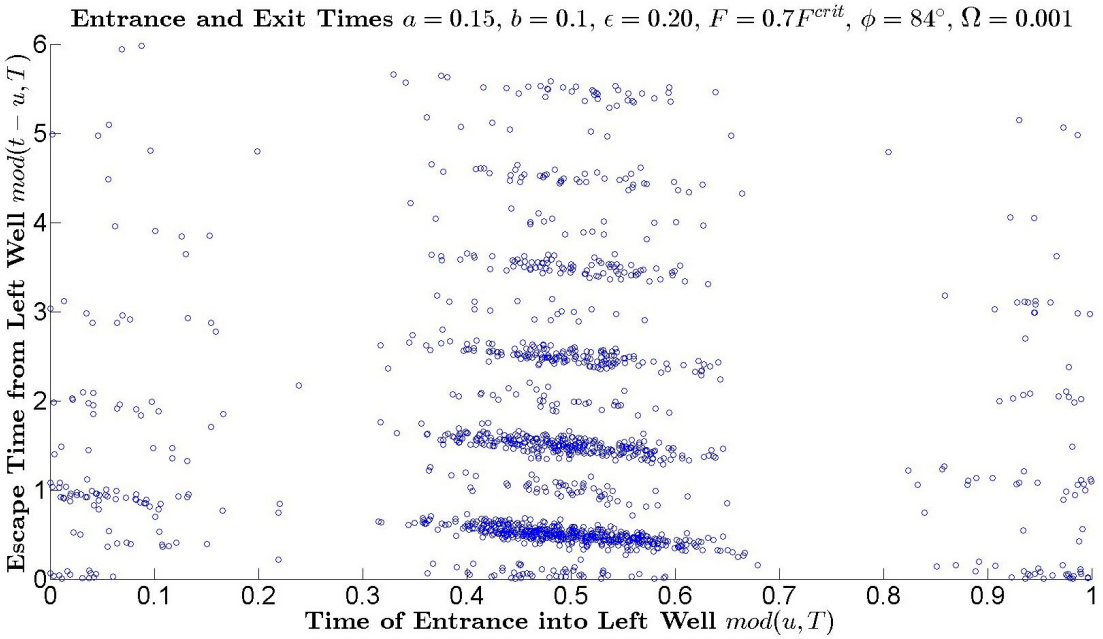

Figure 7.18: The u is the time of entrance into the well and t is the time of exit from the well.

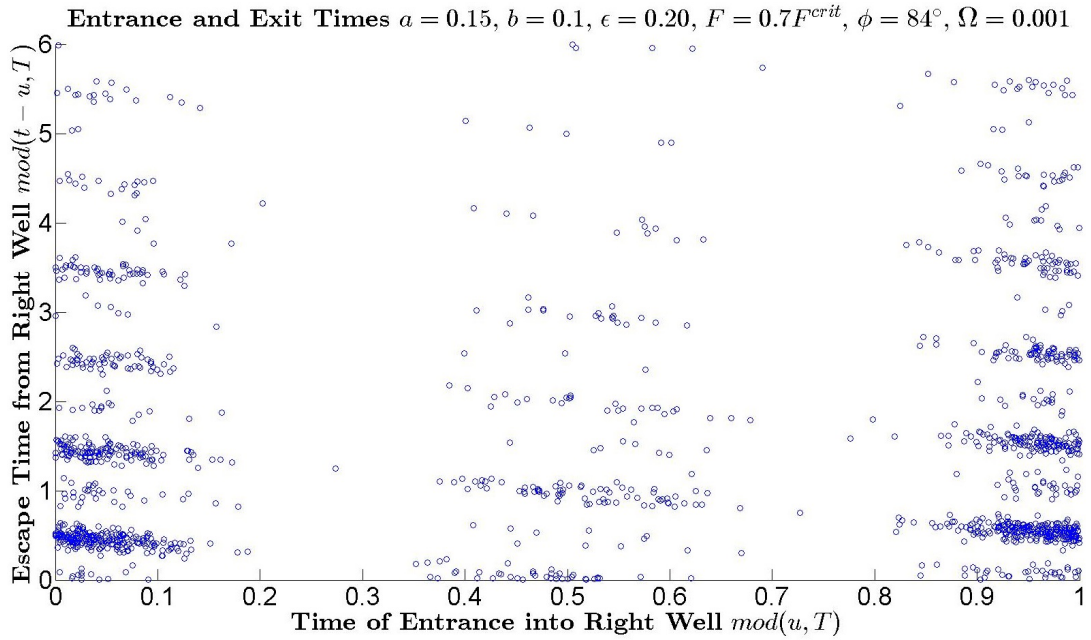

Figure 7.19: The u is the time of entrance into the well and t is the time of exit from the well.

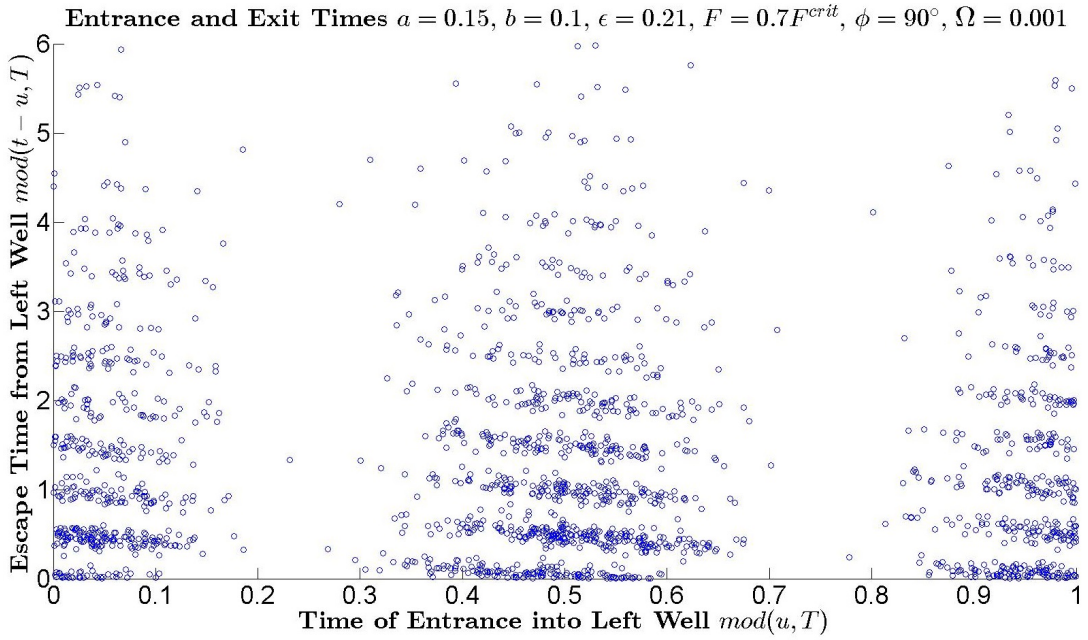

Figure 7.20: The u is the time of entrance into the well and t is the time of exit from the well.

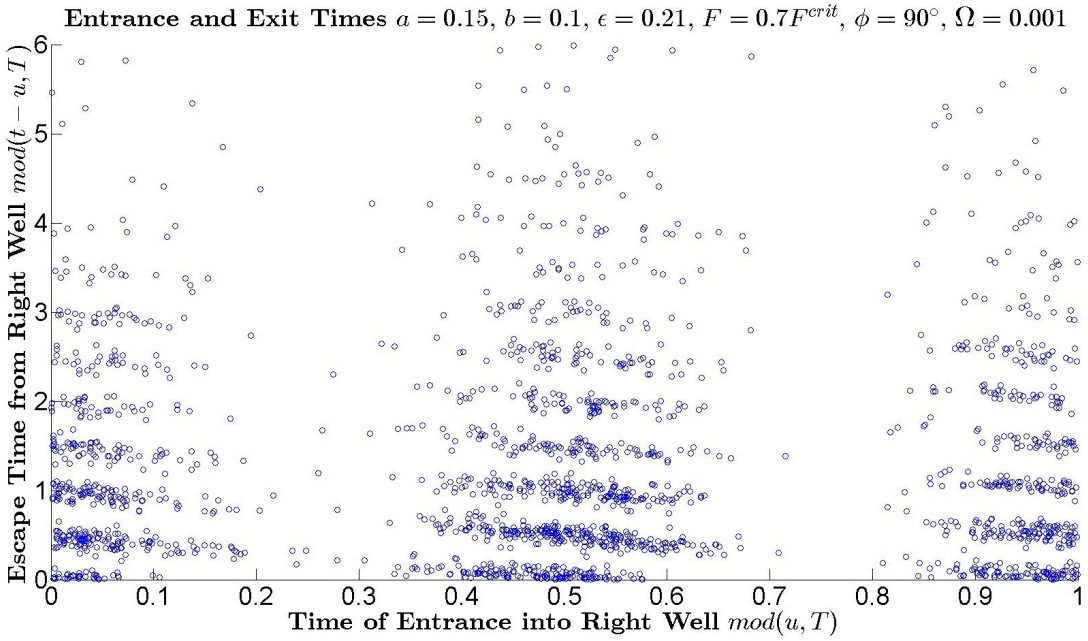

Figure 7.21: The u is the time of entrance into the well and t is the time of exit from the well.

Notice the general behaviour of the data for $mod(u,T)$ and $mod(t-u,T)$. For the $\phi = 0^\circ$ case the wells are alternating and one well is higher than the other. Entrance into the left

well tend to occur near $u = 0.5$ and entrance into the right well tend to occur near $u = 0$ and $u = 1$. For $\phi = 90°$ the wells are synchronised and are always at the same height as each other. Entrance and exit to and from either well tend to occur at $u = 0$, $u = 0.5$ and $u = 1$. Notice that the Single, Intermediate and Double Frequencies can be seen in Figures 7.16, 7.17, 7.18, 7.19, 7.20 and 7.21.

