OSCILLATIONS IN DELAY DIFFERENTIAL EQUATIONS: COMPUTATIONAL PERSPECTIVES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Mark Walth

August 2025

OSCILLATIONS IN DELAY DIFFERENTIAL EQUATIONS: COMPUTATIONAL PERSPECTIVES

Mark Walth, Ph.D.

Cornell University 2025

Delay differential equations are wily beasts. Though they share much of their DNA with ordinary differential equations, their dependence on past states introduces infinite-dimensional subtleties that complicate their analysis. In this thesis, we explore some of the many challenges that arise in studying nonlinear oscillations in delay differential equations, along with techniques we have found useful for taming them.

We begin with an overview of the foundational theory, followed by thoroughly worked examples from the linear case. We then turn to nonlinear oscillations, focusing on the deceptively simple Delayed Duffing Equation. Several approaches to this equation are presented, including the Fourier-based Method of Harmonic Balance.

Next, we introduce a new technique for connecting solutions of delay equations to their undelayed counterparts, which we call the Method of Characteristic Oscillations. Finally, we take on the formidable problem of chaos in delay systems.

Our hope is that by the end of this thesis, the reader will feel well equipped to venture into the wild and fascinating world of delay equations.

BIOGRAPHICAL SKETCH

Mark Walth was born in Colorado Springs. He attained his Bachelor's in Mathematics at Reed College in 2014 and his Master's in Mathematics Education in 2015. He was a Math for America Teaching Fellow, teaching in Washington, DC public schools from 2015-2019. He came to Cornell in 2019.

To Aurora.

ACKNOWLEDGEMENTS

There are, of course, a great many people to thank who helped me to get here. I will focus on just a few.

Thank you to Mauro Camargo, Anthony Graves-McCleary, Benjamin Hoffman, David Mehrle, David Hathcock, Andrew Warren, and Rodrigo Delgado, for all of the curiosity and confusion filled conversations, which are the key to growing understanding.

Thank you to Professor Steven Strogatz for being an inspiration, both as a teacher and as a mathematician, and for all of your kind guidance.

Thank you to Professor Richard Rand. I cannot imagine a better advisor. I will always be proud to have been your student.

Thank you to my Mother, who is still the best math teacher I know.

And thank you to my wife Jennifer, for believing in me through all of it, and for helping to keep the Hausdorff spirit strong.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION AND BACKGROUND

"The past is never dead. It's not even past."

- William Faulkner

1.1 Introduction

This thesis is dedicated to the study of Delay Differential Equations (DDEs). Delay equations are of significant practical importance, as they arise in applications ranging from population dynamics to control systems engineering. However, their appeal extends beyond applications: delay systems exhibit a rich and subtle theoretical structure. For practitioners of nonlinear dynamics, they are often extraordinarily vexing, owing to their intricate behavior and the difficulty of answering even seemingly basic questions about their solutions.

Throughout this work there is a recurring example, which highlights so much of the difficulty and subtlety of delay equations, the so-called Delayed Duffing's Equation

$$\ddot{x}(t) + x(t-T) + x^3(t) = 0.$$

This innocuous-looking equation conceals a wealth of complexity. Despite being arguably the simplest nonlinear second-order delay equation one could write down, many fundamental questions about its solutions remain unresolved. We return to this equation repeatedly as a central thread through the various methods discussed in this thesis.

The overarching goal of this work is to highlight techniques we have found effective in the study of DDEs. Through a series of worked examples and methodological developments, we hope to provide the reader with useful tools—and perhaps a bit of inspiration—for productively struggling with the delay equations they may encounter in their own work.

The thesis is organized as follows:

In Section 1.2, we review the essential theory of delay equations and highlight the key features that distinguish DDEs from ordinary differential equations (ODEs).

In Chapter 2, we examine a collection of linear delay equations, using techniques that underscore how subtle and intricate even simple autonomous DDEs can be—especially in contrast to their ODE counterparts.

Chapter 3 presents the method of Harmonic Balance, which has proven particularly fruitful in analyzing oscillatory behavior in certain classes of DDEs.

Chapter 4 contains our most substantial theoretical contribution to the study of delay equations, the method of Characteristic Oscillations. In analogy to the method of characteristics from PDEs, the method identifies certain special solutions to DDEs which are governed by an ODE. We present the basic ideas of the method, illustrate them thoroughly for the Delay Duffing Equation, and conclude by using the method to address a version of Hilbert's 16th [17] problem for delay systems.

Finally, Chapter 5 explores the onset and structure of chaotic behavior in delay systems, bringing together many of the ideas developed earlier in the thesis.

1.2 Background

In general, dynamical systems study systems which change under a defined evolution rule. In ordinary differential equations, the instantaneous change of a system depends only on the current state of the system. By contrast, a the evolution of a delay system depends on the state of the system at some time in the past. Delay equations arise in many contexts across many fields, including laser dynamics, where the delay is the time it takes for light to travel between points; machine tool vibration, where delay is due to the dependence of the cutting force on the thickness of a rotating work piece; gene dynamics, where delay is due to the time required for messenger RNA to copy the genetic code and export it from the nucleus to the cytoplasm; and finance, where delay comes from the time it takes traders to determine the state of the system [27][10].

In control systems engineering, a frequent theme is a feedback control which responds to the state of the system. Frequently, such systems involve a finite time to sense and respond to information, which introduces delay into the control scheme. Examples of such delay arising in control schemes include Gantry cranes used for moving shipping containers [18], fuel injection in combustion engines [10], queuing strategies [21], among myriad other applications. In these contexts, delay can both be desirable and undesirable, sometimes stabilizing unwanted oscillations [36] and sometimes introducing unwanted oscillations [10].

With so many applications across the sciences, it is no surprise that delay equations have been the object of significant study in recent years. This thesis makes no attempt at giving a complete introduction to the topic. Rather, our goal is to present

enough of the fundamental theory so that the reader will be familiarized with the broad strokes in order to follow the contributions of this thesis.

1.2.1 Mathematical Background

A very general form a delay differential equation is

$$F(t, x(t), x(t-\tau_1), \ldots, x(t-\tau_n), x'(t), x'(t-\mu_1), \cdots, x'(t-\mu_m)) = 0 \tag{1.1}$$

for some function F, where $x(t) \in \mathbb{R}^k$ for some k, and τ_i, μ_i are positive real numbers. Even more general types of delay equations exist — for example, the delays τ_i and μ_i can themselves be functions of time. See [16] for a discussion of the most general forms of delay equations.

We will focus our attention on a smaller class of delay equations, the so called *retarded delay equations,* in which the highest order derivative term appears without delay, and is an explicit function of the other variables:

$$x'(t) = f(t, x(t), x(t-\tau_1), \ldots, x(t-\tau_n)) \tag{1.2}$$

for some function f, and where $\tau_1, \ldots, \tau_n > 0$. Retarded delay equations are in contrast *neutral* delay equations — in which delay appears in the highest order derivative term. Throughout the remainder of this thesis, when we refer to delay equations, we implicitly mean delay equations of retarded type.

It is illustrative to think about what a simple numerical scheme to solve an equation such as (1.2) would look like. Suppose that the value of x is known at a time t_0, and suppose for the sake of argument that $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_n$. Using a

simple Euler's method, we would have that

$$x(t_0 + \Delta t) \approx x(t_0) + \Delta t f(t, x(t_0), x(t_0 - \tau_1), \ldots, x(t_0 - \tau_n)),$$

so that determining the value of $x(t_0 + \Delta t)$ will require evaluating $x(t_0 - \tau_1)$ all the way through $x(t - \tau_n)$. Therefore, to perform a single update step, we need to have access to the values of x at several previous time steps. Given this, what would an initial value problem for equation (1.2) look like? Clearly it wouldn't be sufficient to specify the value of $x(0)$, as there wouldn't be enough information to perform a single update step. Moreover, even knowing the values of $x(-\tau_1), \ldots, x(-\tau_n)$ would not provide enough information to specify a unique solution, because after performing one update step, we would then need then values of $x(-\tau_1 + \Delta t), \ldots, x(-\tau_n + \Delta t)$ to find the value of $x(2\Delta t)$. Continuing with this line of thinking leads us to one of the most essential differences between DDEs and ODEs.

Key Point: The initial data necessary to specify a unique solution to the DDE (1.2) is an entire *initial history function* $\phi : [-\tau_n, 0) \to \mathbb{R}^k$.

This is in stark contrast with the ODE case, in which a unique solution is determined by a single point $x(0) \in \mathbb{R}^k$. This detail is more than just an oddity — it will turn out to fundamentally affect the structure of the solution space of delay equations.

In the next two sections, we further discuss these ideas. First, we elaborate on the basic numerical integration scheme that we laid out here, to discuss the *method of steps*, which is of fundamental importance for the numerical study of DDEs. In the following section, we discuss the nature of the solution space for DDEs, and state the precise existence and uniqueness theorem in this context.

These two perspectives reveal just how multifaceted delay equations can be. On one hand, the method of steps can make DDEs seem like little more than ODEs solved in a peculiar, piecewise fashion. On the other hand, viewing delay equations through the lens of phase space reveals their fundamentally infinite-dimensional nature—something profoundly different from the ODE case. Thus, in some contexts DDEs behave much like ODEs, while in others they are strikingly different.

The Method of Steps

The method of steps is of fundamental importance for the numerical study of delay equations. The method is a classical technique for solving delay differential equations (DDEs) by converting them into a sequence of initial value problems for ordinary differential equations (ODEs). It exploits the fact that once the solution is known on an initial interval—typically provided by a history function—it can be extended step-by-step using the known delay structure. This approach is particularly effective for DDEs with constant delays and allows for a piecewise construction of the solution over successive intervals.

We illustrate the idea with a simple example, first explored in [2]. Consider the equation

$$x'(t) = x(t-1) \tag{1.3}$$

with the initial history function $x(t) = 1$ if $0 \leq t \leq 1$. (Here we set our initial interval to $[0, 1]$ rather than $[-1, 0]$ for algebraic convenience). How shall we proceed with

such an equation? Well, on the interval $[1, 2]$, equation (1.3) becomes simply

$$\begin{aligned} x'(t) &= 1 \qquad \text{for } 1 \leq t \leq 2, \\ x(1) &= 1. \end{aligned}$$

On this interval, clearly a solution is $x(t) = t$, which can be written as $x(t) = 1 + (t - 1)$, for reasons that will become clear soon.

In a similar fashion, on the interval $[2, 3]$, the equation becomes

$$\begin{aligned} x'(t) &= 1 + (t - 1) \qquad \text{for } 2 \leq t \leq 3 \\ x(2) &= 2. \end{aligned}$$

Again, on this interval, we have a simple ODE initial value problem which can readily be solved. We find that on this interval, $x(t) = 1 + (t - 1) + \frac{(t-2)^2}{2}$.

Continuing in this way, a straight forward induction finds that $x(t)$ can be expressed for all $t \geq 1$ as

$$x(t) = \sum_{j=0}^{N} \frac{(t-j)^j}{j!} \qquad \text{for } N \leq t \leq N+1, \qquad \text{where } N = 0, 1, \ldots \tag{1.4}$$

The solution is illustrated below in Figure 1.1.

Inspecting Equation (1.4) reveals that the solution is continuous everywhere, C^1 if $t > 1$, C^2 if $t > 2$, and in general C^m if $t > m$. (Recall that the class C^m denotes those functions which are continuous with m continuous derivatives). This phenomenon is known as the smoothing of discontinuities, because the solution becomes more smooth as t increases. It is a common feature of delay equations that they become increasingly smooth as time increases.

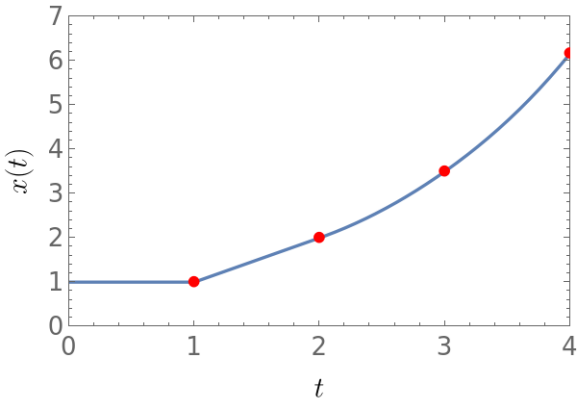

Figure 1.1: The solution to the simple delay equation (1.3) for $t \in [0, 4]$. The solution is defined in piecewise fashion, or "steps," with red dots denoting transition points. Note that at the transition points, it is possible to see the smoothing of solutions: at the first transition point, the solution is clearly continuous but not continuously differentiable; at the second transition point, the solution is differentiable but has discontinuous second derivative; etc.

This example is notable in that it admits a simple, closed-form solution. More importantly, it highlights that, in principle, solving a DDE—at least numerically—amounts to repeatedly solving ODE initial value problems over successive intervals. The upshot is that much of the existing numerical machinery for ODEs can be adapted to DDEs with relatively little modification. In this sense, DDEs may not initially seem much more complex than ODEs. That said, there are of course a number of important subtleties which arise in practice, which we do not explore here. For a detailed discussion of these challenges, see [32].

A Note on Numerical Solutions Among the many approximations schemes that have been developed for the study of nonlinear systems, numerical integration techniques rank among the most important and useful. When a nonlinear system

cannot be solved analytically, there is no way of guaranteeing that a numerical simulation will remain completely faithful to the true behavior of a system. However, there are many indicators used to determine if a numerical scheme is valid. First, in the cases when an analytic solution is available, the numerical solution should agree closely. Additionally, in cases when an analytic solution is not available, two different valid simulation methods should largely agree with each other. Numerical methods that satisfy these two criteria are sometimes referred to as "proper" simulations [9]. Throughout this thesis, we assume that the numerical techniques we use are proper simulations, and that they therefore give an accurate representation of the behavior of the systems we study. See [32] for a more detailed discussion of the numerical analysis of DDEs.

The Phase Space of Delay Equations.

In the previous section, we saw that the method of steps effectively reduces the numerical solution of DDEs to solving a sequence of ODE initial value problems on successive intervals. This resemblance, however, is deceptive: the structure of the solution space for DDEs is fundamentally different from that of ODEs.

To solve an initial value problem for a DDE such as Equation ,(1.2) one must specify an entire history function—not just a finite set of initial values. In effect, this means providing infinitely many initial data points to determine the future evolution of the solution. This requirement reflects the *infinite-dimensional* nature of the phase space for delay equations.