Notice also in Figure 7.16 the data points are tiled near 0.5. This seems to suggest that the use of the Dirac delta function to approximate $p_{tot} \approx p_+(t, 0)$ (see Chapter 3.4.2) may not be very good. The main problem here is the fact that we do not have an explicit formula for a probability measure of the time of entrance into a well, that is we do not have expressions for $m_-(u)$ and $m_+(u)$. This motivates us into developing the conditional KS test.

We want to test whether the escape times we have measured are really distributed by the conditional PDFs $p_-(t, u)$ and $p_+(t, u)$. This is testing the conditional null hypothesis. Define the conditional CDFs by

$$F_u^-(t) = \int_u^t p_-(s, u)\, ds = 1 - \exp\left\{ - \int_u^t R_{-1+1}(s)\, ds \right\}$$

$$F_u^+(t) = \int_u^t p_+(s, u)\, ds = 1 - \exp\left\{ - \int_u^t R_{+1-1}(s)\, ds \right\}.$$

The time coordinates of the entrance and exit from the wells are collected. These are

$$\begin{pmatrix} u_1 & u_2 & \dots & u_n \\ t_1 & t_2 & \dots & t_n \end{pmatrix}$$

where u_i is the time coordinate of the ith entrance into a well and t_i is the time coordinate of the ith exit from a well. The conditional KS statistic is calculated by

$$S_n^- = \sup_{x \in [0,1]} \left\| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{[0,x]} \left(F_{u_i}^-(t_i) - x \right) \right\|$$

$$S_n^+ = \sup_{x \in [0,1]} \left\| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{[0,x]} \left(F_{u_i}^+(t_i) - x \right) \right\|$$

where in S_n^- we sum over the time coordinates of entrance and exit to and from the left well and in S_n^+ we sum over the time coordinates of entrance and exit to and from the right well. Recall that if the conditional null hypothesis is true then S_n^- and S_n^+ are asymptotically distributed by

$$\lim_{n \longrightarrow \infty} P(\sqrt{n} S_n \leq x) = Q(x) \quad \text{where} \quad Q(x) = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2 x^2}.$$

We want 99% confidence. Note that

$$P\left(\sqrt{n} S_n \leq 1.6920 \right) = Q(1.6920) = 0.99.$$

The $Q\left(\sqrt{n}S_n\right)$ is also calculated. The smaller $Q\left(\sqrt{n}S_n\right)$ is the more certain we are in accepting the null hypothesis. A selection of some of the data being implemented with the conditional KS test are given below for various angles of the forcing ϕ and noise level ϵ. These are examples of the KS test being implemented for the histograms of escape times just given in Figures 7.12, 7.13, 7.14 and 7.15

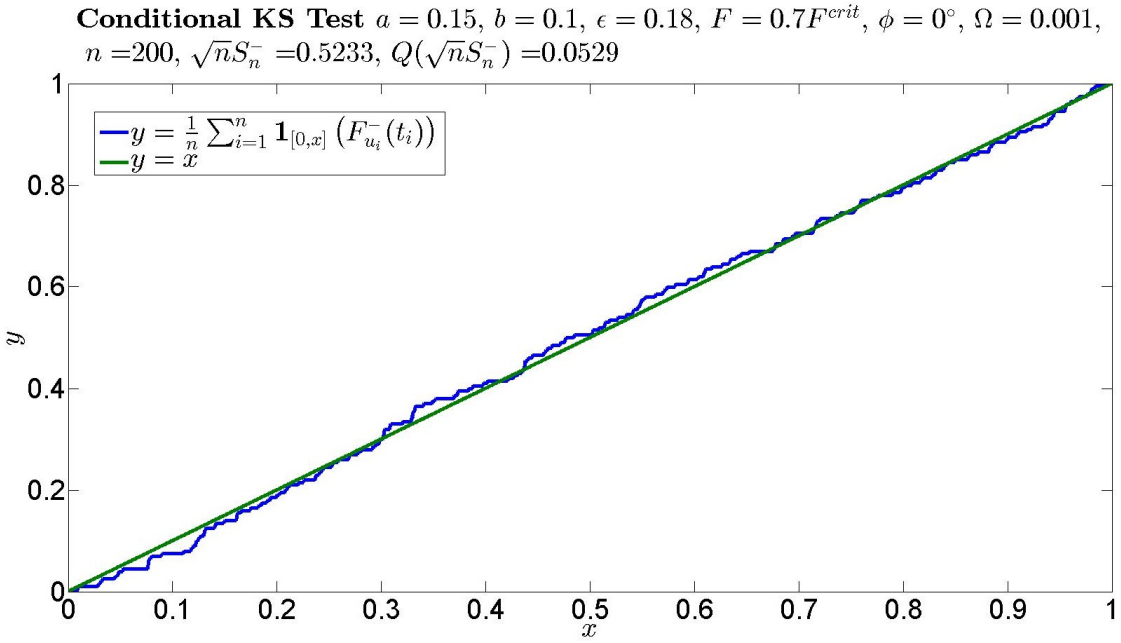

Figure 7.22: This is an example of the conditional KS test being implemented for the data in Figure 7.12. Note that $\epsilon = 0.18$, $\phi = 0°$, $n = 200$, $\sqrt{n}S_n^- = 0.5233$ and $Q\left(\sqrt{n}S_n^-\right) = 0.0529$.

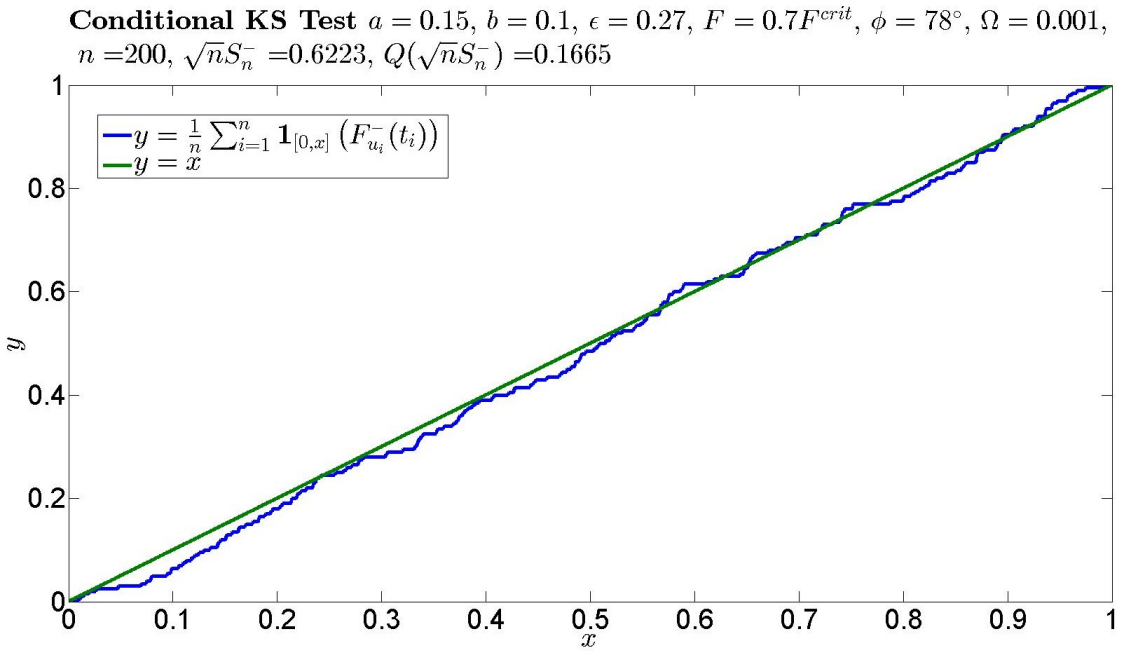

Figure 7.23: This is an example of the conditional KS test being implemented for the data in Figure 7.13. Note that $\epsilon = 0.20$, $\phi = 84°$, $n = 200$, $\sqrt{n}S_n^- = 0.6223$ and $Q\left(\sqrt{n}S_n^-\right) = 0.1665$.

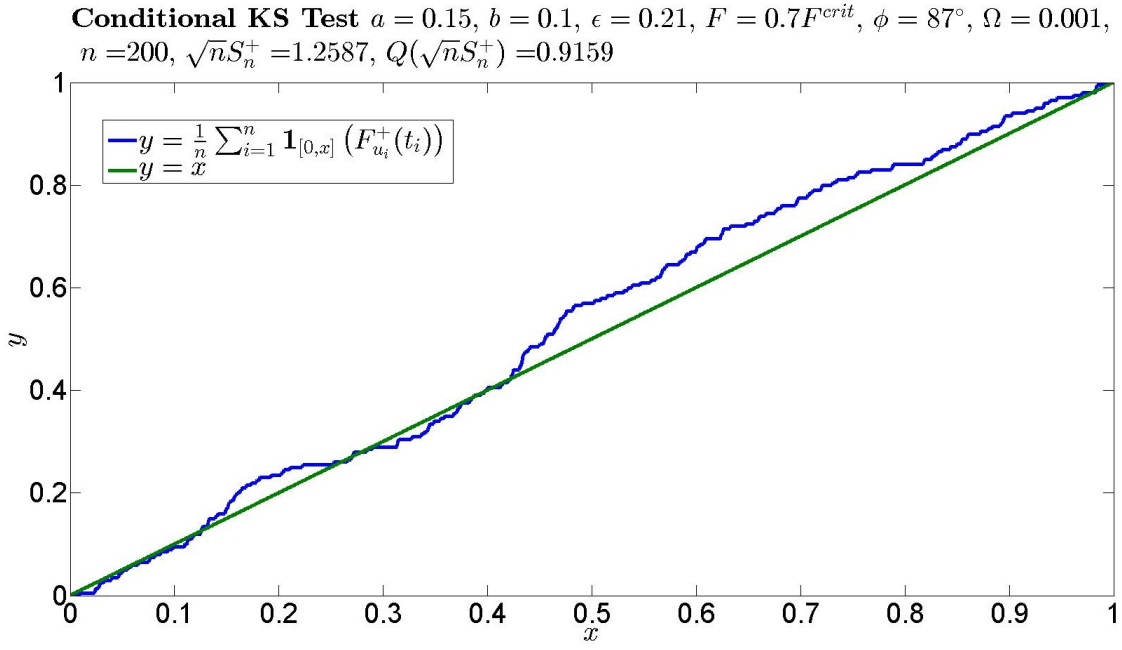

Figure 7.24: This is an example of the conditional KS test being implemented for the data in Figure 7.14. Note that $\epsilon = 0.21$, $\phi = 87°$, $n = 200$, $\sqrt{n}S_n^- = 1.2587$ and $Q\left(\sqrt{n}S_n^+\right) = 0.9159$.

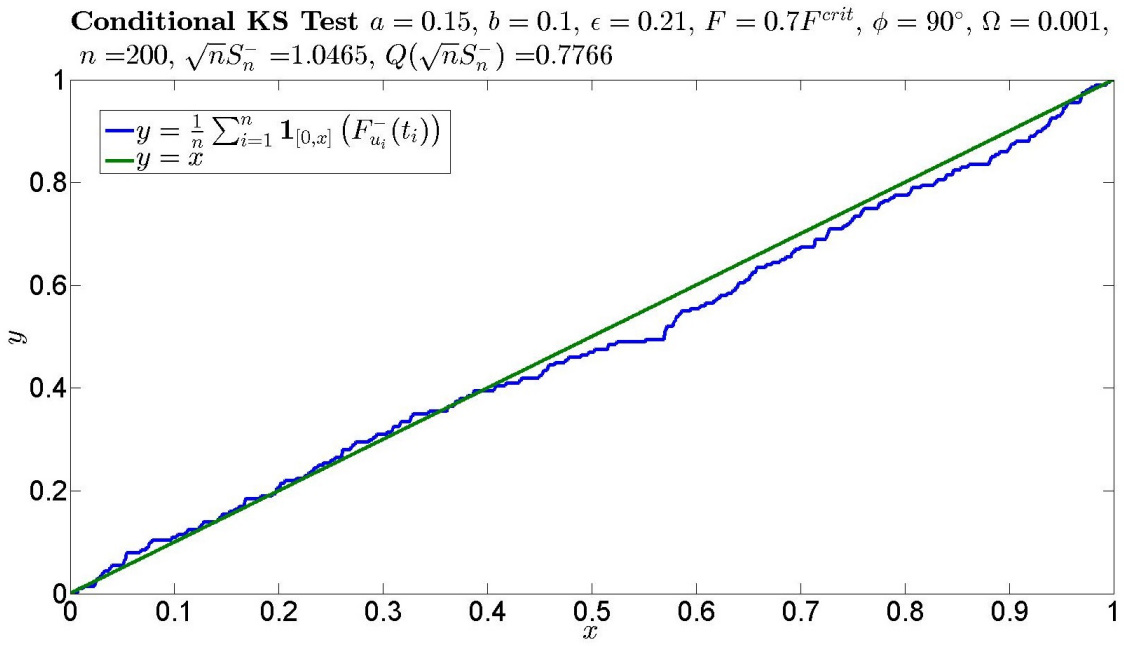

Figure 7.25: This is an example of the conditional KS test being implemented for the data in Figure 7.15. Note that $\epsilon = 0.21$, $\phi = 90°$, $n = 200$, $\sqrt{n}S_n^- = 1.0465$ and $Q\left(\sqrt{n}S_n^-\right) = 0.7766$.

7.3.1 Interpretation of the Escape Time and Conditional KS Test Analysis

When $\phi = 0°$ there were peaks in the empirical PDF of the escape times. These occurred at times $\frac{1}{2}T$, $\frac{3}{2}T$, $\frac{5}{2}T$, ... This effect we call the Single frequency. When $\phi = 90°$ the peaks occurred at $\frac{1}{2}T$, $\frac{3}{2}T$, $\frac{5}{2}T$, ... and 0, T, $2T$, $3T$, $4T$, ... This effect we call the Double Frequency. When $0° < \phi < 90°$ an intermediate effect is seen. There were major peaks at $\frac{1}{2}T$, $\frac{3}{2}T$, $\frac{5}{2}T$, ... and minor peaks at 0, T, $2T$, $3T$, $4T$.

The behaviour of the Single, Intermediate and Double Frequencies can be explained geometrically. When the height between a well and a saddle is minimum, the optimal probability of escape has occurred. When $\phi = 0°$ the frequency of the return of the optimal probability of escape is the same as the driving frequency Ω. This optimal probability comes back very T which is once in a period. When $\phi = 90°$ the frequency of the return of the optimal probability of escape is double the driving frequency at 2Ω. This optimal probability comes back very $\frac{T}{2}$ which is twice in a period. This explains why the peaks in the Single and Double Frequencies are seen where they are.

As the angle changed from $\phi = 0°$ to $\phi = 90°$ the Single Frequency gradually changes into the Double Frequency with the Intermediate Frequency seen in between. Thus the angle of the forcing is leaving a mark in the PDFs of escape times.

When the conditional KS test was implemented, the functions

$$y_0(x) = x, \quad y_-(x) = \sum_{i=1}^{n} \mathbf{1}_{[0,x]}\left(F_{u_i}^-(t_i)\right) \quad \text{and} \quad y_+(x) = \sum_{i=1}^{n} \mathbf{1}_{[0,x]}\left(F_{u_i}^+(t_i)\right)$$

were used to calculate the following distances which are the conditional KS statistics

$$S_n^- = \|y_0 - y_-\|_\infty \quad \text{and} \quad S_n^+ = \|y_0 - y_+\|_\infty .$$

It is reasonable to say that $y_-(\cdot)$ and $y_+(\cdot)$ were close enough to $y_0(\cdot)$ that we can accept the conditional null hypothesis. This can be seen and judged graphically with S_n^- and S_n^+ calculated as well. This is an example of the conditional KS test giving a reasonable result.[28]

7.4 Sparse Data Analysis

We do the same analysis with the escape time and the conditional KS test. But now we artificially make the data sparse by only implementing the conditional KS test for 20 transitions. We want 99% confidence. Thus with $n = 20$ tables for the KS distribution show that

$$P\left(S_{20} \leq 0.356\right) = 0.99$$

and there are two particular examples we want to focus on. These are when $p_{tot} \approx p_+(t, 0)$ is not a good approximation and the conditional KS test is performed in such a situation.

[28]See Appendix C.1 for discussions as to how some of our implementation of the conditional KS test are examples of oversampling.

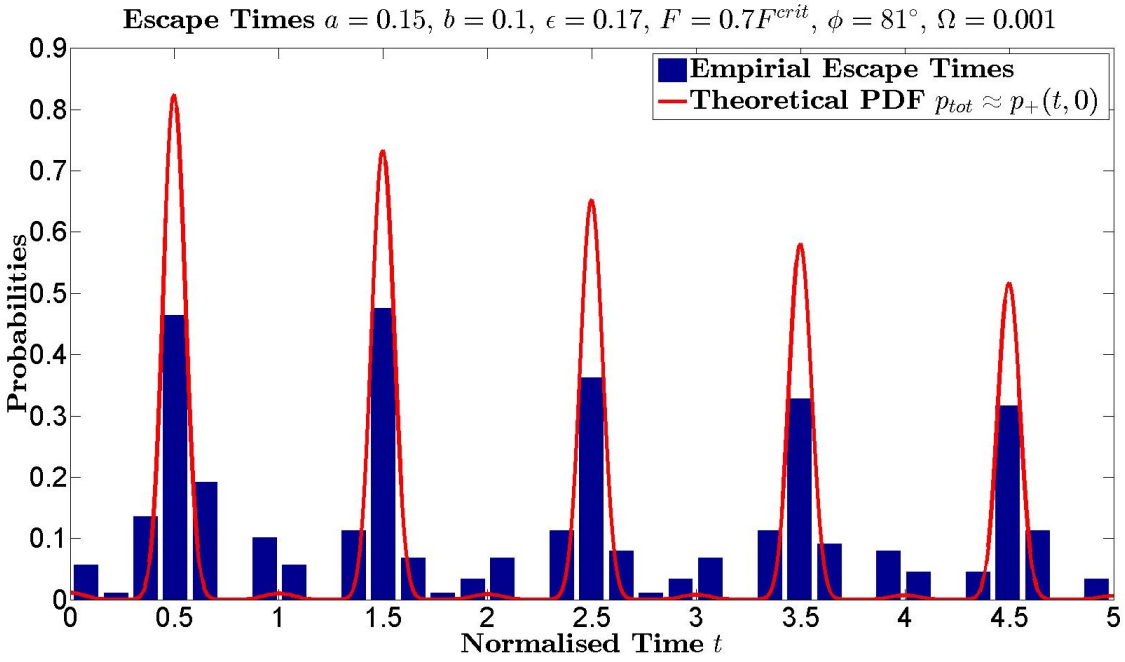

Figure 7.26: The $p_{tot} \approx p_+(t, 0)$ is not a good approximation here.