For comparison, consider ordinary differential equations: a first-order ODE requires

a single initial condition, and its phase space is one-dimensional. An nth-order ODE requires n initial conditions—typically values of x and its first $n-1$ derivatives—resulting in an n-dimensional phase space. In contrast, even a first-order delay differential equation demands an entire function defined over a past interval as its initial condition. Consequently, the phase space of a DDE is infinite-dimensional.

The formal study of delay differential equations necessarily takes place in the framework of Banach spaces. Foundational results—such as existence and uniqueness theorems, and continuous dependence on parameters—must be stated and proved in this setting. Likewise, familiar notions from ODE theory, such as the flow map, are generalized to define a flow on a space of functions. As a result, the mathematical theory of DDEs is significantly more complex than that of ODEs, rooted as it is in functional analysis.

More fundamentally, the infinite-dimensionality of the DDE phase space permits far more intricate dynamics than are possible in finite-dimensional ODEs. For instance, a first-order ODE has a one-dimensional phase space, which imposes strict topological constraints: all trajectories must tend to fixed points (possibly at infinity) [35]. Similarly, second-order systems are constrained by results such as the PoincarÃ©–Bendixson Theorem (see [15], Thm 1.8.1), which implies that nontrivial limit sets are restricted to fixed points, periodic orbits, or connecting trajectories like homoclinic or heteroclinic loops. In particular, chaos is impossible in second-order autonomous ODEs.

No such restrictions apply to delay equations: even the simplest first-order DDE has an infinite-dimensional phase space. As a result, complex behaviors—such as

chaos—can and do arise in scalar DDEs. One well-known example is the Mackey–Glass equation, which exhibits high-dimensional chaotic dynamics [9].

The intuitive idea behind understanding DDEs as defining a flow through function space comes from considering the whole dynamics as playing out on the interval defined by the initial history function. As time progresses, the function evolves, to give a new function on the same interval. Take for example the equation

$$x'(t) = x(t-1) \tag{1.5}$$
$$x(t) = \sin(2\pi t) \qquad \text{for } -1 \leq t \leq 0.$$

We will consider the phase space to be real valued functions on the interval $[-1, 0]$. From this perspective, the initial history $\sin(2\pi t)$ is just a single point in phase space. Now, we have seen that by integrating equation (1.5) using the method of steps, we can find a solution function $x(t)$, defined for time $t \geq 0$, the result of which is pictured in Figure (1.2).

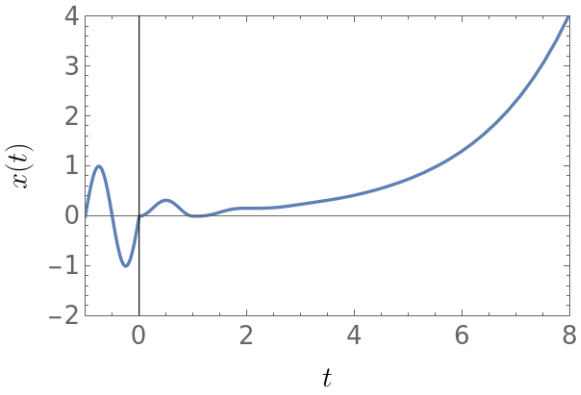

Figure 1.2: The solution to equation (1.5), obtained numerically through the method of steps.

The function space perspective instead considers that at each time, we define a

new function $x_t : [-1, 0] \to \mathbb{R}$, given by $x_t(\theta) = x(t + \theta)$. Therefore, x_0 corresponds to the initial history function $x_0(\theta) = \sin(2\pi\theta)$ for $\theta \in [-1, 0]$. Each x_t for $t > 0$ defines a new function on $[-1, 0]$, obtained by "sliding" the interval to the right, so that x_t is exactly the values of x restricted to the interval $[t - 1, t]$. More generally, when the delay is τ, x_t corresponds to the portion of the full solution on the interval $[t - \tau, t]$. The image to have in mind is that of a movie playing out on the interval $[-1, 0]$, with the initial function x_0 smoothly deforming as time t advances. This is the flow through function space. A few snapshots of the "movie" are shown below for varying values of t.

In the rest of this section, we lay out enough of this formal functional analytic framework to state some of the main theorems. A much more extensive presentation of the theory of delay equations is given in [16]. We include this section for the sake of completeness, but note that the rest of this thesis is from a more applied perspective, so this section can be safely skipped. The main takeaway for the reader is that this infinite dimensionality is always lurking in the background of delay equations.

To begin the formal discussion, fix a delay parameter τ. Consider the Banach space $C = C([-\tau, 0], \mathbb{R}^k)$ consisting of continuous functions from $[-\tau, 0] \to \mathbb{R}^k$, equipped with the supremum norm $|\phi| = \sup_{-\tau \leq \theta \leq 0} |\phi(\theta)|$. This norm induces the so called topology of uniform convergence. Let x be a continuous function, $x : \mathbb{R} \to \mathbb{R}^k$. Then for a fixed and a given $t \in \mathbb{R}$, we can define $x_t \in C$ by $x_t(\theta) = x(t + \theta)$ for $\theta \in [-\tau, 0]$. Now, let $f : \mathbb{R} \times C \to \mathbb{R}^k$ be some functional. With this notation, a retarded delay differential equation as an equation of the form

$$x'(t) = f(t, x_t). \tag{1.6}$$

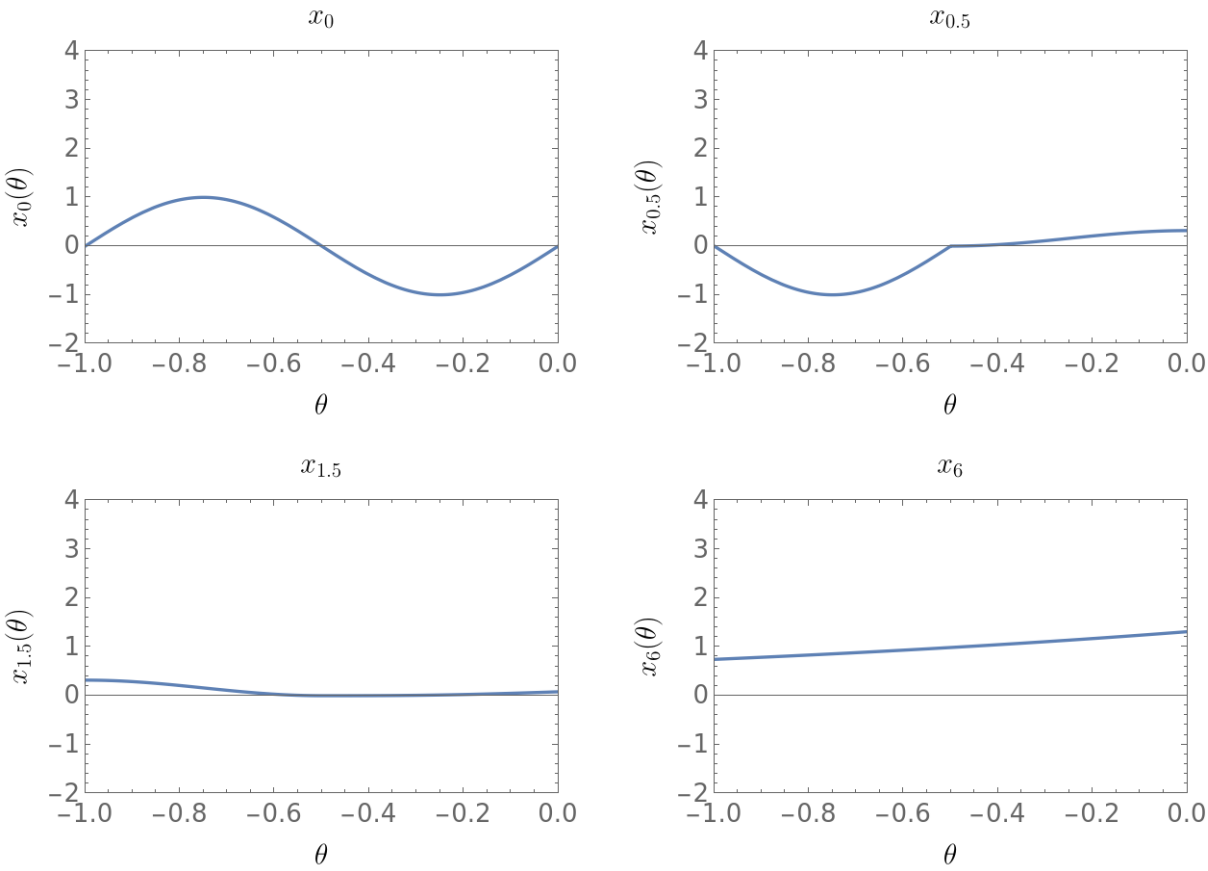

Figure 1.3: Select snapshots of the flow through function space defined by Equation (1.5). At time $t = 0$, we have the initial history function, shown in the top left panel. For each t, the function continuously evolves, giving rise to a new function on the interval $[-1, 0]$. We show the function at times $t = 0.5$ (note the discontinuity in the derivative at $\theta = 0.5$), $t = 1.5$, and $t = 6$.

This equation includes a wide variety of equations as special cases, including ordinary differential equations, integro-differential equations and other more exotic equations still. In this general setting, we can state the major theorems.

Theorem 1. *(Hale) Suppose f in equation (1.6) is continuous. Then for any $(t_0, x_0) \in \mathbb{R} \times C$, there exists a solution to equation (1.6) passing through (t_0, x_0), which we write as $\phi(t_0, x_0)$, which is continuous and defined on some open set of (t_0, x_0). If additionally the functional f is Lipschitz in its second variable, then the solution is unique.*

This is the existence and uniqueness theorem which is relevant for our context. We note that these are not the most general assumptions for which the theorem can be stated. We point the reader to [16] Chapter 2 for the full statement and the proofs (as well as what precisely is meant by "solution" in this context).

It is interesting to compare Theorem (1) to the equivalent theorem for ODEs, the Picard-Lindelöf Theorem. Indeed, the statement is essentially identical, down to the condition that f be Lipshitz in the second variable. Indeed, the proof of Theorem (1) directly mirrors the proof of the Picard-Lindelöf Theorem — the proof of Theorem 1 on the Schauder Fixed-Point Theorem to guarantee the existence of the desired solution, in much the same way as the Picard-Lindelöf Theorem applies the Banach Fixed Point Theorem. (See [16] ch 2 and [15] ch 1 details. See [39] ch 4 and 5 for a discussion of the Fixed Point Theorems). The similarity in the statements and proofs of the two theorems highlights the curious fact that in some ways, delay equations are something of a halfway between ODEs and PDEs. Owing to the method of steps, aspects of delay equations can be reduced to the theory of ODEs. But, like PDEs,

the phase space of a DDE is generally a function space. This tension is part of the delight of studying DDEs — aspects of our intuition from ODEs carries over well, but nonetheless there are always infinite dimensional subtleties and surprises to keep you on your toes!

Other results akin to the familiar foundational theorems of ODEs exist for DDEs as well, such as the continuous dependence of solutions on parameters, continuation of solutions, and the existence of a flow map. All of the details can be found in [16].

CHAPTER 2

LINEAR DELAY EQUATIONS

2.1 Stability

In this section, we present some useful computational techniques in the context of the simplest type of delay equations, autonomous linear equations. This discussion is far from exhaustive; we refer readers to [2] and [16] for a thorough discussion of linear DDEs. We wish to illustrate just enough to get a working understanding of some important concepts and techniques.

For the sake of this introduction, we restrict attention to the simple class of second order linear autonomous equations, of the form

$$x''(t) + a_1 x'(t) + a_0 x + cx(t-T) = 0. \tag{2.1}$$

As in the case of ordinary differential equations, the long term behavior behavior of such an equation is determined by the stability of any fixed points. Inspired by linear ODEs, we may seek a solution of the form $Ae^{\lambda t}$. Plugging such a guess into Equation (2.1) yields the *characteristic equation*

$$\lambda^2 + a_1\lambda + a_0 + ce^{-\lambda T} = 0. \tag{2.2}$$

Solutions λ to Equation (2.2) are known as *characteristic roots.* In general, because of the transcendental exponential term, there are infinitely many characteristic roots in the complex plane. The existence of infinitely many roots is in contrast to ODEs, in which the characteristic equation is a polynomial of degree equal to the degree

of the ODE. The infinitude of roots in the DDE setting is reflective of the infinite dimensional phase space, and is one of the main sources of complexity encountered in the delay setting. The distribution of characteristic roots governs the stability of the origin, according to the following theorem.

Theorem 2. *If all of the characteristic roots have negative real part, the origin is exponentially stable. If one of the characteristic roots has positive real part, the origin is exponentially unstable.*

A proof of Theorem 2 can be found in [16] and [2].

Understanding the distribution of roots is therefore critical to understanding the behavior of the system (2.1). Even in relatively simple cases, these roots can have a complicated structure which can be subtle to understand. In the following two examples, we illustrate some techniques which can be very useful for understanding the characteristic roots of linear DDEs.

Example 3. Consider the equation

$$x''(t) + cx(t-T) = 0.$$

To apply Theorem 2, we ansatz a solution of the form $e^{\lambda t}$. This results in the characteristic equation

$$\lambda^2 + ce^{-\lambda T} = 0. \tag{2.3}$$

As it turns out, this equation can actually be solved exactly, using the Lambert W function. Recall that the Lambert W function is the inverse of the product-exponential function, defined by $we^w = z \iff W(z) = w$. Actually, because the

function $w \mapsto we^w$ is far from injective, a single inverse function does not exist. For this reason, the Lambert W function is a multivalued, or multibranched, function. A more correct statement would therefore be $we^w = z \iff W_k(z) = w$ where k is an integer specifying a particular branch of the W function. See the standard reference [4] for a thorough discussion of the W function.

Now, we can manipulate our characteristic equation (2.3) as follows to solve for λ using the W function:

$$\begin{aligned}
\lambda^2 e^{\lambda T} &= -c \\
\lambda e^{\lambda T/2} &= \pm i\sqrt{c} \\
\frac{\lambda T}{2} e^{\lambda T/2} &= \pm \frac{i\sqrt{c}T}{2} \\
W_k\left(\frac{\lambda T}{2} e^{\lambda T/2}\right) &= W_k\left(\pm \frac{i\sqrt{c}T}{2}\right) \\
\frac{\lambda T}{2} &= W_k\left(\pm \frac{i\sqrt{c}T}{2}\right) \\
\lambda_k &= \frac{2}{T} W_k\left(\pm \frac{i\sqrt{c}T}{2}\right).
\end{aligned} \tag{2.4}$$

This gives all of the infinitely many characteristic roots, one for each $k \in \mathbb{Z}$, as a function of T and c.