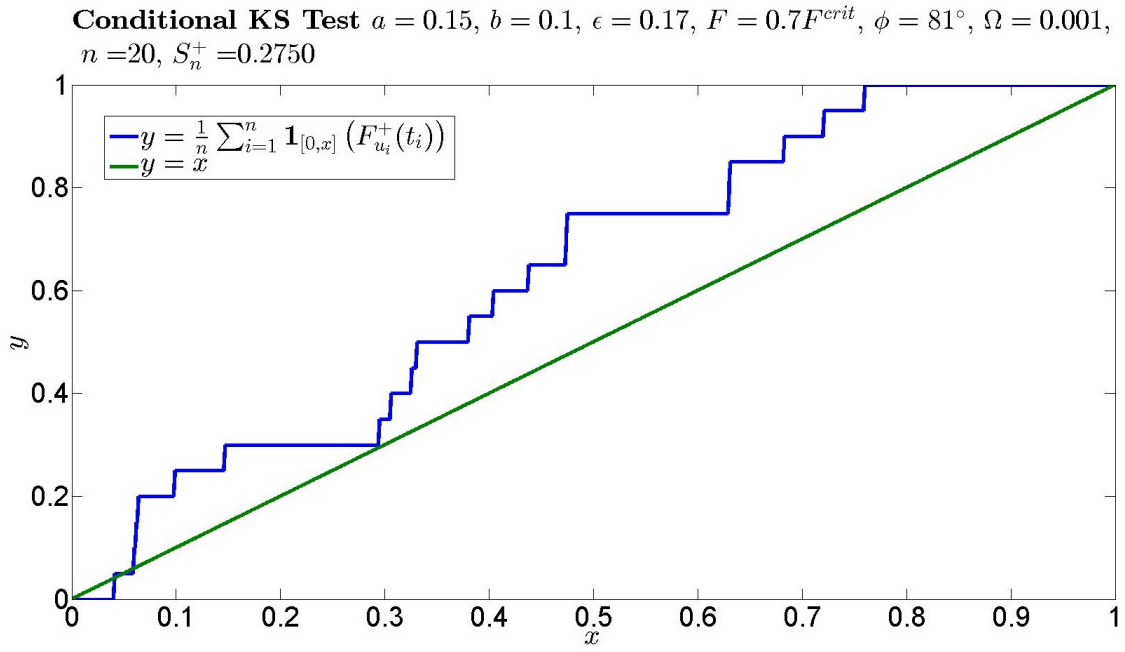

Figure 7.27: This is a KS test on the data in Figure 7.26. The conditional null hypothesis can be reasonably accepted. Note that $\epsilon = 0.17$, $\phi = 81°$, $n = 20$ and $S_n^+ = 0.2750$. $Q(\sqrt{n}S_n^+) = 0.9029$.

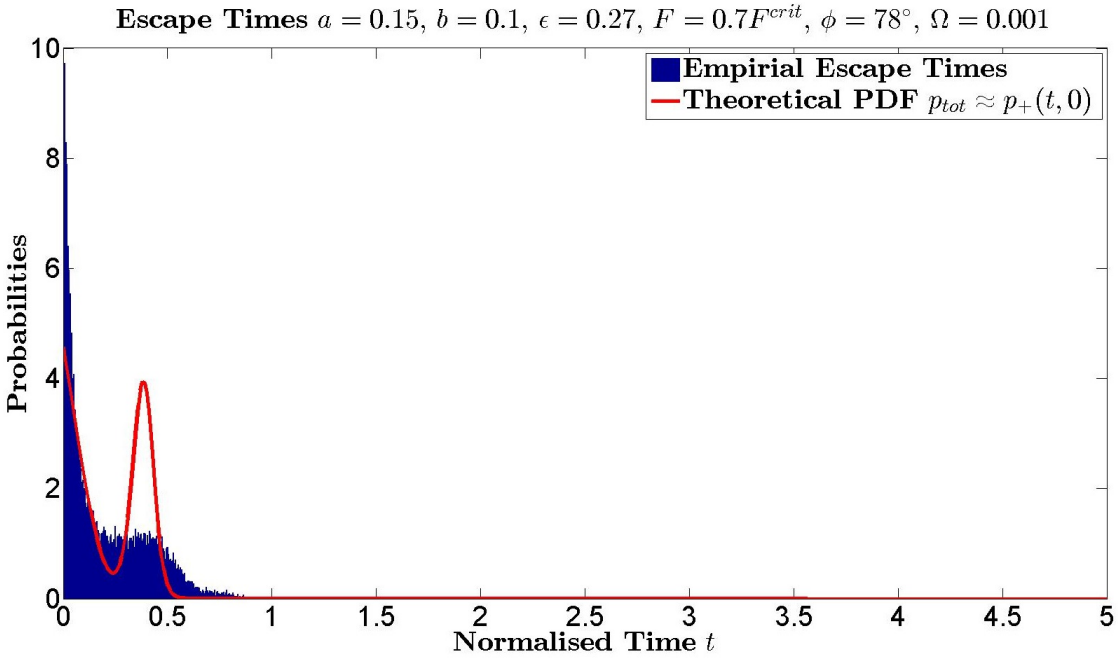

Figure 7.28: The $p_{tot} \approx p_+(t, 0)$ is not a good approximation here.

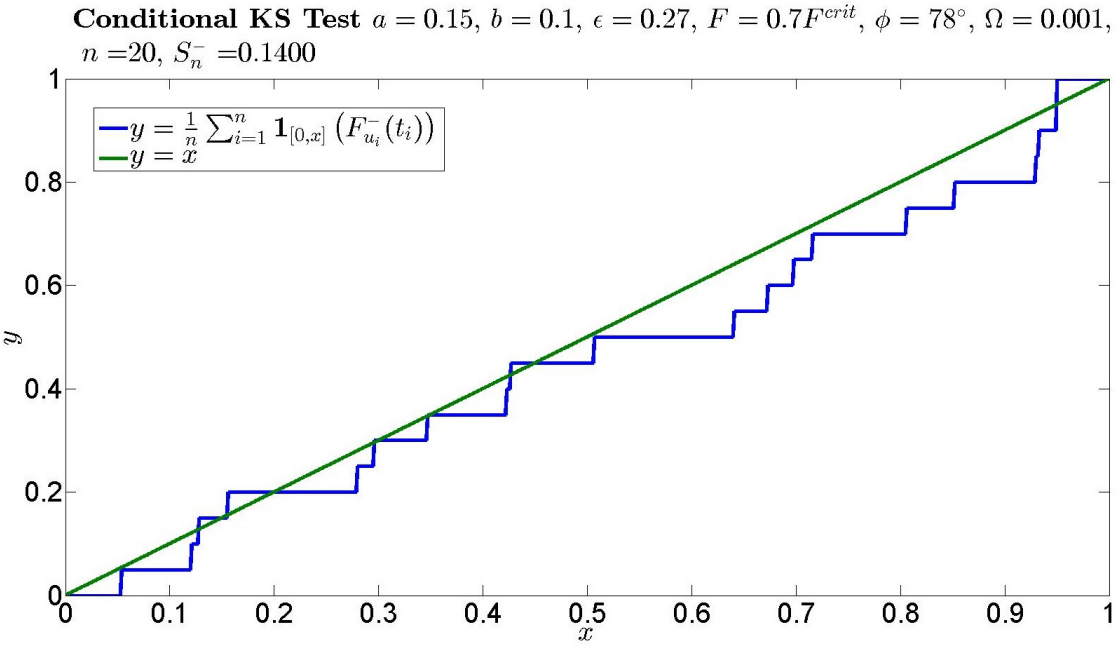

Figure 7.29: This is a conditional KS test on the data in Figure 7.28. The conditional null hypothesis can be reasonably accepted. Note that $\epsilon = 0.27$, $\phi = 78°$, $n = 20$ and $S_n^- = 0.1400$. Also note that $Q(\sqrt{n}S_n^-) = 0.1720$.

7.4.1 Interpretation of the Sparse Data Analysis

The aim of Sparse Data Analysis is to see how the conditional KS test performs even if less data is available. This is done by looking at two cases where $p_{tot} \approx p_+(t,0)$ is not a good approximation and implementing the conditional KS test on them after artificially making the data sparse.

Consider the case for the parameters in Figure 7.26. Figure 7.26 is an example of when the noise is so small there is very little escape times being detected in the range $[0, 5T]$. The $p_{tot} \approx p_+(t,0)$ is not a good approximation here. In Figure 7.27 the conditional KS test was performed on the data in 7.26 and the distance between the two functions is small. This means we can accept the conditional null hypothesis even when there is fewer data and $p_{tot} \approx p_+(t,0)$ is not a good approximation.

Now consider the case for the parameters in Figure 7.28. The noise is so large $p_{tot} \approx p_+(t,0)$ is no longer a good approximation. But in Figure 7.29 the conditional KS test was performed on the data in Figure 7.28. Again this is an example of us being able to accept the conditional null hypothesis even if $p_{tot} \approx p_+(t,0)$ is not a good approximation.

Only 20 escape times were implemented in the conditional KS test and the conditional null hypothesis can still be accepted with a reasonable degree of certainty. But the $p_{tot} \approx p_+(t,0)$ was not a good approximation for the empirical PDF of escape times. These are examples of the conditional KS test giving reasonable conclusions even when there are sparse data. It also shows that the conditional KS test can still be used even if there is no good approximation of the PDF of escape times.