The first several of these roots are shown in Figure 2.1 for $T = 1$ and $c = 1$.

It is not too hard to show that if $c > 0$ and $T > 0$, Equation (2.4) always has at least one root with positive real part. Hence, in these cases, the origin the system is unstable. To see this, we use a strategy that will prove useful in future problems. The strategy involves two steps. First, we investigate what condition must hold in order to have a characteristic root crossing the imaginary axis, which potentially

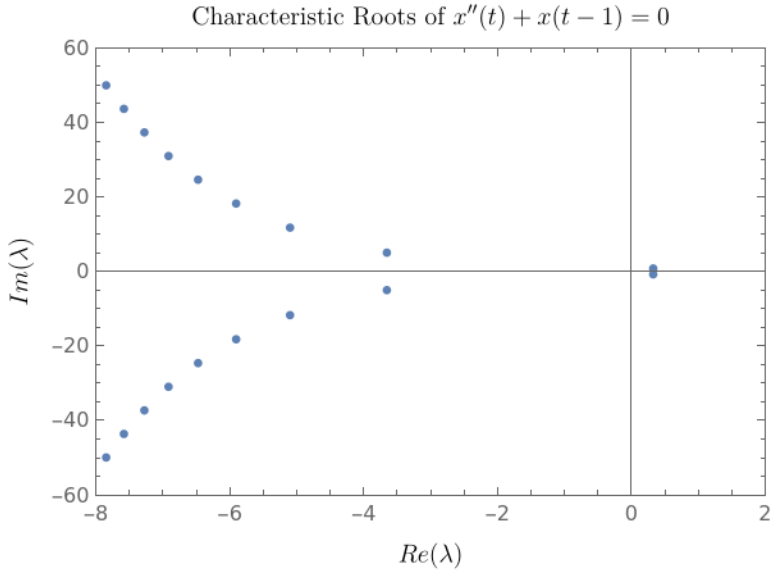

Figure 2.1: The characteristic roots of Equation (2.3) with $c = 1$ and $T = 1$. Note the pair of roots with positive real part — hence, for these parameters, the origin is unstable.

corresponds to a change in stability. Second, we investigate the derivative of the real part of the characteristic root with respect to the parameter T, to investigate if the root is crossing from left to right (stable to unstable) or if it is crossing from right to left (unstable to stable). To simplify the presentation, for this portion of the argument we take $c = 1$ and just focus on the effects of the delay T.

To carry out the strategy just described, begin by searching for a purely imaginary solution of Equation (2.3) of the form $\lambda = i\omega$, where $\omega \in \mathbb{R}$. Applying Euler's formula to simplify the result, we find that

$$-\omega^2 + (\cos(\omega T) - i\sin(\omega T)) = 0.$$

Now, the imaginary and real parts must both be equal to 0. From setting the imaginary part to 0, we find that $\omega T = n\pi$. Plugging this result into the real part,

and solving, and noting that $\cos(n\pi) = (-1)^n$, we find that

$$\left(\frac{n\pi}{T}\right)^2 = (-1)^n$$
$$T = \frac{n\pi}{\sqrt{(-1)^n}}. \tag{2.5}$$

Finally, because the delay T is assumed to be real, we conclude that n must be even, so that $T = 2m\pi$ for $m \in \mathbb{N}$. This is the condition to have a root (or in this case, a pair of roots) cross the imaginary axis. Notice that $T = 0$ is a special case. In this case, the imaginary roots are exactly $\pm i$.

For the next step, we wish to compute the direction in which the characteristic roots are moving as T increases. To this end, we write $\lambda = \alpha + i\omega$. Then equation (2.3) becomes

$$\alpha^2 - \omega^2 + ce^{-\alpha T}\cos(\omega T) + i(2\alpha\omega - ce^{-\alpha T}\sin(\omega T)) = 0.$$

Next, we implicitly differentiate with respect to T. Breaking the result into real and imaginary components, we obtain

$$-e^{-\alpha T}\left(T\frac{d\omega}{dT} + \omega\right)\sin(T\omega) + e^{-\alpha T}\left(-\alpha - T\frac{d\alpha}{dT}\right)\cos(T\omega) + 2\alpha\frac{d\alpha}{dT} - 2\omega\frac{d\omega}{dT} = 0 \tag{2.6}$$

$$2\omega\frac{d\alpha}{dT} + 2\alpha\frac{d\omega}{dT} - e^{-\alpha T}\left(-\alpha - T\frac{d\alpha}{dT}\right)\sin(T\omega) - e^{-\alpha T}\left(T\frac{d\omega}{dT} + \omega\right)\cos(T\omega) = 0. \tag{2.7}$$

Now, our goal is to solve for $\frac{d\alpha}{dT}$ when an eigenvalue is on the imaginary axis. To this end, we set the real part α to 0 in equations (2.6) and (2.7). Additionally, as just computed in equation (2.5), we set $T = 2m\pi$ and $\omega = 1$. This simplifies the above

system substantially:

$$-2m\pi\frac{d\alpha}{dT} - 2\frac{d\omega}{dT} = 0$$
$$2\frac{d\alpha}{dT} - 2m\pi\frac{d\omega}{dT} = 1.$$

From this we conclude that when an eigenvalue $\lambda = \alpha + i\omega$ crosses the imaginary axis, we have

$$\left.\frac{d\alpha}{dT}\right|_{T=2m\pi} = \frac{1}{2 + 2m^2\pi^2} > 0.$$

We conclude that eigenvalues *only ever* cross the imaginary axis from left to right. We therefore conclude that for $T > 0$, the origin is unstable.

Interestingly, we can also conclude that as T increases, more and more characteristic roots have positive real part. A new pair of roots cross the imaginary axis going from left to right whenever $T = 2m\pi$. In Figure 2.2, we can see that when $T = 200$, there are many characteristic roots with positive real part.

In a sense, the number of roots with positive real part correspond to the number of modes of the solution which do not damp down as time goes to infinity. This can be viewed as giving some kind of measure of the "effective dimension" of the system — after long times, a solution will tend to be dominated by the few modes corresponding to positive characteristic roots, each of which can be understood as giving one dimension of the solution. We see in this example that the effective dimension of the solution space tends to increase as the delay increases. This phenomenon is not specific to this example — in Chapter (5), we will see that the dimension of a chaotic attractor in phase space tends to increase as the delay increases.

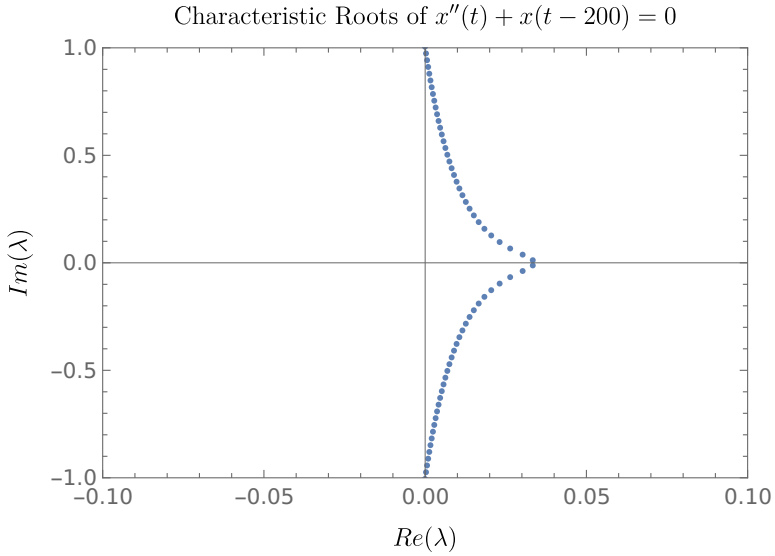

Figure 2.2: A plot of the characteristic roots of Equation (2.3) with $c = 1$ and $T = 200$. Notice that there are now many characteristic roots with positive real part. In general, as $T \to \infty$, the number of positive characteristic roots will tend to increase without bound.

The previous example gave a relatively rare case in which we could directly compute the characteristic roots in terms of a special function. Generally, this will not be the case, as illustrated by the next example.

Example 4. This example shows that even in the simple linear case, DDEs can have a very intricate stability structure. The analysis is more intricate than the previous case.

Consider the equation

$$x''(t) + x(t) + \beta x(t - T) = 0, \tag{2.8}$$

where $\beta \in \mathbb{R}$ and $T \in \mathbb{R}_{\geq 0}$. As in Example 3, the stability of the origin is determined

by the characteristic roots. The resulting characteristic equation is

$$\lambda^2 + 1 + \beta e^{-\lambda T} = 0. \tag{2.9}$$

Perhaps surprisingly, the additional "$+1$" term in the characteristic equation makes it impossible to solve exactly for the characteristic roots in terms of the W function. Instead, we begin by first studying the stability for $0 < T \ll 1$. We will then find conditions on T and β that lead to stability transitions.

To begin, suppose $0 < T \ll 1$ and $|\beta| \ll 1$. In this case, the characteristic equation is approximately $\lambda^2 + 1 + \beta(1 - \lambda T) + \mathcal{O}(T^2) = 0$. This equation is quadratic in λ, making it easy to analyze stability. We can conclude that for small T, the origin is stable for $\beta < 0$ and unstable for $\beta > 0$. The $\beta = 0$ case degenerates to the simple harmonic oscillator.

Next, we compute conditions for a stability change, when T and β are not necessarily small. To this end, we look for a purely imaginary characteristic root $\lambda = i\omega$. Plugging this into Equation (2.9) yields a real and imaginary part, both of which must be 0:

$$\begin{aligned} \text{Re}: \quad & 1 - \omega^2 + \beta\cos(\omega T) && = 0 \\ \text{Im}: \quad & -i\beta\sin(\omega T) && = 0. \end{aligned}$$

From the imaginary equation, we find that $\omega T = n\pi$ for $n \in \mathbb{Z}$. Plugging this result into the real equation, and solving for T yields the stability transition condition

$$T = \frac{n\pi}{\sqrt{1 + (-1)^n \beta}}. \tag{2.10}$$

These transition curves are shown in Figure 2.3. Notice that there are two families of curves: one family corresponding to even n and one corresponding to odd n.

Note that an additional solution comes from setting $\omega = 0$ and $\beta = -1$. This solution corresponds to a real root λ crossing the imaginary axis, rather than a pair of imaginary roots.

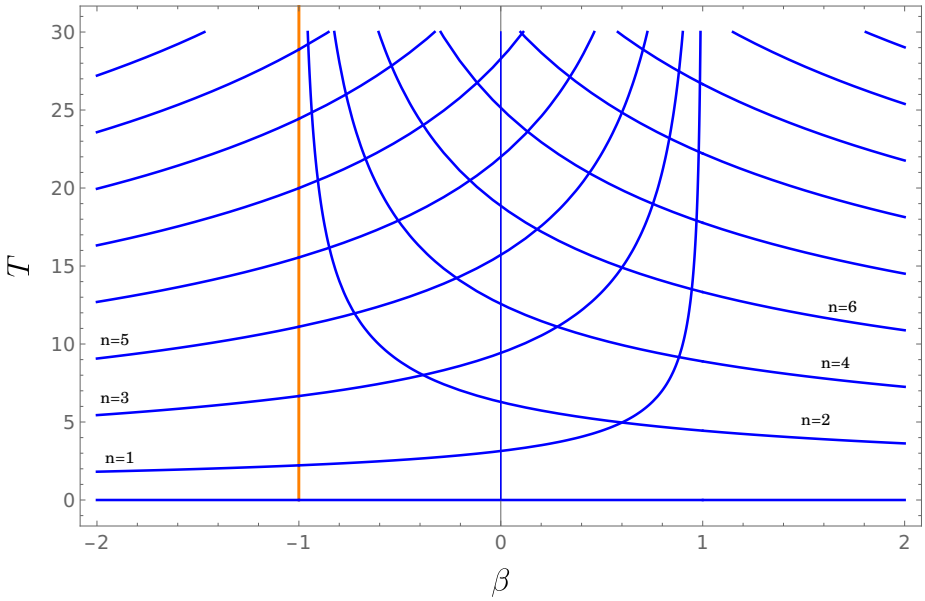

Figure 2.3: The stability transition curves in Equation (2.10). Along blue curves, a pair of complex conjugate eigenvalues cross the imaginary axis. The value of n giving rise to each curve is indicated for the first several curves. The orange vertical line at $\beta = -1$ denotes a single real root crossing the imaginary axis.

Now, having computed the curves along which an eigenvalue has a real part of 0, we next calculate if the eigenvalue is travelling from left to right (stable to unstable) or right to left (unstable to stable). To do so, as we did in Example (3), we will calculate the derivative of the real part of an eigenvalue $\lambda = \alpha + i\omega$ along the transition curves; that is to say, we will calculate $\frac{\partial \alpha}{\partial T}$ along the curves calculated in Equation (2.10) To do so, we plug in $\lambda = \alpha + i\omega$ into the characteristic equation

and set real and imaginary parts to 0. We find that

$$\alpha^2 - \omega^2 + \beta e^{\alpha(-T)}\cos(T\omega) + 1 = 0$$
$$2\alpha\omega - \beta e^{\alpha(-T)}\sin(T\omega) = 0.$$

Next, we implicitly differentiate both equations with respect to the delay, finding that

$$\beta e^{-\alpha T}\left(\alpha + T\frac{\partial\alpha}{\partial T}\right)\sin(T\omega) - \beta e^{-\alpha T}\left(T\frac{\partial\omega}{\partial T} + \omega\right)\cos(\omega T) + 2\omega\frac{\partial\alpha}{\partial T} + 2\alpha\frac{\partial\omega}{\partial T} = 0$$
$$-\beta e^{-\alpha T}\left(T\frac{\partial\omega}{\partial T} + \omega\right)\sin(T\omega) - \beta e^{-\alpha T}\left(\alpha + T\frac{\partial\alpha}{\partial T}\right)\cos(T\omega) + 2\alpha\frac{\partial\alpha}{\partial T} - 2\omega\frac{\partial\omega}{\partial T} = 0.$$

We are interested in $\frac{\partial\alpha}{\partial T}$ at a stability transition, which occurs when $\alpha = 0$. We have found that when $\alpha = 0$, $\omega T = n\pi$. Plugging in these conditions and simplifying gives

$$\frac{2n\pi}{T}\frac{\partial\alpha}{\partial T} - \beta(-1)^n\left(T\frac{\partial\omega}{\partial T} + \frac{n\pi}{T}\right) = 0$$
$$\beta(-1)^n T\frac{\partial\alpha}{\partial T} + 2\frac{n\pi}{T}\frac{\partial\omega}{\partial T} = 0.$$

At this point, we can solve for $\frac{\partial\alpha}{\partial T}$ and $\frac{\partial\omega}{\partial T}$. We find

$$\frac{\partial\alpha}{\partial T} = (-1)^n\frac{2\pi^2 n^2\beta}{4\pi^2 n^2 + \beta^2 T^4}$$
$$\frac{\partial\omega}{\partial T} = -\frac{\pi\beta^2 T^2\beta}{4\pi^2 n^2 + \beta^2 T^4}.$$

We are only really concerned with the change in the real part, α, along the transition curves found in equaiton (2.10). To this end, we focus just on the $\frac{\partial\alpha}{\partial T}$ equation. Notice that the denominator is strictly positive for all β and T. We therefore see that the sign of $\frac{\partial\alpha}{\partial T}$ alternates with n, and also depends on the sign of β. We can use this fact to calculate the **degree of instability** of the origin for various values of β and T.

Definition. The **degree of instability** of the origin is defined as the number of characteristic roots λ with $\mathrm{Re}\,(\lambda) > 0$.

Notice that the origin is stable if the degree of instability is 0. To calculate we the degree of instability in every region of the $\beta - T$ plane, we begin with the regions whose stability we already know, shown in Figure 2.3, at the bottom of the figure. When $\beta > 0$ and $T \approx 0$, we know there are two roots with positive real part, so the degree of instability is 2. When $-1 < \beta < 0$ and $T \approx 0$, we know the origin is stable, so the degree of instability is 0. From here, we color the transition curves blue or red based on whether they correspond whether $\frac{\partial \alpha}{\partial T}$ is positive (blue) or negative (red) on the curves. When crossing a blue curve in the increasing T direction, a pair of eigenvalues crosses the imaginary axis to the left, meaning the degree of instability decreases. When crossing a red curve in the increasing T direction, a pair of eigenvalues crosses the imaginary axis moving to the right, so the degree of instability increases.

Finally, a similar analysis reveals that when cross the vertical line $\beta = -1$ from left to right, a real root moves from the right half plane to the left half plane. Hence, the degree of instability is greater by one to the left of the orange vertical line.

In this way, we can fill in the degree of instability in the whole $\beta - T$ plane. The result is picture in Figure (2.4).

As the above two examples illustrate, even relatively linear simple DDEs have an extremely intricate stability structure. It will come as no surprise, then, that in nonlinear DDEs, the structure can be more vexing still. Before moving on to the

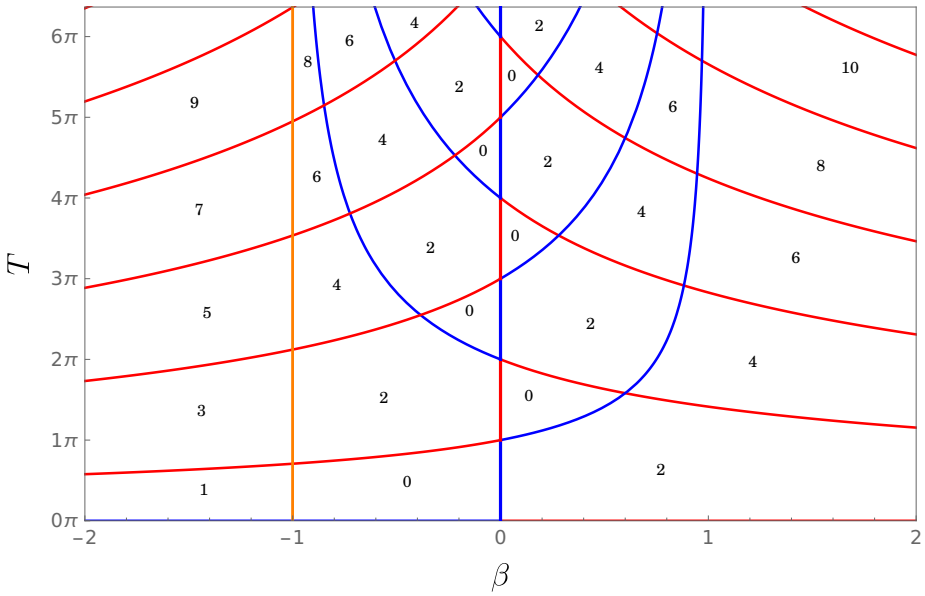

Figure 2.4: A depiction of the degree of instability of the origin in Equation (2.8). A region with instability 0 corresponds to a stable origin. When a red curve is crossed in the upwards direction, a pair of characteristic roots crosses to the right half plane, thereby increasing the degree of instability by 2. When a blue curve is crossed in the upward direction, a pair of roots crosses to the left half plane, thereby decreasing the degree of instability by 2. Finally, when the vertical orange curve is crossed in the rightward direction, a single real root crosses to the left half plane, decreasing the instability by 1.
Note the intricate distribution of the parameter regions corresponding to stability.

nonlinear case, however, we explore how we can obtain a representation of the exact solution for linear DDEs.

2.2 Solution Representations

In the previous section we have seen that the long term behavior of a linear DDE is determined by its characteristic roots. However, combing from the world of ODEs, we may be interested in more — not just the stability, but an actual representation of the solution in terms of the characteristic roots and the initial condition. After all, because the equation is linear, any linear combination of the characteristic functions $e^{\lambda t}$ will again give a solution. A natural question arises:

Question. Given a linear DDE (we will focus on second order)

$$x'' + a_1 x' + a_0 x + b_1 x'(t-T) + b_0 x(t-T) = 0 \tag{2.11}$$
$$x(t) = \phi(t) \qquad \text{for } t \in [-T, 0]$$

let Λ be the set of characteristic roots. Is it possible to write the solution as a linear combination of the characteristic functions? That is to say, do there exist coefficients a_λ such that

$$x(t) = \sum_{\lambda \in \Lambda} a_\lambda e^{\lambda t}?$$

We focus on second order linear equations as it is sufficient to show the main points without getting mired down in calculations. This question is full of subtleties, of course, because of the infinitude of characteristic roots. A complete treatment would of course require a careful study of convergence issues, etc. Nonetheless, it is a nice exercise which we show here, as it does elucidate some nice connections.

The key idea is to exploit the *Laplace transform,* as is often in done in linear control ODE problems. Recall that the Laplace transform of a function f is computed by

$$\mathcal{L}(f)(s) = \int_0^\infty f(t)e^{-st}dt.$$

Here we sweep away many technical details and simply assume that our delay equation possesses appropriate regularity to ensure that all Laplace transforms converge for sufficiently large $\mathrm{Re}(s)$. See [8] for technical details concerning convergence of the Laplace transform.

Several standard and easy to prove facts about the Laplace transform make it well suited for usage in linear delay equations. Recall that

- $\mathcal{L}\left(f^{(n)}\right)(s) = s^n\mathcal{L}(f)(s) - \sum_{k=1}^n s^{n-k}f^{(k-1)}(0^-)$, where $f^{(n)}$ denotes the n^{th} derivative of f, and $f(0^-) := \lim_{t\uparrow 0} f(t)$ denotes the left handed limit of f at 0; and
- $\mathcal{L}(f(t-T))(s) = e^{-sT}\left(\mathcal{L}(f)(s) + \int_{-T}^0 f(t)e^{-st}dt\right)$.

Both of these facts can be shown using basic techniques — the former using integration by parts and an easy induction, the latter by a simple change of coordinates.

Additionally, recall that the Laplace transform is one to one up to functions which differ on a set of measure 0, and hence is invertible and whose inverse can be computed by

$$\mathcal{L}^{-1}(F)(t) = \frac{1}{2\pi i}\lim_{h\to\infty}\int_{\alpha-ih}^{\alpha+ih} F(s)e^{st}ds$$

where α is a real number such that the contour integral converges. Again, we point the reader to any number of standard works, such as [8] for the many technical details

which we do not address here.

With these techniques in hand, we can compute the Laplace transform of Equation (2.11). Note that the left hand limits $x(0^-)$ are evaluated using the initial history function, $x(0^-) = \phi(0)$ (assuming ϕ is continuous, which we do). Let $X(s) = \mathcal{L}(x)(s)$. Then, computing the Laplace transform of Equation (2.11), we find that

$$\begin{aligned} &s^2X(s) - s\phi(0) - \phi'(0) + a_1sX(s) - a_1\phi(0) + a_0X(s) \\ &\quad + e^{-sT}\left(b_1sX(s) - b_1\phi(0) + b_0X(s) + \int_{-T}^{0}\left(b_1\phi'(t) + b_0\phi(t)\right)e^{-st}dt\right) = 0 \end{aligned}$$

Next, solving for $X(s)$, we find

$$X(s) = \frac{s\phi(0) + \phi'(0) + a_1\phi(0) + b_1\phi(0)e^{-sT} - \int_{-T}^{0}\left(b_1\phi'(t) + b_0\phi(t)\right)e^{-st}dt}{s^2 + a_1s + a_0b_1se^{-sT} + b_0e^{-sT}}. \quad (2.12)$$

Note that the denominator is precisely the characteristic equation for Equation (2.11), obtained by plugging in the ansatz $x(t) = e^{st}$. This will generally be the case. We can see the influence of the initial condition ϕ in the terms in the numerator.

Next, we need to invert the above expression in order to find our solution $x(t)$. For sufficiently large c, we have that

$$\mathcal{L}^{-1}(X)(t) = \frac{1}{2\pi i}\lim_{h\to\infty}\int_{\alpha-ih}^{\alpha+ih}X(s)e^{st}\,ds$$

where X is given in (2.12). But how to evaluate the integral and perform the inversion? Moreover, what does this have to do with our original goal of expressing the solution $x(t)$ as a series combination of its characteristic functions?

The attentive reader may already start to see where this is going. The Residue Theorem can be used to express the above inverse Laplace transform as a sum of

residues, which in this case we see will be related to the characteristic roots. Recall that the Residue Theorem states that for a function f which is holomorphic off of a set of points $\{z_1, z_2, \ldots, z_m\}$, we have

$$\oint_\gamma f(z)dz = 2\pi i \sum_{j=1}^{m} \text{Res}(f, z_j)$$

where γ is a simple closed curve enclosing the points $\{z_j\}$.

The idea for computing $\mathcal{L}^{-1}(X)(t)$, is to define a sequence of closed contours γ_j which contain the segment from $\alpha - ih$ to $\alpha + ih$ and enclose some number of the singularities of $X(s)$. This sequence γ_j should get larger as j increases in such a way that the segments from $\alpha - ih$ to $\alpha + ih$ converge to the whole vertical line $\alpha - i\infty$ to $\alpha + i\infty$, and to eventually enclose all of the singularities of $X(s)$. Once such a family of contours is found, the residues of $X(s)$ can be computed, and finally $x(t)$ can be expressed as a series involving the characteristic functions $e^{\lambda t}$.

This procedure is generally quite effective. The difficulty comes from the fact that computing the location of singularities of $X(s)$ is in general quite a nontrivial problem in its own right, making the problem of finding an appropriate sequence of contours γ_j and the corresponding residues rather complex.

However, for the sake of illustrating the technique fully, we conclude this discussion with an example which can be worked out completely.

Turn our attention to the delay equation from Example 3,

$$x'' + cx(t - T) = 0 \tag{2.13}$$

$$x(t) = x_0 \qquad \text{for } t < 0.$$

where x_0 is taken to just be a constant. For this example, Equation (2.12) becomes simply

$$X(s) = \frac{x_0(s^2 + ce^{-sT} - c)}{s(s^2 + ce^{-sT})},$$

so that the inverse can be computed as

$$\mathcal{L}^{-1}(X)(t) = \frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} \frac{x_0(s^2 + ce^{-sT} - c)}{s(s^2 + ce^{-sT})} e^{st} ds. \tag{2.14}$$

A fact which we state without proof is that there that all of the poles of the integrand have real part less than some vertical line. We chose α to be to the right of this vertical line so that the inversion formula holds. See [2] for a proof about the distribution of roots of exponential-polynomials.

We evaluate the above integrand using the residue theorem. The integrand has simple poles at the zeros of $s(s^2 + ce^{-sT})$. We break these into two groups: $s = 0$ and $s^2 + ce^{-sT} = 0$. Let $\Lambda = \{\lambda \mid \lambda^2 + ce^{-\lambda T} = 0\}$. Note that Λ is exactly the set of characteristic roots of Equation (2.13).

Recall that for simple poles, a residue can be calculated by $\text{Res}(f, a) = \lim_{z \to a}(z - a)f(z)$. We begin by calculating the residue at 0:

$$\begin{aligned}
\text{Res}(f, 0) &= \lim_{s \to 0} s \cdot \frac{x_0(s^2 + ce^{-sT} - c)}{s(s^2 + ce^{-sT})} e^{st} \\
&= \lim_{s \to 0} \frac{x_0(s^2 + ce^{-sT} - c)}{(s^2 + ce^{-sT})} e^{st} \\
&= 0.
\end{aligned}$$

Next, we calculate the residue at $\lambda \in \Lambda$:

$$\begin{aligned}\text{Res}(f,\lambda) &= \lim_{s\to\lambda}(s-\lambda)\cdot\frac{x_0(s^2+ce^{-sT}-c)}{s(s^2+ce^{-sT})}e^{st}\\ &= \frac{x_0\left(-ce^{\lambda T}+c+\lambda^2e^{\lambda T}\right)}{-c\lambda T+c+3\lambda^2e^{\lambda T}}e^{\lambda t}\\ &= -\frac{cx_0}{\lambda^2(2+\lambda T)}e^{\lambda t}.\end{aligned}$$

where in the second line, we have applied L'Hôpital's rule and then evaluated the limit, and in the last line, we have replaced $e^{\lambda T}$ with $-c/\lambda^2$, using the fact that $\lambda \in \Lambda$.