Back in Chapter 3.4.2 we approximated $m_-(u)$ and $m_+(u)$ by

$$m_-(u) \approx \delta\left(u - T/2\right)$$

$$m_+(u) \approx \frac{1}{2}\delta\left(u\right) + \frac{1}{2}\delta\left(u - T\right).$$

Although the escape times (represented as dots on a scatter graph) tend to cluster around $u = 0$, $u = 0.5$ and $u = 1$ there are spread around them. As the noise levels ϵ increases the spread around $u = 0$, $u = 0.5$ and $u = 1$ would increase and $p_{tot} \approx p_+(t,0)$ would stop to be a good approximation. Despite this the conditional KS test still shows sensible results, in that we can accept the conditional null hypothesis.

7.5 Remarks on Analysis of Stochastic Resonance

There are a few subtleties and setbacks to the analysis which is worth mentioning here.

7.5.1 Remarks on Implementing the Conditional KS Test

Notice that all the theories developed about the KS Test were based on the assumption that the null hypothesis is true. This means strictly speaking a small KS statistic, that is a small S_n^- or S_n^+ does not immediately allow us to accept the null hypothesis but good

reasons not to reject it. Also when there were many transitions, that is for large n, the terms $Q(\sqrt{n}S_n^-)$ and $Q(\sqrt{n}S_n^+)$ were also calculated. The smaller $Q(\sqrt{n}S_n^-)$ and $Q(\sqrt{n}S_n^+)$ are the more confidence we have in not rejecting the null hypothesis. This is because for very large n, we would expect

$$\lim_{n\to\infty} \sqrt{n}S_n^- = 0 \quad \text{and} \quad \lim_{n\to\infty} \sqrt{n}S_n^+ = 0$$

so the smaller $Q(\sqrt{n}S_n^-)$ and $Q(\sqrt{n}S_n^+)$ are the more certain we are in not rejecting the null hypothesis.

7.5.2 Remarks on Adiabatic Approximation

Notice that in the PDFs $p_-(t,u)$, $p_+(t,u)$ and $p_{tot}(t)$ expressions for the escape rates $R_{-1+1}(t)$ and $R_{+1-1}(t)$ were required. These rates were also required for the conditional KS test. Strictly speaking these rates are dependent on the driving frequency Ω, but we stress that these rates were calculated using Kramers' formula as though the particle is escaping from a static potential. This is the adiabatic approximation where an oscillatory potential is approximated by a static potential. Considerations for whether the adiabatic approximation would fail in our calculations were done back in Chapter 6.4.3

It is worth summarising all the approximations which the analysis of the data have been based. There is the small noise approximation and slow forcing approximation from Kramers' formula, the adiabatic approximation and the perfect phase approximation where p_{tot} is approximated by $p_{tot} \approx p_{tot}(t,0)$.

Conclusion

Outline of Results

In this thesis we have considered the following problem. Let X_t^ϵ be a stochastic process in $\mathbb{R}^2$ which is described by the SDE

$$dX_t^\epsilon = b\left(X_t^\epsilon, t\right) dt + \epsilon\, dW_t$$

and the drift term $b(\cdot,\cdot)$ is expressed by

$$b(x,t) = -\nabla V_0(x) + F\cos\Omega t$$

where $V_0 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is a time independent function, the unperturbed potential, with two metastable states, and two pathways between these states. The $F \in \mathbb{R}^r$ is the magnitude of the forcing and Ω is the driving frequency. Our aim was to see characteristics of the trajectory X_t^ϵ which only depends on the qualitative structure of V_0, that is the existence of two metastable states and two pathways.

For concreteness we considered a model, which we call the Mexican Hat Toy Model

$$V_0(x,y) = \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 \quad \text{where} \quad r = \sqrt{x^2+y^2}.$$

The magnitude and angle of the forcing are given by

$$F = \sqrt{F_x^2 + F_y^2} \quad \text{and} \quad \phi = \tan^{-1}\left(\frac{F_y}{F_x}\right).$$

The angle ϕ and noise level ϵ were varied. At $\phi = 0$ the wells were alternating, that is one well is higher than the other, in the sense that it is easer to jump from one well to the other than vice versa. At $\phi = 90^\circ$ the wells are synchronised, that is both wells are always at the same height but the heights of the barrier for the two paths is alternating.

A potential with two pathways has never been considered before in the context of stochastic resonance. We studied it using approximation techniques and direct simulations. In an adiabatic regime the Freidlin-Wentzell theory allows one to give analytical solutions of the jump type distributions asymptotically in this regime. This theory predicted the appearance of additional resonance peaks at half the frequency when the angle approaches $\phi = 90^\circ$.

We simulated X_t^ϵ for different values of ϕ and ϵ and computed for the values of angle increasing from $\phi = 0$ to $\phi = 90°$ the six measures M_1, M_2, M_3, M_4, M_5 and M_6 as functions of the noise level. The first major surprise was that the graphs showed less and less pronounced minima (or maxima) and hence suggests that the phenomena of stochastic resonance gets less and less pronounced, see Chapter 7.2. The effect of resonance seems to disappear overall.

However, considering the path X_t^ϵ itself, one sees that there may be nevertheless some synchronisation, see Figure 7.9, 7.10 and 7.11. We carefully controlled our simulation and checked it for consistency, see Chapter 6.4. To properly quantify synchronisation we considered the histograms of the escape times, which to our knowledge has been not considered thoroughly before. The histograms showed a clear periodicity and also the emergence of peaks at the Double Frequency for increasing angle. For a quantitative consideration we assume that the entrance time is in perfect phase (this is when $m_-(u)$ and $m_+(u)$ *can* be approximated by Dirac delta functions). This gives for several cases good quantitative and in general good qualitative agreement with the combined adiabatic and small noise approximation. A more sophisticated analysis based on a Kolmogorov-Smirnov test developed here shows that this approximation works for a larger range of parameters where the approximation of the perfect phase of the entrance time is not appropriate (this is when $m_-(u)$ and $m_+(u)$ *cannot* be approximated by Dirac delta functions) see Chapter 7.4.1. Summarizing, the theoretical and the simulation results are in very good agreement. We want to stress that in the comparison no free parameters were present and so no fitting took place.

The fact that the six measures are blind can be explained using Markov chain models approximating the SDE. As one expects from large deviation theory, for small noise and in an adiabatic regime the SDE can be approximated by a continuous time Markov chain. In this Markov chain model we showed that the invariant measures are constant when $\phi = 90°$. Hence we expect that the invariant measure gives in the diffusion case equal weights to the left and the right well. Together, this gives us the following qualitative picture of the dynamics for any angle. At a fixed time the probability that one sees a jump from the left to the right well or vice versa has the same probability. However, conditioned on the phase and the direction of the last jump, for concreteness assume that it was at phase u and from the left to the right (that is to say the particle entered the well at time u) the next jump will be at phase which is near to a multiple of $T/2$ (that is to say the particle will leave the well near the times $t = nT/2$ where n is an integer). The jump rates will be given by the height of the potential barriers.

At $\phi = 90°$, the path X_t^ε and $-X_t^\varepsilon$ will appear with the same probability if one starts in the invariant measure. This explains why the six measures are all insensitive in this case. The equilibration happens because the process will skip some of the jump opportunities and in this way the left-right synchronization will get lost quickly.

This new phenomena we discovered has added an additional motivation to the observation of Hermann, Imkeller, Pavlyukevich, Berglund and Gentz that the appropriate consideration has to be on the path level. Averaged quantities like the six measures can be very misleading and masking the real behaviour of the system. The escape time dis-

tribution shows a clear signal of stochastic resonance in accordance with the theoretical consideration. The presence of a two pathways manifests itself in an appearance of peaks at the Double Frequency. We showed that adiabatic small noise approximation gives a good statistical model. We demonstrated that this appearance can be detected also when only a limited number of transitions is available. Our analysis provides us with a clear footprint indicating the existence of a second pathway. The angle dependence of our result should also allow us to predict the orientation of the saddles with respect to the wells.

Further Studies

The invariant measures studied in this thesis are for a simple two state model. One could try to generalise this to continuous states, that is a space-time phase PDF for the position of the particle could be derived.

The conditional KS test gives us confidence that one could develop statistical inference, using maximum likelihood for example, to develop a statistic test to estimate the basic parameters of the system, if they are unknown to us. Instead of the approximation $p_{tot} \approx p_+(t,0)$ used in parts of the consideration, a better approximation may be found by studying the PDFs of $m_-(u)$ and $m_+(u)$ theoretically and statistically.

A theory beyond adiabatic approximation may be developed for very slow to fast frequencies. Higher order of approximation to the escape times than Kramers' formula could be studied. Analytic and theoretical developments to go beyond adiabatic approximation and potential theory may be a real mathematical challenge. But experimental simulations may provide an idea of what this new theory may be like.