Finally, to finish the calculation, we define a sequence of contours γ_j which eventually enclose all of the characteristic roots, and enable us to evaluate the the integral in Equation (2.14). Recall that there exists a vertical line in the complex plane such that all the roots are to the left of this line. We choose α in Equation (2.14) to be to the right of all characteristic roots. From here, we define a sequence of contours γ_j to be a sequence of left-facing semicircles of radius j centered on α. The residue theorem implies that

$$\oint_{\gamma_j}\frac{x_0(s^2+ce^{-sT}-c)}{s(s^2+ce^{-sT})}e^{st}ds = 2\pi i\sum_{\lambda\in\text{int}(\gamma_j)}-\frac{cx_0}{\lambda^2(2+\lambda T)}e^{\lambda t}.$$

Now, as $j\to\infty$, the vertical radius of γ_j converges to the vertical line from $\alpha - i\infty$ to $\alpha + i\infty$. It can also be shown that as the radius increases, the contribution of the integrand on the circular to the integral $\oint_{\gamma_j}$ goes to 0. Hence,

$$\begin{aligned}\lim_{j\to\infty}\oint_{\gamma_j}\frac{x_0(s^2+ce^{-sT}-c)}{s(s^2+ce^{-sT})}e^{st}ds &= \int_{\alpha-i\infty}^{\alpha+i\infty}\frac{x_0(s^2+ce^{-sT}-c)}{s(s^2+ce^{-sT})}e^{st}ds\\ &= 2\pi i\sum_{\lambda\in\Lambda}\left(-\frac{cx_0}{\lambda^2(2+\lambda T)}\right)e^{\lambda t}\end{aligned}$$

which finally gives us that

$$\begin{aligned}\mathcal{L}^{-1}(X)(t) &= \frac{1}{2\pi i}\int_{\alpha-i\infty}^{\alpha+i\infty}\frac{x_0(s^2+ce^{-sT}-c)}{s(s^2+ce^{-sT})}e^{st}ds \\ &= \sum_{\lambda\in\Lambda}\left(-\frac{cx_0}{\lambda^2(2+\lambda T)}\right)e^{\lambda t} \\ &= x(t).\end{aligned} \tag{2.15}$$

Figure (2.5) illustrates the contours γ_j in the case of $c = 1$ and $T = 10$.

After all of that effort, we have finally been able to compute an exact expression for the solution $x(t)$ to Equation (2.13) in terms of the characteristic functions (if you are willing to accept that everything is convergent). Since we also have an exact expression for the roots λ in terms of the Lambert W function, Equation (2.15) gives us an exactly computable form for the solution of the DDE (2.13).

Finally, let we check our solution (2.15) to the results of numerical integration. As discussed at the end of Chapter 1, comparing the results of a numerical integration scheme to those cases when an exact solution is available is necessary to establish that the numerical integration scheme is proper, and can be trusted for cases where an analytic solution is not available.

The results are shown in Figures 2.6 and 2.7.

Conclusions

In this section, we have seen some of the essential theory of linear delay equations, and worked out a few examples in detail. Hopefully these examples have helped to convince the reader about the claim we made in the Introduction, that delay equations sit somewhere in between ODEs and PDEs. On the one hand, a number

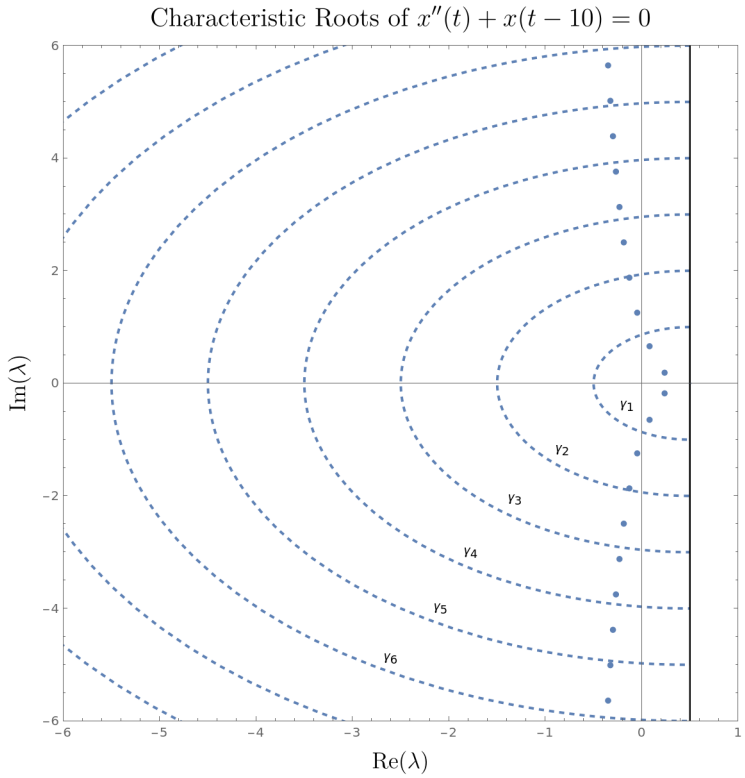

Figure 2.5: An illustration of the contours γ_j for computing the inverse Laplace transform in Equation (2.14) for $x'' + x(t-10)$. Characteristic roots are shown as dots, as computed by Equation (2.3). The contours include the vertical radius at $\alpha = 1/2$ and semicircular arc or radius j. As j goes to infinity, the contours eventually enclose all characteristic roots. The contribution of the integrand on the semicircular arc goes to 0 as $j \to \infty$, so $\oint_{\gamma_j} \to \int_{\alpha-i\infty}^{\alpha+i\infty}$, allowing for the computation of the inverse Laplace transform.

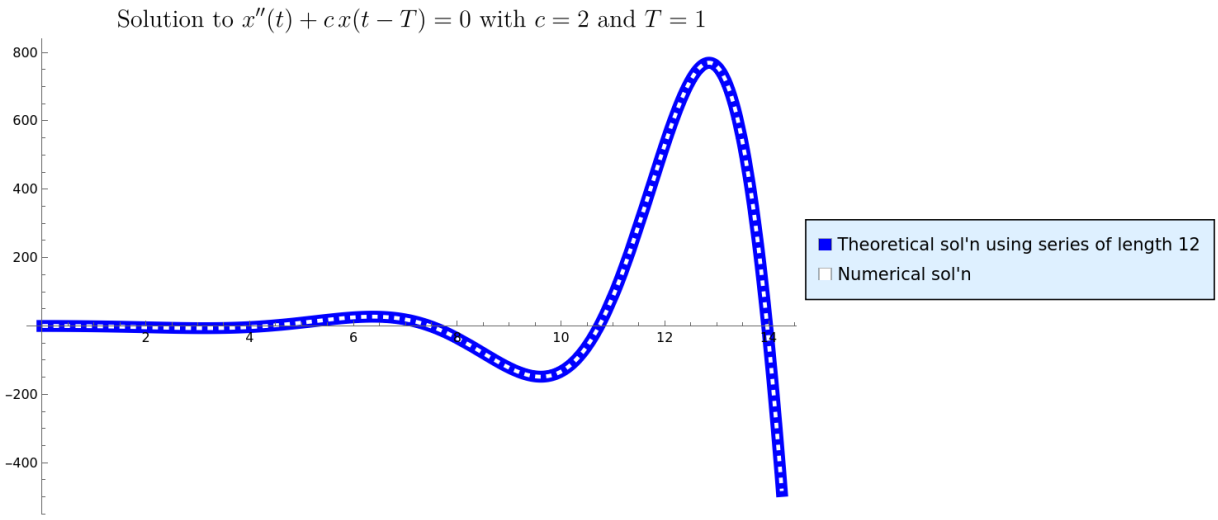

Figure 2.6: A plot of the solution to Equation 2.13 computed numerically and using the series expression calculated in Equation 2.15. The white dashed line is the numerical solution, while the thick solid curve is the result of evaluating the series in 2.15 with 12 terms. We take $c = 2$, $T = 1$, and $x_0 = 1$. Note the extremely close agreement. This gives us reassurance that our numerical integration scheme is "proper."

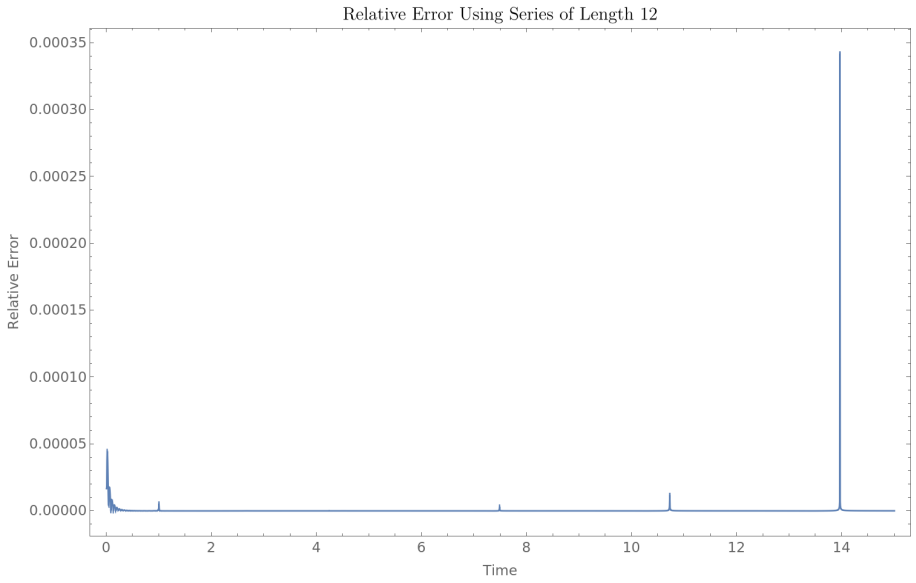

Figure 2.7: The relative error between the series computation and the numerical integration scheme, computed as

$$\left|(x_{\text{series}}(t) - x_{\text{num}}(t))/x_{\text{series}}(t)\right| .$$

The relative error is generally about 1×10^{-5}, with small peaks near the zeros of $x(t)$, owing to the divide-by-zero in the relative error computation. Even in these cases, the relative error is on the order of 1×10^{-4}.

of ideas from linear ODEs transfer almost directly to the study of linear DDEs: the behavior of solutions is determined by the stability of the origin, which can be determined by studying the characteristic equation, and that the solution can be expressed as a linear combination of the characteristic functions. However, hopefully the examples about have illustrated that, while the basic theory may be similar, actually performing any of the necessary computations in the infinite dimensional delay setting can be quite subtle!

With these examples behind us, we now move on to the study of some nonlinear delay equations.

CHAPTER 3

NONLINEAR OSCILLATIONS AND THE METHOD OF HARMONIC BALANCE

Having studied the basic theory of delay equations and focused on linear equations, we now move on to nonlinear delay systems. In this chapter, we study a phenomenon in delay equations which is quite distinct from anything found in in ordinary differential equations, and yet nonetheless intimately related to ODEs. This is the existence of infinitely many self sustaining oscillations — also known as limit cycles — often with amplitudes and frequencies going to infinity. We illustrate the idea with a motivating example, the Delay Duffing Equation, throughout our analysis. We begin by presenting the system and an almost unbelievably simple calculation — known as the method of harmonic balance — which reveals this remarkable behavior. After introducing the technique of harmonic balance, we extend the method to higher orders, revealing bifurcations not previously discussed in the literature.

In Chapter 4, we develop a general and formal framework for understanding this phenomenon of infinitely many limit cycles, which we call the method of characteristic oscillations—a powerful tool for analyzing oscillatory behavior in delay equations. The method involves finding special solutions to a given DDE along which the trajectory is governed by a related ODE. This is in analogy to the method of characteristics from partial differential equations. We give a precise statement of this correspondence, and again illustrate it in detail in the case of the Delay Duffing Equation. Finally, we use the method of characteristic oscillations to prove a result about a whole class of delay systems which exhibit this infinite limit cycle behavior.

These chapters constitute our main contributions to study of nonlinear delay equations — our hope is that the reader finds the proposed method useful in predicting and explaining oscillations in other nonlinear delay systems.

3.1 Introduction

At the turn of the 20th century, at the International Congress of Mathematicians in Paris, the eminent mathematician David Hilbert presented a list of 23 unsolved problems which he believed would define mathematics of the 20th century [17]. While some problems would be solved in a matter of months after his address, others remain almost completely open to this day. Progress on the Hilbert problems has been extraordinarily influential on the direction of mathematical development in the one and a quarter century that has followed. Among the problems which remain open is the 16th, which has been divided into two parts. It is the second part of the 16th problem which is of interest to us.

Problem. (Hilbert's 16th, 1900) Given a vector field $x'(t) = P(x, y)$, $y'(t) = Q(x, y)$ where$P$ and Q are polynomials of degree n, is there an *upper bound* for the possible limit cycles that can appear in the phase space, and if so, what is it as a function of n?

Recall that a limit cycle is a closed trajectory in the phase space of a dynamical system toward which nearby trajectories converge (or from which they diverge) over time, representing a stable (or unstable) periodic solution. Limit cycles capture sustained oscillations in nonlinear systems and are fundamental in modeling rhythmic behaviors in biology, engineering, and physics. See [15, 27, 35] for background on

limit cycles in theory and application.

There has been very little progress towards answering Hilbert's 16th problem. The trivial case of $n = 1$ is known, of course: in this case, the system is linear, and there can be no limit cycles. However, even in the case of $n = 2$, the problem remains wide open. One of the only concrete theorems in this direction is that planar polynomial vector fields have at most finitely many limit cycles. This result can be found in the monumental work [19]. Note that this does not necessarily imply the existence of an upper bound on the number of limit cycles that can exist as a function of n.

However, a little publicized fact is that in the context of *delay differential equations*, this question has been resolved in many cases. Namely, their exist second order polynomial delay equations which contain infinitely many limit cycles. We spend the remainder of this chapter and the next exploring this phenomenon. We will study one example system, the delayed Duffing equation, in detail. In Chapter 4, we then situate this phenomenon in the more general theoretical context of *characteristic oscillations.*

As mentioned in Chapter 1, Compared to ordinary differential equations, DDEs have a much richer structure. Whereas an n^{th} order ODE generally has an n dimensional phase space, the phase space of a DDE is in general an infinite dimensional function space [16]. As such, the topological restrictions which constrain the behavior of low-dimensional ODEs (for example, the Poincaré-Bendixson Theorem, see [35] or [15]) do not apply to DDEs, and for this reason, even low order DDEs can display extremely rich dynamics. It is perhaps this distinction that makes Hilbert's 16th

problem in the delayed case so much more tractable.

3.2 The Delayed Duffing Equation

As a case study for the appearance of infinitely many limit cycles in a second order DDE, we consider a version of the delayed Duffing's Equation,

$$\ddot{x} + cx(t-T) + x^3 = 0 \tag{3.1}$$

for $T > 0$ and $c \in \mathbb{R}_{\neq 0}$. Equations similar to this one have been studied at least since the late 1980s [26], as they have applications in many engineering contexts. It has been known since at least 2003 that for some parameter values, systems similar to this one can contain infinitely many limit cycles [38]. Here, we give an especially simple derivation, which not only illustrates the existence of infinitely many limit cycles, but also finds approximations for their amplitudes and frequencies. This argument, based on the method of Harmonic Balance first appeared in [5] in 2017.