Appendix A

Conventions in Defining the SDEs, Potential, Time Dependency and Forcing

This is more of a clarification on the notation being used. In this thesis only two toy model potentials are studied. These in their most unperturbed forms are denoted by

$$\begin{aligned} V_0(x) &= \frac{x^4}{4} - a\frac{x^2}{2} \\ V_0(x,y) &= \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 \end{aligned}$$

where $r = \sqrt{x^2+y^2}$, $a > 0$ and $b > 0$ which when perturbed by a force are denoted by

$$\begin{aligned} V_F(x) &= \frac{x^4}{4} - a\frac{x^2}{2} + Fx \\ &= V_0(x) + Fx \\ V_F(x,y) &= \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 + F_x x + F_y y \\ &= V_0(x,y) + F_x x + F_y y \\ &= V_0(x,y) + \mathbf{F}\cdot\mathbf{x} \end{aligned}$$

and when given a periodic forcing are denoted by

$$\begin{aligned} V_t(x,F) &= \frac{x^4}{4} - a\frac{x^2}{2} - Fx\cos\Omega t \\ &= V_0(x) - Fx\cos\Omega t \\ V_t(x,y,F_x,F_y) &= \frac{1}{4}r^4 - \frac{1}{2}r^2 - ax^2 + by^2 - F_x x\cos\Omega t - F_y y\cos\Omega t \\ &= V_0(x,y) - F_x x\cos\Omega t - F_y y\cos\Omega t \\ &= V_0(x,y) - \mathbf{F}\cdot\mathbf{x}\cos\Omega t. \end{aligned}$$

This is so that the SDEs can be written in the form

$$dX_t^\epsilon = -\nabla V_t\, dt + \epsilon dw$$

which when expanded can be written as

$$dx = \left[-\frac{\partial V_0}{\partial x} + F_x \cos \Omega t \right] dt + \epsilon\ dw_x$$
$$dy = \left[-\frac{\partial V_0}{\partial y} + F_y \cos \Omega t \right] dt + \epsilon\ dw_y$$

meaning more details about the system can be quickly seen in the notation. This also implies that the SDEs are always defined with a negative forcing. When the critical points of the system are being studied (in Chapter 5 for example) we can study the critical points with a positive force and V_F would be an appropriate notation to use. Using V_0, V_F and V_t may seem like an abuse of notation, but if anything specific is being referred to, we can denote $V_{F=F^{crit}}$ for example. When the most general expression for a potential V is being used, it should be deduced from context whether $V = V_0$, $V = V_F$ or $V = V_t$ is being referred to.

Note also that for a stochastic process in $\mathbb{R}^r$ which is described by the SDE

$$\dot{X}_t^\epsilon = -\nabla V + F\cos(\Omega t) + \epsilon \dot{W}_t$$

and the magnitude of the forcing is sometimes denoted by

$$F = \sqrt{F_1^2 + F_2^2 + \ldots + F_r^2}.$$

Again this may seem like an abuse of notation, but it should be clear from context whether F is a vector or scalar.

Appendix B

Further Numerical Methods

B.1 Numerical Methods for measuring Escape Times

The Markov Chain takes the values $Y_t^\epsilon = \pm 1$. But in simulations time is discrete with a time step t_{step}, that is

$$0, t_{step}, 2t_{step}, \ldots, Nt_{step}.$$

The reduction from the diffusion X_t^ϵ to the Markov Chain at the $(n+1)$th time step is actually given by

$$Y^\epsilon_{(n+1)t_{step}} = \begin{cases} -1 & \text{if} \quad \left| X^\epsilon_{nt_{step}} - w_l(nt_{step}) \right| < R \\ +1 & \text{if} \quad \left| X^\epsilon_{nt_{step}} - w_r(nt_{step}) \right| < R \\ Y^\epsilon_{nt_{step}} & \text{if otherwise} \end{cases}$$

which is slightly different from the way Y_t^ϵ was defined in Chapter 3 (see page 37). This is so that the definition of Y_t^ϵ was easier to write down theoretically, such that the sets

$$\{t : |X_t^\epsilon - w_l(t)| \leq R\} \quad \text{and} \quad \{t : |X_t^\epsilon - w_r(t)| \leq R\}$$

are compact sets given the continuity of X_t^ϵ, $w_l(t)$ and $w_r(t)$. This meant

$$Y_t^\epsilon = \begin{cases} -1 & \text{if} \quad |X_t^\epsilon - w_l(t)| \leq R \\ +1 & \text{if} \quad |X_t^\epsilon - w_r(t)| \leq R \\ Z & \text{if neither} \end{cases}$$

then Y_t^ϵ would be easier to define for $t \notin \{t : |X_t^\epsilon - w_l(t)| \leq R\} \cup \{t : |X_t^\epsilon - w_r(t)| \leq R\}$. But alternatively if we had

$$Y_t^\epsilon = \begin{cases} -1 & \text{if} \quad |X_t^\epsilon - w_l(t)| < R \\ +1 & \text{if} \quad |X_t^\epsilon - w_r(t)| < R \\ Z & \text{if neither} \end{cases}$$

then

$$\{t : |X_t^\epsilon - w_l(t)| < R\} \quad \text{and} \quad \{t : |X_t^\epsilon - w_r(t)| < R\}$$

would be open sets and Y_t^ϵ would be harder to define for $t \notin \{t : |X_t^\epsilon - w_l(t)| < R\} \cup \{t : |X_t^\epsilon - w_r(t)| < R\}$ which is the neither case. Nevertheless the simulations should gloss out all these details.

B.2 Numerical Methods for calculating Fourier Transform and Linear Response

Fourier Transforms are involved in finding the linear response. The trajectory of the particle is in theory a continuous object, but in practice when simulations are done it is a finite discrete object. The exact mechanism of obtaining the linear response from a simulated trajectory is now being discussed.

When the trajectory is being numerically realised it is a finite discrete set. Let the x (or y) coordinate of the particle at time nt_{step} where $0 \leq n \leq (N-1)t_{step}$ be denoted by $X_{nt_{step}}$. This gives rise to the set

$$\begin{aligned} X &= \left\{X_0, X_{t_{step}}, X_{2t_{step}}, X_{3t_{step}}, \ldots, X_{(N-1)t_{step}}\right\} \\ &= \{x_0, x_1, x_2, x_3, \ldots, x_{N-1}\} \end{aligned}$$

where $x_n = X_{nt_{step}}$ etc. Notice that time is discrete here. When this is Discrete Fourier Transformed (being quickly implemented by the Fast Fourier Transform algorithm) it is denoted by

$$\begin{aligned} \tilde{X} &= \left\{\tilde{X}_0, \tilde{X}_{\omega_{step}}, \tilde{X}_{2\omega_{step}}, \tilde{X}_{3\omega_{step}}, \ldots, \tilde{X}_{(N-1)\omega_{step}}\right\} \\ &= \{\tilde{x}_0, \tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \ldots, \tilde{x}_{N-1}\} \end{aligned}$$

where $\tilde{x}_n = \tilde{X}_{n\omega_{step}}$ etc and the transform is given by

$$\tilde{x}_k = \sum_{n=0}^{N-1} x_n e^{-2\pi ikn/N}$$

and the following relation is used

$$\omega_{step} = \frac{1}{(N-1)t_{step}}$$

which is the highest detectable frequency divided by the number of steps. If we want to find the linear response at driving frequency Ω, then Ω needs to be approximated by a finite number of ω_{step} as in

$$\frac{\Omega}{2\pi} \approx n\omega_{step}$$

and the linear response at this driving frequency is then given by

$$X_{lin}^{\Omega} = 2 \times \left| \tilde{X}_{n\omega_{step}} \right|.$$

Notice the factor of 2 being used here. Suppose that the trajectory can be approximated by

$$X_t^{\epsilon} \approx A\cos(\Omega t + \phi)$$

then a good approximate expression for A and ϕ would be

$$A \approx X_{lin}^{\Omega} \quad \text{and} \quad \phi \approx \arg\left(\tilde{X}_{n\omega_{step}}\right) = \tan^{-1}\left\{\frac{\text{Im}\left(\tilde{X}_{n\omega_{step}}\right)}{\text{Re}\left(\tilde{X}_{n\omega_{step}}\right)}\right\}$$

where ϕ is the angle of the complex number $\tilde{X}_{n\omega_{step}}$.

B.3 Numerical Methods for calculating M_5 and M_6

Here we present how we computed M_5 and M_6 numerically. This is how M_5 and M_6 are calculated in theory

$$M_5 = \int_0^T \phi^-(t) \ln\left(\frac{\phi^-(t)}{\overline{\nu}_-(t)}\right) + \phi^+(t) \ln\left(\frac{\phi^+(t)}{\overline{\nu}_+(t)}\right) dt$$
$$M_6 = \int_0^T -\overline{\nu}_-(t) \ln \overline{\nu}_-(t) - \overline{\nu}_+(t) \ln \overline{\nu}_+(t)\, dt$$

where

$$\phi^-(t) = \begin{cases} 1 & \text{if} \quad mod(t,T) \leq T/2 \\ 0 & \text{if} \quad mod(t,T) > T/2 \end{cases}$$

$$\phi^+(t) = \begin{cases} 0 & \text{if} \quad mod(t,T) \leq T/2 \\ 1 & \text{if} \quad mod(t,T) > T/2. \end{cases}$$

When the invariant measures are generated numerically they are finite discrete objects described by

$$\nu_- = \left\{\nu_1^-, \nu_2^-, \cdots, \nu_N^-\right\}$$
$$\nu_+ = \left\{\nu_1^+, \nu_2^+, \cdots, \nu_N^+\right\}.$$

The real invariant measure were close to zero sometimes and in the numerical approximation they became actually zero or even negative which lead to numerical artefacts. Note that

$$\lim_{x \longrightarrow 0} \ln\left(\frac{1}{x}\right) = \infty \quad \text{and} \quad \lim_{x \longrightarrow 0} x \ln(x) = 0.$$

Define

$$\nu_{-}^{lim} = \min_{\substack{i=1,2,\ldots,N \\ \nu_i^- > 0}} \left\{ \nu_1^-, \nu_2^-, \cdots, \nu_N^- \right\}$$
$$\nu_{+}^{lim} = \min_{\substack{i=1,2,\ldots,N \\ \nu_i^+ > 0}} \left\{ \nu_1^+, \nu_2^+, \cdots, \nu_N^+ \right\}.$$

The quantities M_5 and M_6 are computed numerically in the following way

$$M_5 = \sum_{\substack{i \leq \frac{N}{2} \\ \nu_i^- > 0}} t_{step} \ln\left(\frac{1}{\nu_i^-}\right) + \sum_{\substack{i \leq \frac{N}{2} \\ \nu_i^- \leq 0}} t_{step} \ln\left(\frac{1}{\nu_{-}^{lim}}\right) + \sum_{\substack{i > \frac{N}{2} \\ \nu_i^+ > 0}} t_{step} \ln\left(\frac{1}{\nu_i^+}\right) + \sum_{\substack{i > \frac{N}{2} \\ \nu_i^+ \leq 0}} t_{step} \ln\left(\frac{1}{\nu_{+}^{lim}}\right)$$

$$M_6 = \sum_{\substack{i=1,2,\cdots,N \\ \nu_i^- > 0}} \nu_i^- \ln(\nu_i^-)(-t_{step}) \quad + \quad \sum_{\substack{i=1,2,\cdots,N \\ \nu_i^+ > 0}} \nu_i^+ \ln(\nu_i^+)(-t_{step}).$$

Appendix C

Further Commentary on Sparse Data Analysis

C.1 Examples of Oversampling

Subjectively one may think that Figures 7.24 and 7.25 are so bad the conditional null hypothesis may be rejected. This is actually an example of oversampling, where too many transitions were used in the implementation of the conditional KS test. We know the PDF we are fitting is not the real PDF but an approximation in the limit of small noise and adiabatic forcing. Hence if one has enough data points this should be picked up and the conditional KS test will refuse the approximate PDF as it will pick up even slight deviation from the real PDF. When $n = 20$ are used we have the following.