Begin by supposing that there exists a solution of the form $x(t) = A\cos(\omega t)$, for some values of A and ω to be determined. Plugging this ansatz into (3.1) and performing some basic trigonometry, one obtains

$$-\omega^2 A\cos(\omega t)+c+Ac\sin(t\omega)\sin(T\omega)+\cos(t\omega)\cos(T\omega)+\frac{1}{4}\left(3A^3\cos(t\omega)+A^3\cos(3\omega t)\right)=0.$$

Next, we neglect any terms of frequency higher than some value, in this case, 3ω. This yields the equation

$$\left(\frac{3}{4}A^3 - A\omega^2 + Ac\cos(\omega T)\right)\cos(\omega t) + Ac\sin(\omega T)\sin(\omega t) = 0.$$

By linear independence, the coefficients of $\sin(\omega t)$ and $\cos(\omega t)$ must each individually be 0. For the coefficient of $\sin(\omega t)$ to be 0, we conclude that $Ac\sin(\omega T) = 0$, which gives

$$\omega = \frac{n\pi}{T} \tag{3.2}$$

for $n \in \mathbb{Z}$. Now, setting the coefficient for $\cos(\omega t)$ to 0, and solving the resulting equations for A gives

$$A = 2\sqrt{\frac{\left(\frac{n\pi}{T}\right)^2 + c(-1)^{n+1}}{3}} \tag{3.3}$$

We therefore end up with solutions of the form

$$x_n(t) = 2\sqrt{\frac{\left(\frac{n\pi}{T}\right)^2 + c(-1)^{n+1}}{3}} \cos\left(\frac{n\pi}{T}t\right). \tag{3.4}$$

What's remarkable about the above calculation is the appearance of an integer n in the expression, and that the approximate solution (3.4) appears to be valid for any value of $n \in \mathbb{Z}$, suggesting the existence of infinitely many limit cycles. While this may seem like an artifact of the approximation being so crude, the authors of [5] went on to validate this result, first numerically, then by Melnikov's integral, and finally by regularizing equation (3.1) by introducing damping. All of these methods supported the surprising conclusion that (3.1) does indeed contain infinitely many limit cycles. A formal proof of this fact was given in [12, 29] in 2020. We will explore and generalize the ideas behind the formal proof in the Chapter 4

Numerical evidence suggests that the limit cycles alternate stability. When $c > 0$, the limit cycles are stable for odd n and unstable for even n, at least for small values of T. The stability is reversed when $c < 0$. The situation is somewhat more complicated in the case when $c < 0$, due to the existence of additional equilibrium solutions at

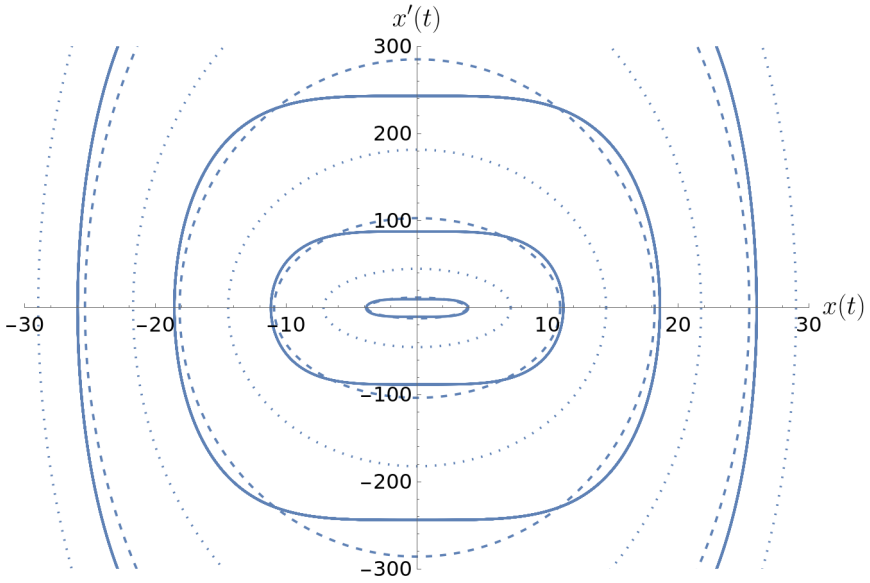

Figure 3.1: A comparison of numerical integration of Equation (3.1) vs the harmonic balance predictions of the limit cycle given by (3.4). Dashed line corresponds to the predicted limit cycles with \$n\$ being odd (stable), while dotted corresponds to predicted limit cycles with n being even(unstable). Solid line corresponds to the results of numerical integration for various initial conditions. (Constant initial conditions were used, of size $2 + 5k$ for $k \in \{0, 1, \ldots, 7\}$). The initial transient approach to the limit cycle is not pictured, only the long time behavior. Note that numerical integration only converges to the odd n limit cycles — this suggests that the limit cycles with even n are unstable. Here, $c = 1$ and $T = 1$. The first 8 limit cycles are pictured.

$x = \pm\sqrt{-c}$. Figure 3.1 illustrates the approximation (3.4) compared to numerical integration.

In this section, we showed how a simple harmonic balance calculation can reveal the existence of infinitely many limit cycles in the delayed Duffing equation. In the next subsection, we see how these oscillations can be explained by a broader framework, which we call "characteristic oscillations."

3.3 Higher Order Harmonic Balance

From the success of the Harmonic Balance approach, which essentially approximates a solution $x(t)$ by a single term of its Fourier series expansion, it is reasonable to attempt to find an approximation $x(t) = A_0 + A_1\cos(\omega t) + A_2\cos(2\omega t) + \cdots + B_1\sin(\omega t) + B_2\sin(2\omega t) + \cdots$. We begin by only considering terms up to frequency 2ω, assuming that later terms will contribute less to the overall value of $x(t)$. Also note that because equation (3.1) is autonomous, we can without loss of generality assume that $B_1 = 0$ by appropriate choice of initial conditions. We therefore seek an approximate solution of the form $A_0 + A_1\cos(\omega t) + A_2\cos(2\omega t) + B_2\sin(2\omega t)$. Plugging this ansatz into (3.1), simplifying the trig, dropping terms of frequency 3ω and higher, and finally setting the coefficients of each trig term to zero, we obtain the following equations:

$$0 = \frac{1}{4}\left(A_0\left(6A_1^2 + 6A_2^2 + 6B_2^2 + 4c\right) + 4A_0^3 + 3A_1^2A_2\right) \tag{3.5}$$

$$0 = \frac{1}{4}A_1\left(12A_0^2 + 3A_1^2 + 6A_2^2 + 12A_0A_2 + 6B_2^2 + 4c\cos(T\omega) - 4\omega^2\right) \tag{3.6}$$

$$0 = A_1\left(3A_0B_2 + c\sin(T\omega)\right) \tag{3.7}$$

$$0 = \frac{1}{2}\left(\frac{1}{2}B_2\left(12A_0^2 + 6A_1^2 + 3A_2^2 + 3B_2^2 - 16\omega^2\right) + 2A_2c\sin(2T\omega) + 2B_2c\cos(2T\omega)\right) \tag{3.8}$$

$$0 = \frac{1}{2}\left(\frac{3}{2}A_2B_2^2 + 2A_2c\cos(2T\omega) - 8A_2\omega^2 + \frac{3A_2^3}{2} + 6A_0^2A_2 + 3A_1^2A_2 + 3A_0A_1^2 - 2B_2c\sin(2T\omega)\right) \tag{3.9}$$

This is a system of five cubic equations with trigonometric coefficients. Solving such a system in general is extremely difficult, and extracting interpretable dynamic

information from such a solution is just as difficult. However there are several special cases in which this system can be solved, each of which reveals information about bifurcations contained in equation (3.1) as the parameters c and T vary. We break our study of these complicated equations into several special cases, each of which sheds light on some new phenomena.

3.3.1 Case 1: $x(t) = A_0 + A_1\cos(\omega t)$

To begin, consider the special case when $A_0 \neq 0$ and $A_1 \neq -0$, but $A_2 = B_2 = 0$. In other words, we seek a solution of the form $x(t) = A_0 + A_1\cos(\omega t)$. This corresponds to an oscillatory solution not centered on the origin. Equations (3.5)-(3.7) can then be simplified to

$$0 = 6A_1^2 + 4c + 4A_0^2 \tag{3.10}$$

$$0 = 3A_0^2 + \frac{3}{4}A_1^2 + c\cos(T\omega) - \omega^2 \tag{3.11}$$

$$0 = c\sin(T\omega) \tag{3.12}$$

while equations (3.8) and (3.9) become identically 0. From equation (3.12), we see that we must have $T\omega = n\pi$ for $n \in \mathbb{Z}$. One can then solve equations (3.10) and (3.11) for A_0 and A_1, which leads one to the equations

$$A_0 = \pm\frac{\sqrt{2\left(\frac{n\pi}{T}\right)^2 - c(2(-1)^n - 1)}}{\sqrt{5}} \tag{3.13}$$

$$A_1 = \pm\frac{2\sqrt{-\left(\frac{n\pi}{T}\right)^2 + c((-1)^n - 3)}}{\sqrt{15}} \tag{3.14}$$

Inspection of the radical in equation (3.14) reveals that A_1 can only be real if the radicand is positive, which can only occur if $c < 0$. There are good reasons for

this: the form of our Ansatz in this case assumes that x is a limit cycle centered on some point other than the origin. In the case that $c < 0$, equation (3.1) supports two additional fixed points, $x^* = \pm\sqrt{-c}$. Therefore, the limit cycles predicted in equations (3.13) and (3.14) can be thought of as being centered on these non-zero fixed points.

To make matters more concrete, consider the case $n = 1$. Our expressions then become

$$A_0 = \pm\frac{\sqrt{3cT^2 + 2\pi^2}}{\sqrt{5}T}$$
$$A_1 = \pm\frac{2\sqrt{-4cT^2 - \pi^2}}{\sqrt{15}T}$$

These results predict the existence of a symmetric pair of limit cycles centered on the non-zero fixed points $\pm\sqrt{-c}$, given by

$$\pm\frac{\sqrt{3cT^2 + 2\pi^2}}{\sqrt{5}T} + \frac{2\sqrt{-4cT^2 - \pi^2}}{\sqrt{15}T}\cos\left(\frac{\pi}{T}t\right).$$

In order for these values to be real, the radicands must be positive. This condition then gives

$$\frac{\pi}{2}\sqrt{-\frac{1}{c}} \leq T \leq \pi\sqrt{-\frac{2}{3c}}.$$

The lower value of T exactly agree with the predicted change of stability of the fixed points $\pm\sqrt{-c}$ calculated in the Subsection 3.4.1, suggesting that these limit cycles are born in a Hopf bifurcation. Since the fixed point stability calculation is somewhat of an aside from the higher order harmonic balance, it is presented at the end of this chapter.

Figure 3.2 compares this theoretical prediction against numerical integration, in the case $c = -1$ and $T = 2$.

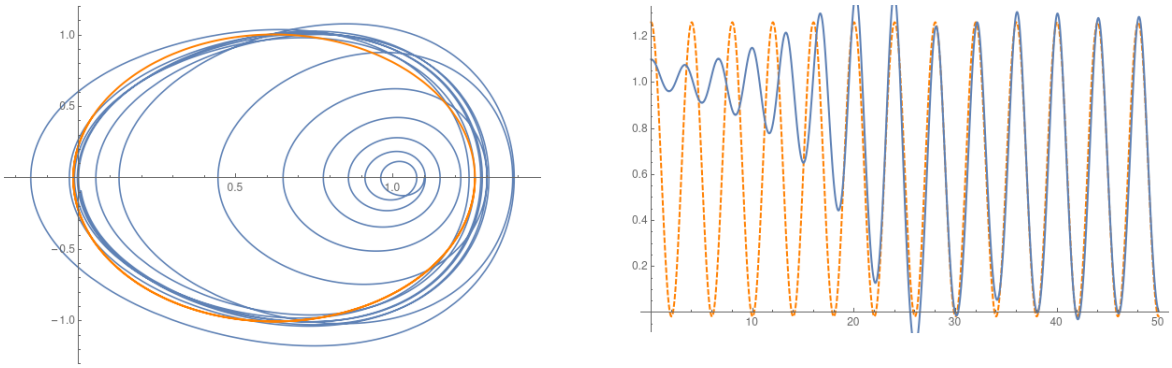

Figure 3.2: In this figure, $c = -1$ and $T = 2$. Orange represents the predicted location of the limit cycle, and blue represents the results of numerical integration, beginning with a more or less random initial condition. The left is a plot in phase space, the right is the time history of the trajectory.

Furthermore, it can be shown that the two symmetric centered on the non-zero equilibrium points limit cycles collide with one of the limit cycles found in section 3.2 in a pitchfork bifurcation of limit cycles, which, if we do say so ourselves, is a pretty wild bifurcation.

3.3.2 Case 2: $x(t) = A_1 \cos(\omega t) + B_2 \sin(2\omega t)$: Period Doubling

Let us consider solutions which are centered on the origin — that is, we take $A_0 = 0$. It turns out that these considerations will reveal a different sort of bifurcation that the limit cycles undergo, in which there spontaneously emerges a second frequency of oscillation. It turns out that this corresponds with a *period doubling bifurcation,* in which the time to complete a full period doubles. With the assumption $A_0 = 0$, the system of equations (3.5)-(3.9) simplifies substantially. Most notably, equation

(3.5) becomes simply

$$\frac{3}{4}A_1^2 A_2 = 0.$$

There are only two possible ways this equation can be true, either $A_2 = 0$ or $A_1 = 0$. It turns out that the case $A_1 = 0$ leads only to the trivial solution where all coefficients are 0. We therefore are forced to take $A_2 = 0$. The remaining system then becomes

$$0 = \frac{1}{4}A_1\left(3A_1^2 + 6B_2^2 + 4c\cos(T\omega) - 4\omega^2\right) \tag{3.15}$$

$$0 = A_1 c\sin(T\omega) \tag{3.16}$$

$$0 = \frac{1}{4}B_2\left(6A_1^2 + 3B_2^2 - 16\omega^2\right) + B_2 c\cos(2T\omega) \tag{3.17}$$

$$0 = B_2 c\sin(2T\omega) \tag{3.18}$$

Notice that both equations 3.16 and 3.18 will be satisfied if $T\omega = n\pi$, where $n \in \mathbb{Z}$. We therefore again write $\omega = \frac{n\pi}{T}$. This simplifies equations (3.15) and (3.17) to a system of quadratic equations in A_1 and B_2 :

$$0 = \frac{1}{4}A_1\left(3A_1^2 + 6B_2^2 + 4c(-1)^n - 4\left(\frac{n\pi}{T}\right)^2\right) \tag{3.19}$$

$$0 = \frac{1}{4}B_2\left(6A_1^2 + 3B_2^2 - 16\left(\frac{n\pi}{T}\right)^2\right) + B_2 c \tag{3.20}$$

We can solve the resulting system to find the following approximate solutions of the form $x(t) = A_1\cos\left(\frac{n\pi}{T}t\right) + B_2\sin\left(\frac{2n\pi}{T}t\right)$ where $n \in \mathbb{Z}$ is arbitrary. In so doing, we obtain several types of solutions; however, most of the solutions turn out to be redundant with those found in the first part of Section 3.2, Equation (3.4). It turns out, however, that in addition to these redundant solutions, we also end up with a new solution form:

$$x(t) = \pm \frac{2}{3}\sqrt{7\left(\frac{\pi n}{T}\right)^2 + c\left((-1)^n - 2\right)}\cos\left(\frac{\pi n t}{T}\right)$$
$$\pm \frac{2}{3}\sqrt{-2\left(\frac{\pi n}{T}\right)^2 + c\left(2(-1)^{n+1} + 1\right)}\sin\left(\frac{2\pi n t}{T}\right) \tag{3.21}$$

These solutions are genuinely new, and must be studied separately.