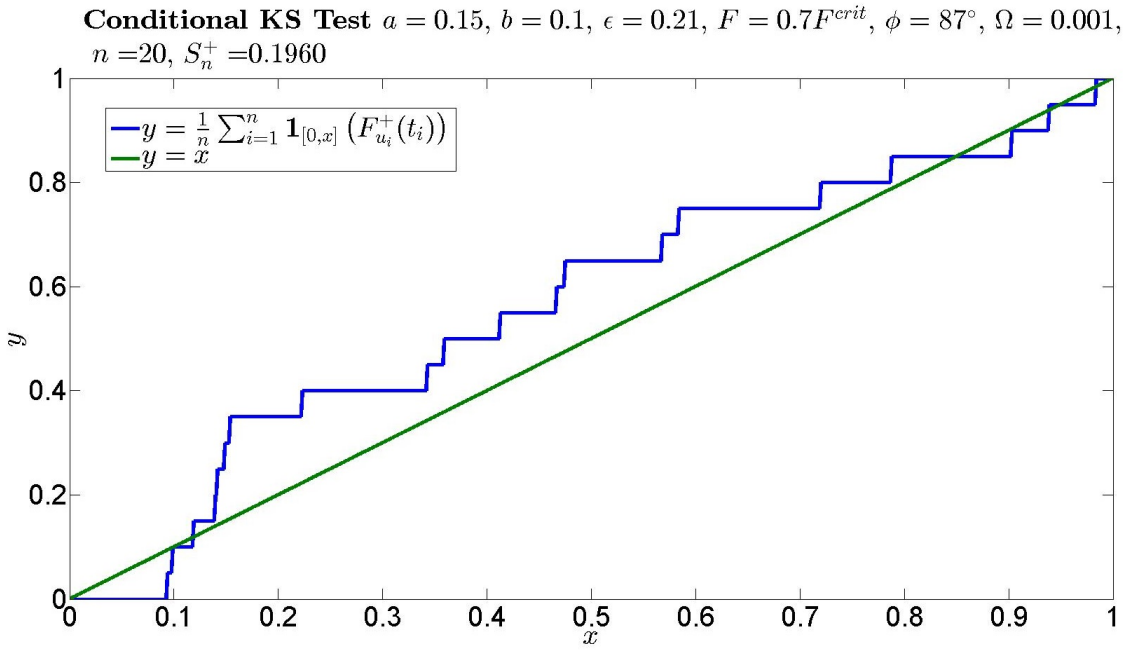

Figure C.1: This is Figure 7.24 redone with 20 transitions. Note that $\epsilon = 0.21$, $\phi = 87°$, $n = 20$, $S_n^+ = 0.1960$.

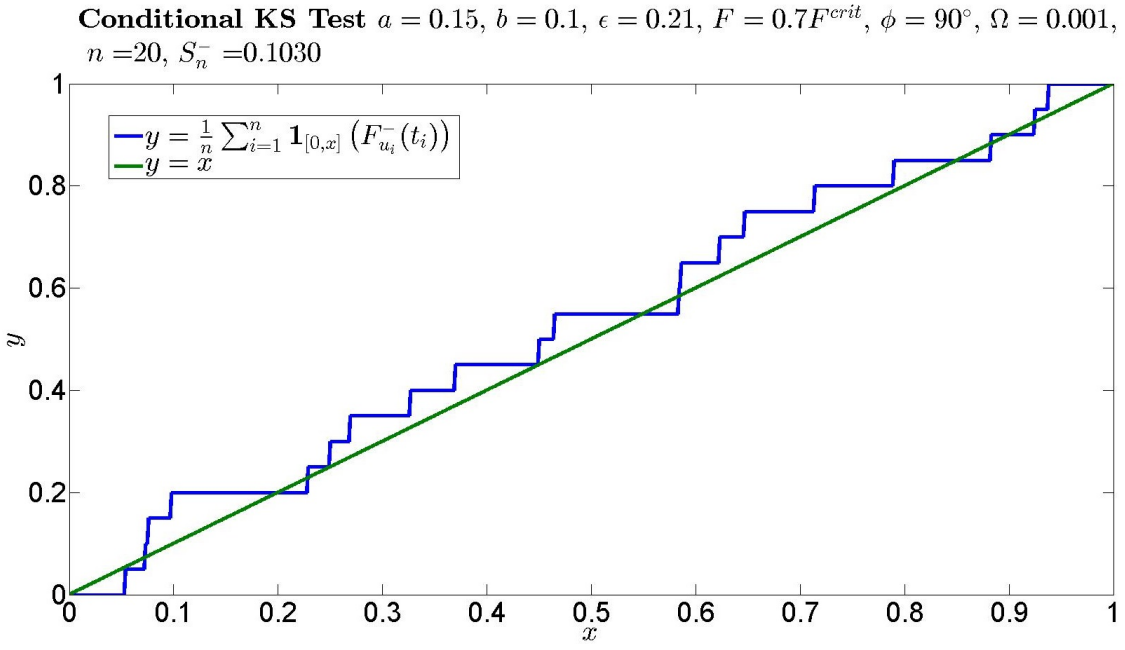

Figure C.2: This is Figure 7.25 redone with 20 transitions. Note that $\epsilon = 0.21$, $\phi = 90°$, $n = 20$, $S_n^- = 0.1030$.

C.2 Empirical CDF

Consider Figure 7.27. Notice that the empirical CDF is on top on the $y = x$ line. There were enough data to give 10 more realisations of the random variable S_n^+. Note that all ten of these S_n^+ with $n = 20$ were calculated from 200 transitions divided into ten sets for the ten S_n^+. This meant 10 more versions of the Figure 7.27 were plotted. Out of these 10 plots, one had the empirical CDF to the bottom of the $y = x$ line and one had roughly half the empirical CDF above and below the $y = x$ line. The noise level was very low at $\epsilon = 0.17$, which meant the escape times were very long with a very large spread, which gave rise to data looking unreasonable. Only 200 transitions were detected which is significantly less than other parameters, which meant only 10 realisations of the S_n^+ random variable was possible. No further conclusions are drawn here.

References

[1] A. Neiman, A. Silchenko, V. Anishchenko, and L. Schimansky-Geier, "Stochastic resonance: Noise-enhanced phase coherence," *Physical Review E*, vol. 58, no. 6, p. 7118, 1998.

[2] B. Shulgin, A. Neiman, and V. Anishchenko, "Mean switching frequency locking in stochastic bistable systems driven by a periodic force," *Physical Review Letters*, vol. 75, no. 23, p. 4157, 1995.

[3] N. Berglund and B. Gentz, "A sample-paths approach to noise-induced synchronization: Stochastic resonance in a double-well potential," *Annals of Applied Probability*, pp. 1419–1470, 2002.

[4] R. Benzi, A. Sutera, and A. Vulpiani, "The mechanism of stochastic resonance," *Journal of Physics A: Mathematical and General*, vol. 14, no. 11, p. L453, 1981.

[5] C. Nicolis and G. Nicolis, "Stochastic aspects of climatic transitionsadditive fluctuations," *Tellus*, vol. 33, no. 3, pp. 225–234, 1981.

[6] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani, "A theory of stochastic resonance in climatic change," *SIAM Journal on applied mathematics*, vol. 43, no. 3, pp. 565–578, 1983.

[7] B. McNamara, K. Wiesenfeld, and R. Roy, "Observation of stochastic resonance in a ring laser," *Phys. Rev. Lett.*, vol. 60, pp. 2626–2629, Jun 1988.

[8] L. Guidoni, R. Mannella, V. Isaia, P. Verkerk, and E. Arimondo, "Stochastic resonance in a laser with saturable absorber," *Il Nuovo Cimento D*, vol. 17, no. 7, pp. 803–810, 1995.

[9] J. Grohs, S. Apanasevich, P. Jung, H. Issler, D. Burak, and C. Klingshirn, "Noise-induced switching and stochastic resonance in optically nonlinear cds crystals," *Phys. Rev. A*, vol. 49, pp. 2199–2202, Mar 1994.

[10] A. Simon and A. Libchaber, "Escape and synchronization of a brownian particle," *Phys. Rev. Lett.*, vol. 68, pp. 3375–3378, Jun 1992.

[11] S. Fauve and F. Heslot, "Stochastic resonance in a bistable system," *Physics Letters A*, vol. 97, no. 1, pp. 5 – 7, 1983.

[12] R. N. Mantegna and B. Spagnolo, "Stochastic resonance in a tunnel diode," *Phys. Rev. E*, vol. 49, pp. R1792–R1795, Mar 1994.