First, we need to investigate when both radicals give rise to real numbers, i.e. where

$$7\left(\frac{\pi n}{T}\right)^2 + c\left((-1)^n - 2\right) > 0 \quad \text{and} \quad -2\left(\frac{\pi n}{T}\right)^2 + c\left(2(-1)^{n+1} + 1\right) > 0$$

The inequalities depend on the parity of n. Separating into the cases when n is even and odd, we obtain the conditions following conditions on T :

$$\begin{cases} \frac{\sqrt{2}\pi n}{\sqrt{3c}} \leq T \leq \frac{\sqrt{7}\pi n}{\sqrt{3c}} & \text{if } n \text{ is odd} \\ T \geq \frac{\sqrt{2}n\pi}{\sqrt{-c}} & \text{if } n \text{ is even and } c < 0 \end{cases}$$

What are we to make of these new solutions? To begin with, let n be odd and consider what happens right at the value $T = \frac{\sqrt{2}\pi n}{\sqrt{3c}}$. In this case, (3.21) reduces to

$$x(t) = \frac{2}{3}\sqrt{\frac{15c}{2}}\cos\left(\sqrt{\frac{3c}{2}}t\right).$$

Comparing this to the solutions found in the first part of Section 3.2 when we plug in $T = \frac{\sqrt{2}\pi n}{\sqrt{3c}}$ to Equation (3.4), we also find

$$x(t) = \frac{2}{3}\sqrt{\frac{15c}{2}}\cos\left(\sqrt{\frac{3c}{2}}t\right).$$

Therefore, the solution with two component frequencies is "born out of" the solution with a single frequency. We illustrate this bifurcation in Figure 3.3. We take $c = 3$

and $n = 1$ for these figures. This means that equation (3.21) predicts a bifurcation to occur at $T = \frac{\sqrt{2}}{3}\pi \approx 1.48$.

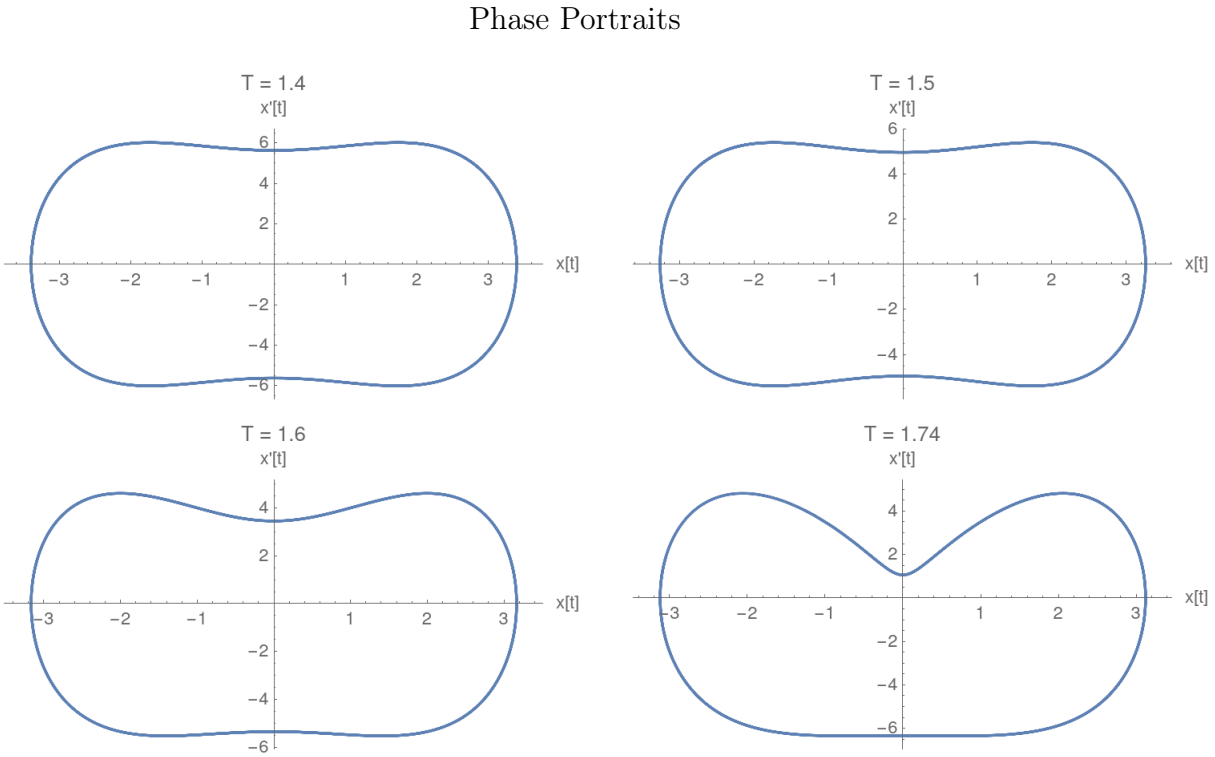

Figure 3.3: Numerical simulations of equation (3.1) with $c = 3$ and varying values of T. Initial condition is $x(t) = .9$ for $t \leq 0$. The solution is plotted from time $t = 500$ to $t = 510$ to allow initial transience to settle. Notice that as T increases past the threshold $T_c = \frac{\sqrt{2}}{3}\pi \approx 1.48$, a notable "dimple" appears. This corresponds to the appearance of a new frequency component in the solution's Fourier series expansion.

Figure 3.4 shows a comparison of numerical simulation against the approximation predicted by equation (3.21).

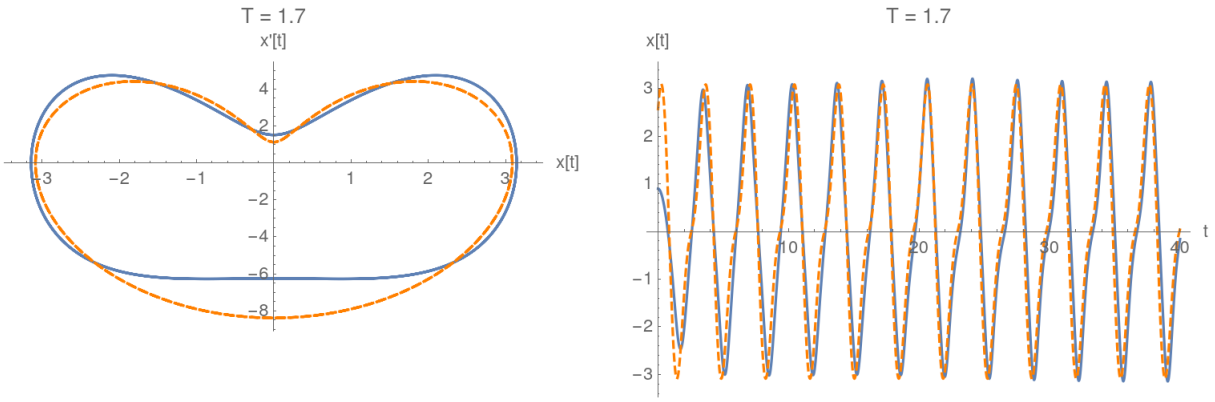

Figure 3.4: A comparison of numerical integration vs theoretically predicted result. Here $T = 1.7$ and $c = 3$. The orange dashed curves are the theoretically predicted curves from equation (3.21), while the solid blue curves are the results of numerical integration.

For another perspective on the period doubling bifurcation, we compute the Fourier transform of the numerically calculated solution, for various values of the delay T. The Fourier transform identifies the oscillatory frequencies present in the signal. As T increases, we distinctly see the appearance of a second frequency, corresponding to the period doubling bifurcation.

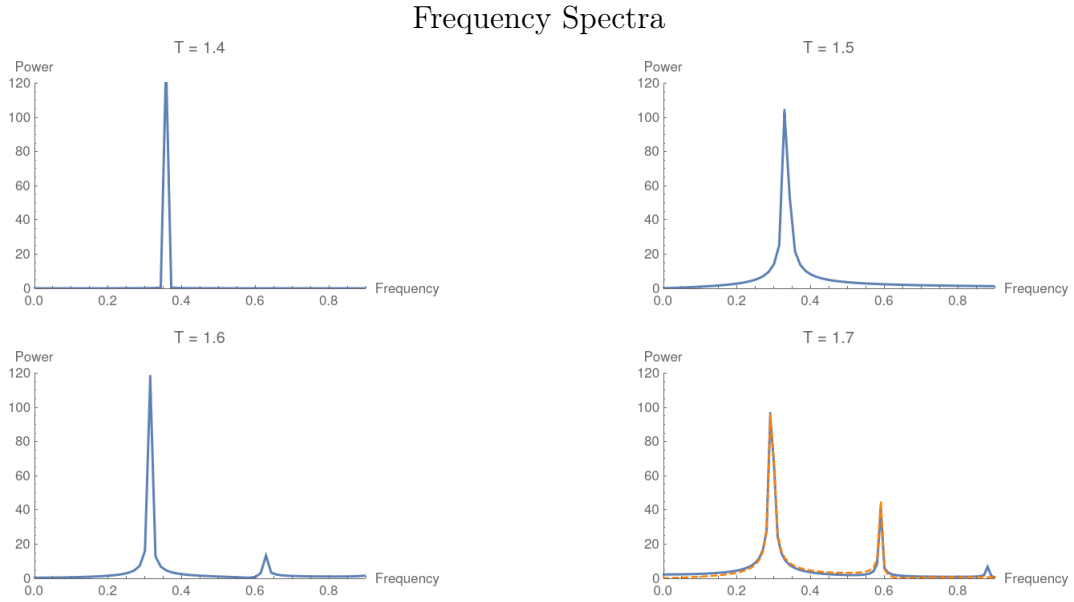

Figure 3.5: Here we show the frequency spectra of the numerically computed solution for various values of T using the Fast Fourier Transform. Notice that at $T = 1.4$ the frequency is completely concentrated a single peak, but at $T = 1.5$, the frequency spectrum has become more spread out. At $T = 1.6$ a noticeable second peak has appeared. In the bottom right panel, we compare the frequency spectrum of the numerical solution in blue with the frequency spectrum predicted in equation (3.21) (dotted orange). Notice the excellent agreement between the prediction and the numerical solution.

3.4 Further Questions

Despite the success of the harmonic balance approach in studying equation (3.1), there still remain many questions. Numerical simulation suggests that there are further bifurcations that are not currently predicted by the harmonic balance method. Numerical results suggest that the limit cycle that we studied in Section (3.3.2) appears to undergo a phase shift, followed by further period doublings, and eventually becomes chaotic, before finally becoming unstable. Figures 3.6, 3.7 and (3.8) depict

some of the bifurcation process that occurs.

We see that between $T \approx 3.55$ and $T \approx 3.6$, the orbit undergoes repeated period doublings, in a *period doubling cascade*, leading to the onset of chaos at $T \approx 3.61$. As T continues to be increased, one sees parameter regions of chaos, interspersed with windows of stable periodic orbits. This situation is reminiscent of the well known bifurcation diagram for the logistic map. For more information on period doubling cascades, see [35] Chapter 10, as well as Feigenbaum's seminal work on the subject [11].

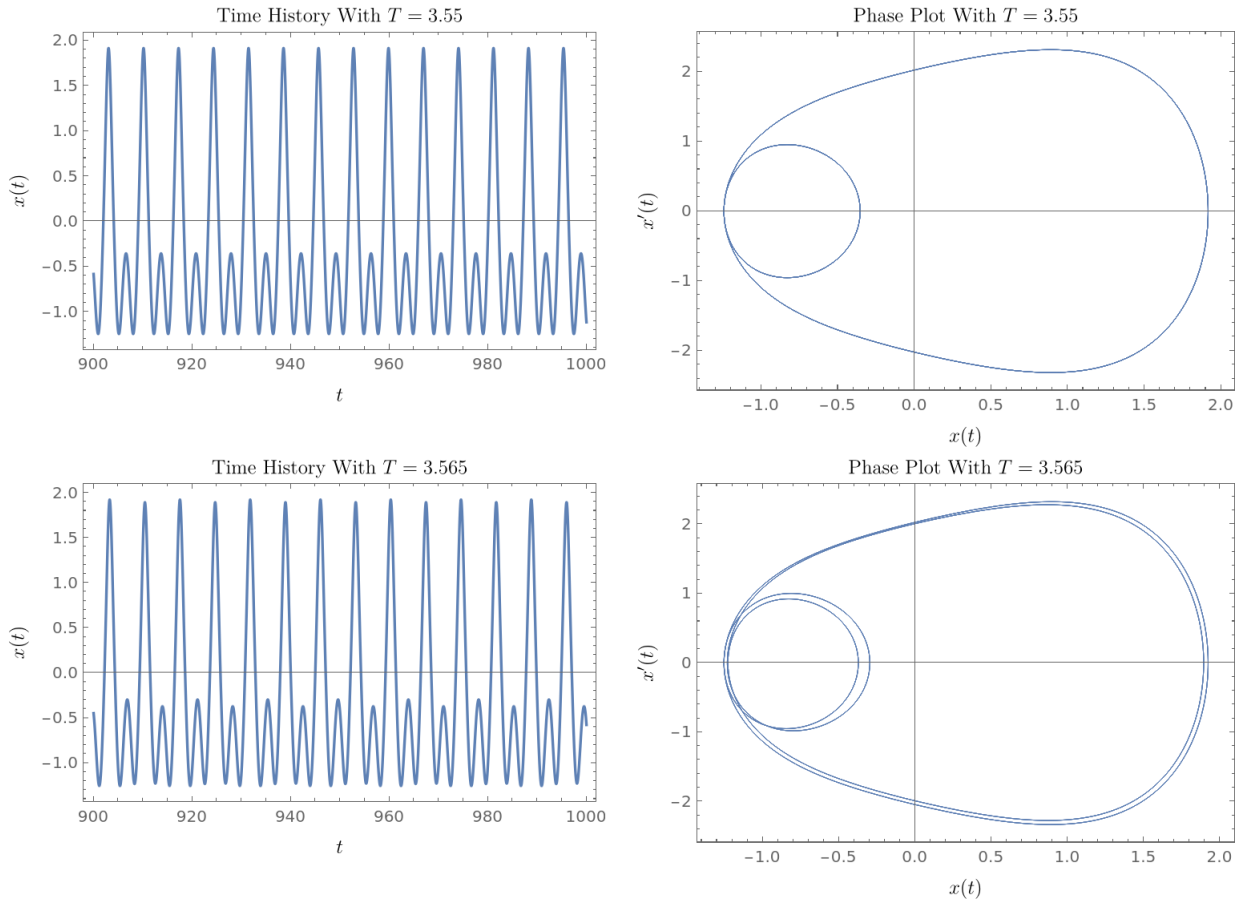

Figure 3.6: Phase portrait and time series of the innermost periodic motion. At $T \approx 3.565$, we see that the limit cycle undergoes a *period doubling bifurcation.* For these figures, we integrate $x'' + x(t-T) + x^3 = 0$ with initial condition $x(t) = .1$ for $t \leq 0$ on the interval $t \in [0, 1000]$. We plot just $t \in [900, 1000]$ to avoid capturing initial transience.