[13] R. N. Mantegna and B. Spagnolo, "Stochastic resonance in a tunnel diode in the presence of white or coloured noise," *Il Nuovo Cimento D*, vol. 17, no. 7, pp. 873–881, 1995.

[14] R. N. Mantegna and B. Spagnolo, "Noise enhanced stability in an unstable system," *Phys. Rev. Lett.*, vol. 76, pp. 563–566, Jan 1996.

[15] I. Lin and J.-M. Liu, "Experimental observation of stochastic resonance like behavior of autonomous motion in weakly ionized rf magnetoplasmas," *Physical Review Letters*, vol. 74, no. 16, p. 3161, 1995.

[16] A. N. Grigorenko, P. I. Nikitin, A. N. Slavin, and P. Y. Zhou, "Experimental observation of magnetostochastic resonance," *Journal of Applied Physics*, vol. 76, no. 10, 1994.

[17] G. Debnath, T. Zhou, and F. Moss, "Remarks on stochastic resonance," *Phys. Rev. A*, vol. 39, pp. 4323–4326, Apr 1989.

[18] L. Gammaitoni, F. Marchesoni, E. Menichella-Saetta, and S. Santucci, "Multiplicative stochastic resonance," *Phys. Rev. E*, vol. 49, pp. 4878–4881, Jun 1994.

[19] L. Gammaitoni, M. Martinelli, L. Pardi, and S. Santucci, "Observation of stochastic resonance in bistable electron-paramagnetic-resonance systems," *Phys. Rev. Lett.*, vol. 67, pp. 1799–1802, Sep 1991.

[20] A. Longtin, A. Bulsara, and F. Moss, "Time-interval sequences in bistable systems and the noise-induced transmission of information by sensory neurons," *Phys. Rev. Lett.*, vol. 67, pp. 656–659, Jul 1991.

[21] A. D. Hibbs, A. L. Singsaas, E. W. Jacobs, A. R. Bulsara, J. J. Bekkedahl, and F. Moss, "Stochastic resonance in a superconducting loop with a josephson junction," *Journal of Applied Physics*, vol. 77, no. 6, 1995.

[22] R. Rouse, S. Han, and J. E. Lukens, "Flux amplification using stochastic superconducting quantum interference devices," *Applied Physics Letters*, vol. 66, no. 1, 1995.

[23] P. E. Greenwood, L. M. Ward, D. F. Russell, A. Neiman, and F. Moss, "Stochastic resonance enhances the electrosensory information available to paddlefish for prey capture," *Phys. Rev. Lett.*, vol. 84, pp. 4773–4776, May 2000.

[24] J. A. Freund, L. Schimansky-Geier, B. Beisner, A. Neiman, D. F. Russel, T. Yakusheva, and F. Moss, "Behavioral stochastic resonance: How the noise from a daphnia swarm enhances individual prey capture by juvenile paddlefish," *Journal of Theoretical Biology*, vol. 214, no. 1, pp. 71 – 83, 2002.

[25] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani, "Stochastic resonance in climatic change," *Tellus*, vol. 34, no. 1, pp. 10–16, 1982.

[26] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani, "A theory of stochastic resonance in climatic change," *SIAM Journal on Applied Mathematics*, vol. 43, no. 3, pp. 565–578, 1983.

[27] G. Vemuri and R. Roy, "Stochastic resonance in a bistable ring laser," *Phys. Rev. A*, vol. 39, pp. 4668–4674, May 1989.

[28] T. Zhou and F. Moss, "Analog simulations of stochastic resonance," *Phys. Rev. A*, vol. 41, pp. 4255–4264, Apr 1990.

[29] T. Zhou, F. Moss, and P. Jung, "Escape-time distributions of a periodically modulated bistable system with noise," *Phys. Rev. A*, vol. 42, pp. 3161–3169, Sep 1990.

[30] R. Löfstedt and S. N. Coppersmith, "Stochastic resonance: Nonperturbative calculation of power spectra and residence-time distributions," *Phys. Rev. E*, vol. 49, pp. 4821–4831, Jun 1994.

[31] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, "Stochastic resonance," *Rev. Mod. Phys.*, vol. 70, pp. 223–287, Jan 1998.

[32] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*. New York, NY: Springer US, 1984.

[33] M. V. Day, "On the exponential exit law in the small parameter exit problem," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 8, no. 4, pp. 297–323, 1983.

[34] Y. I. Kifer, "Certain results concerning small random perturbations of dynamical systems," *Theory of Probability & Its Applications*, vol. 19, no. 3, pp. 487–505, 1975.

[35] A. Galves, E. Olivieri, and M. E. Vares, "Metastability for a class of dynamical systems subject to small random perturbations," *The Annals of Probability*, vol. 15, no. 4, pp. 1288–1305, 1987.

[36] E. Olivieri and M. E. Vares, *Large deviations and metastability*. Cambridge University Press, 2005.

[37] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, "Metastability in reversible diffusion processes i: Sharp asymptotics for capacities and exit times," *Journal of the European Mathematical Society*, vol. 6, no. 4, pp. 399–424, 2004.

[38] N. Berglund and B. Gentz, "The eyring-kramers law for potentials with nonquadratic saddles," *Markov Processes and Related Fields*, vol. 16, no. 3, pp. 549–598, 2010.

[39] N. Berglund, "Kramers' law: Validity, derivations and generalisations," *Markov Processes and Related Fields*, vol. 19, no. 3, pp. 459–490, 2011.

[40] M. I. Freidlin, "Quasi-deterministic approximation, metastability and stochastic resonance," *Physica D: Nonlinear Phenomena*, vol. 137, no. 34, pp. 333 – 352, 2000.

[41] S. Herrmann, P. Imkeller, I. Pavlyukevich, and D. Peithmann, *Stochastic Resonance: A Mathematical Approach in the Small Noise Limit*, vol. 194. American Mathematical Soc., 2013.

[42] M. V. Day, "Some phenomena of the characteristic boundary exit problem," *Diffusion processes and related problems in analysis*, vol. 1, pp. 55–71, 1990.

[43] N. Berglund and B. Gentz, "Universality of first-passage-and residence-time distributions in non-adiabatic stochastic resonance," *EPL (Europhysics Letters)*, vol. 70, no. 1, p. 1, 2005.

[44] N. Berglund and B. Gentz, "On the noise-induced passage through an unstable periodic orbit ii: General case," *SIAM Journal on Mathematical Analysis*, vol. 46, no. 1, pp. 310–352, 2014.

[45] I. Pavlyukevich, *Stochastic Resonance*. PhD thesis, Humboldt University Berlin, 2002.

[46] P. Imkeller and I. Pavlyukevich, "Model reduction and stochastic resonance," *Stochastics and Dynamics*, vol. 2, no. 4, pp. 463–506, 2002.

[47] P. Imkeller and I. Pavlyukevich, "Stochastic resonance in two-state markov chains," *Archiv der Mathematik*, vol. 77, no. 1, pp. 107–115, 2001.

[48] P. Imkeller and I. Pavlyukevich, "Stochastic resonance: a comparative study of two-state models," in *Seminar on Stochastic Analysis, Random Fields and Applications IV*, pp. 141–154, Springer, 2004.

[49] S. Herrmann, P. Imkeller, and I. Pavlyukevich, *Two Mathematical Approaches to Stochastic Resonance*, pp. 327–351. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.

[50] S. Herrmann, P. Imkeller, and D. Peithmann, "Large deviations for diffusions with time periodic drift and stochastic resonance," *HU Berlin and U Nancy*, 2005.

[51] S. Herrmann and P. Imkeller, "The exit problem for diffusions with time-periodic drift and stochastic resonance," *Ann. Appl. Probab.*, vol. 15, pp. 39–68, 02 2005.

[52] R. Mannella, "Integration of Stochastic Differential Equations on a Computer," *International Journal of Modern Physics C*, vol. 13, pp. 1177–1194, 2002.

[53] I. I. Gihman and A. V. Skorohod, *Stochastic Differential Equations*. Berlin: Springer, 1972.

[54] S. Herrmann, P. Imkeller, and D. Peithmann, "Transition times and stochastic resonance for multidimensional diffusions with time periodic drift: A large deviations approach," *Ann. Appl. Probab.*, vol. 16, pp. 1851–1892, 11 2006.

[55] A. N. Kolmogorov, "Sulla Determinazione Empirica di una Legge di Distribuzione," *Giornale dell'Istituto Italiano degli Attuari*, vol. 4, pp. 83–91, 1933.

[56] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *Ann. Math. Statist.*, vol. 19, pp. 279–281, 06 1948.

[57] W. Feller, "On the kolmogorov-smirnov limit theorems for empirical distributions," *Ann. Math. Statist.*, vol. 19, pp. 177–189, 06 1948.

[58] J. F. Monahan, "Evaluating the smirnov distribution function," *Center for Research in Scientific Computation*, vol. 89, p. 2, 1989.

[59] G. Marsaglia, W. W. Tsang, and J. Wang, "Evaluating kolmogorov's distribution," *Journal of Statistical Software*, vol. 8, no. 1, pp. 1–4, 2003.

[60] R. Simard and P. L'Ecuyer, "Computing the two-sided kolmogorov-smirnov distribution," *Journal of Statistical Software*, vol. 39, no. 1, pp. 1–18, 2011.

[61] E. Platen and N. Bruti-Liberati, *Numerical solution of stochastic differential equations with jumps in finance*, vol. 64. Springer Science & Business Media, 2010.

[62] C. Moler, "Cleve's corner, random thoughts, 10^{435} years is a very long time," *Matlab News & Notes*, pp. 12–13, Fall 1995.