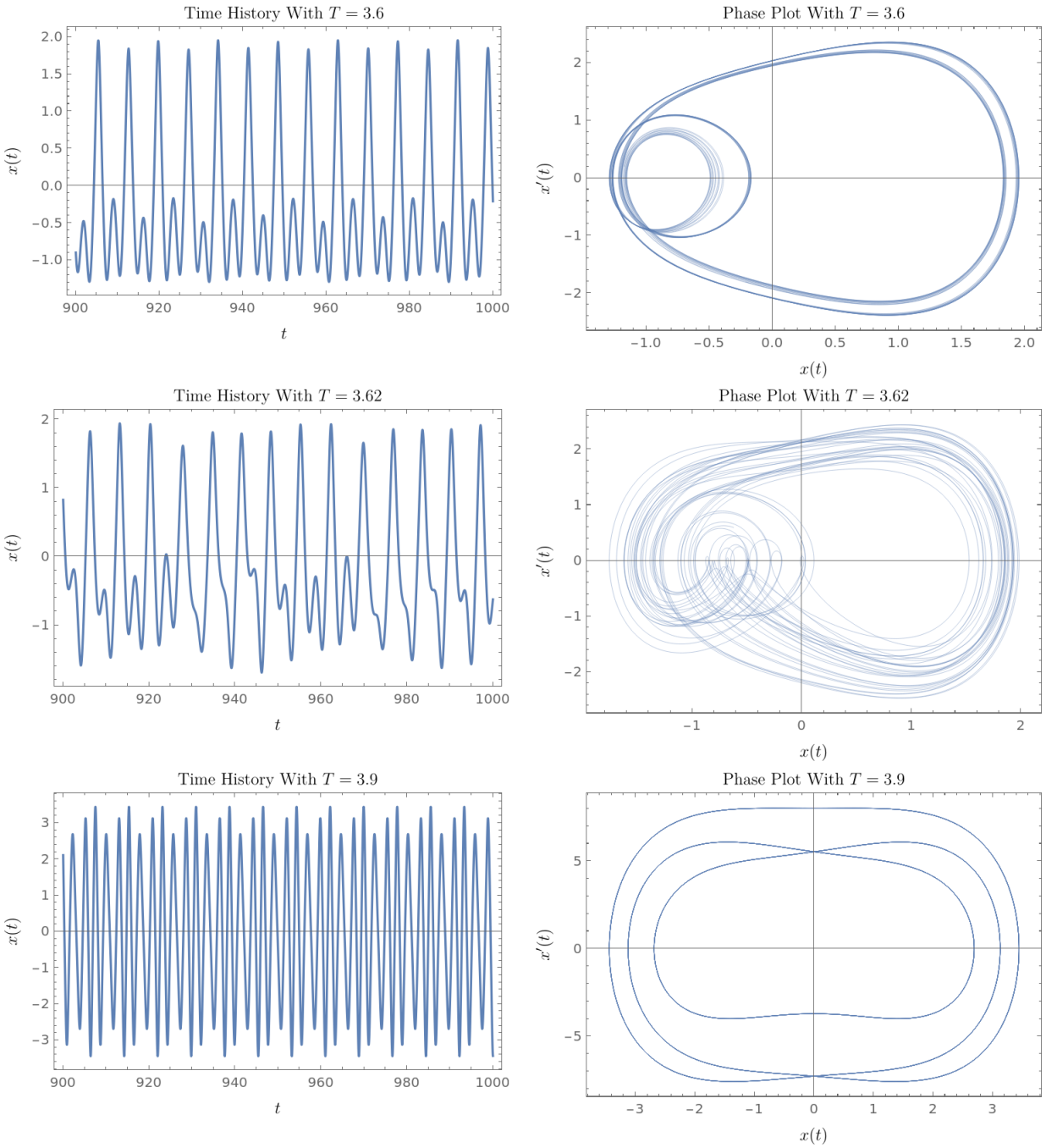

Figure 3.7: Further phase portraits of the innermost limit cycle. We see that when $T = 3.6$, the limit cycle has undergone many more period doublings. Motion is still periodic, but with very large period. By $T = 3.62$, the motion has become aperiodic, and appears chaotic. Chaotic behavior persists between $T \approx 3.62$ and $T \approx 3.9$.

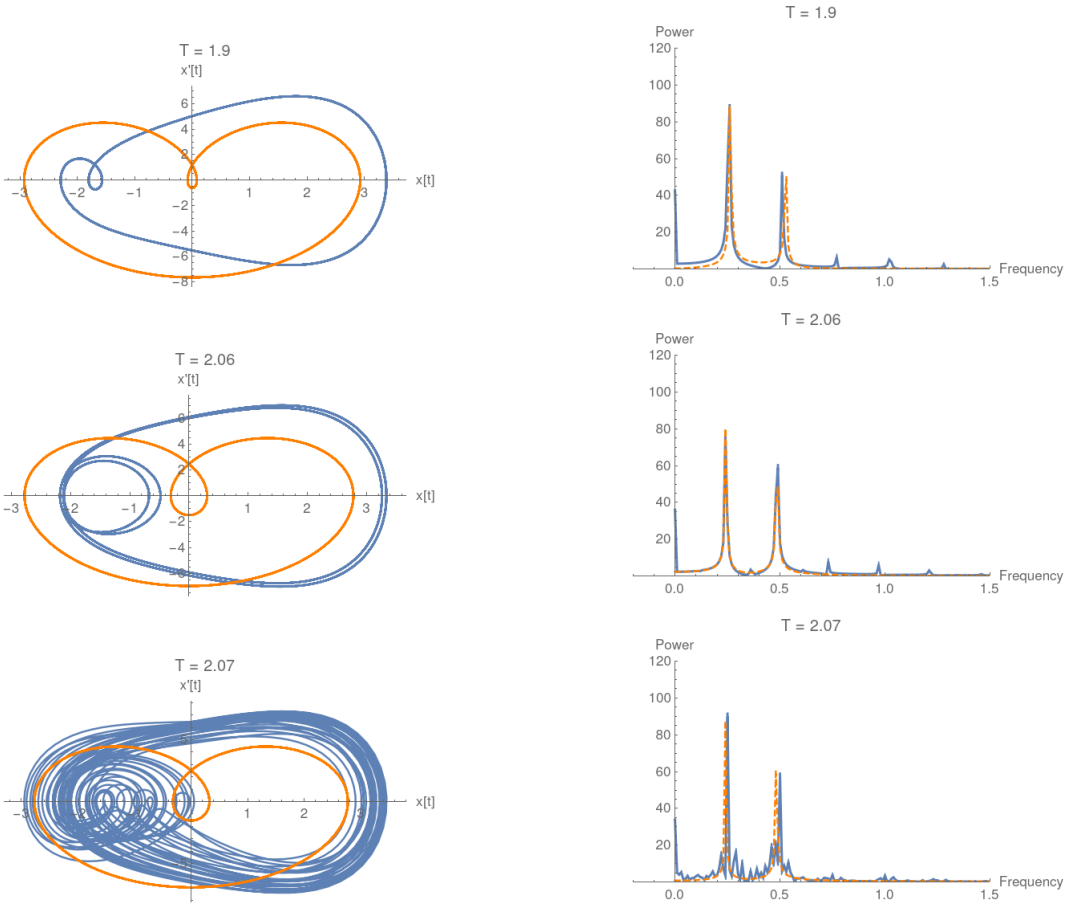

Figure 3.8: As T continues to increase, there are a series of bifurcations that are currently not explained by our approximations. In particular, there is a regime from $T \approx 2.07$ to $T \approx 2.21$ in which there appears to be a chaotic attractor, characterized by an orbit which fills a bounded area of phase space and has a broad band frequency spectrum — see the bottom part of the figure. Interestingly, even in this chaotic regime, the frequency still is dominated by the two peaks predicted by equation (3.21). All figures use the parameter $c = 3$ and show the results of numerical integration from time $t = 420$ to $t = 520$.

Figures 3.6, 3.7, and 3.8 show some of the complexity present in the Delay Duffing equation. A complete treatment of the entire bifurcation process is beyond the scope of this paper, and indeed, is not yet fully understood. We include this discussion

primarily to caution the reader that there is much going on than initially meets the eye in this simple looking DDE.

In this chapter we have seen that, even an equation which is seemingly as simple as a nonlinear delay equation can be, $x'' + x(t-T) + x^3 = 0$, contains a bewildering complexity within it. In the next chapter, we make a few steps towards taming this complexity, but putting this zoo of oscillations into the framework of *characteristic oscillations.* In the final chapter, Chapter (5), we discuss the chaotic aspects of Equation (3.1) in much greater detail.

3.4.1 Stability of the Fixed Points $\pm\sqrt{-c}$

As an addendum to this chapter, we investigate the linear stability of the fixed points at $\sqrt{-c}$ in the case when $c < 0$. This analysis compliments the calculations in this chapter by confirming that the symmetric pair of limit cycles found in Section (3.3.1)3.3.1 are indeed born in a Hopf bifurcation. We begin by letting $y(t) = x(t) - \sqrt{-c}$. (Note that by symmetry, the two fixed points $\pm\sqrt{-c}$ have the same stability).

Equation (3.1) becomes

$$\begin{aligned} \frac{d^2y}{dt^2} + c(y(t-T) + \sqrt{-c}) + \left(y + \sqrt{-c}\right)^3 &= 0 \\ \frac{d^2y}{dt^2} + cy(t-T) + y^3 + 3y^2\sqrt{-c} - 3cy &= 0 \end{aligned} \tag{3.22}$$

Now, linearizing (3.22) about the origin $y = 0$ gives

$$\frac{d^2y}{dt^2} + cy(t-T) - 3cy = 0 \tag{3.23}$$

Equation (3.23) has characteristic equation

$$\lambda^2 + ce^{-\lambda T} - 3c = 0 \tag{3.24}$$

and it is known that the origin is asymptotically stable only when all of the solutions to (3.24) have negative real part. Therefore, a change in stability can occur in (3.22) when (3.24) has a solution with real part equal to 0. Let $\lambda = \alpha + i\beta$. Breaking (3.24) into its real and imaginary parts, we find that

$$\begin{aligned} \alpha^2 - \beta^2 + ce^{-\alpha T}\cos(\beta T) - 3c &= 0 \\ 2\alpha\beta - ce^{-\alpha T}\sin(\beta T) &= 0 \end{aligned}$$

and now setting $\alpha = 0$ to look for changes of stability, we find

$$\begin{aligned} -\beta^2 + c\cos(\beta T) - 3c &= 0 \\ -c\sin(\beta T) &= 0 \end{aligned}$$

Assuming $c \neq 0$, this system has solutions of the form

$$c = \left(\frac{n\pi}{T}\right)^2 \frac{1}{((-1)^n - 3)} \tag{3.25}$$

or equivalently

$$T = \frac{n\pi}{\sqrt{c\left((-1)^n - 3\right)}}$$

where $n \in \mathbb{Z}_{\geq 0}$.

Moreover, we can compute $\frac{d\alpha}{dT}$ at these values of T, using the techniques described in Section 2.1, to find that at these values of T, we have

$$\frac{d\alpha}{dT} = \frac{2c(-1)^n\left((-1)^n - 3\right)^2}{\pi^2 n^2 + 4\left((-1)^n - 3\right)^2}.$$

Note that the denominator is always positive, and that the numerator alternates sign with n. This confirms that the eigenvalues alternate crossing the real axis from left to right and from right to left. Therefore, the fixed points repeatedly alternate stability as T increases, going from stable to unstable and then back again. This fact reveals that the fixed points undergo an infinite sequence of Hopf bifurcations, in which a limit cycle is born, increases in amplitude, and then decreases in amplitude until it eventually disappears, only to be reborn again. We note that Equation (3.25) precisely agrees with the condition from the harmonic balance method in Equation (3.14) for the appearance of an oscillation centered on the non-zero fixed points.

Hence, we have revealed yet another surprising piece of structure in the delayed Duffing equation — when $c < 0$, it features an infinite sequence of Hopf bifurcations as the delay increases!

Numerical results shown in Figures (3.9), and (3.10) confirm for the first several values of n that stability changes do occur at the values of c given by equation 3.25. For these figures, we investigate the system with $T = 1$ and initial condition $x(t) = \sqrt{-c} + 0.1$ for $t \leq 0$. We plot the $(x(t), x'(t))$ plane. Stability changes are expected at $c = \frac{(n\pi)^2}{((-1)^n - 3)}$, the first several values of which are $c = 0, -2.4674, -19.739$.

Figure 3.9: Left: $c = -2.3$. The system quickly decays to the fixed point at $\sqrt{-c}$. Right: $c = -2.6$. The system is spiraling outwards and accumulating on a limit cycle centered on $\sqrt{-c}$. A stability change has occurred very near the predicted value of $c = -2.467$.

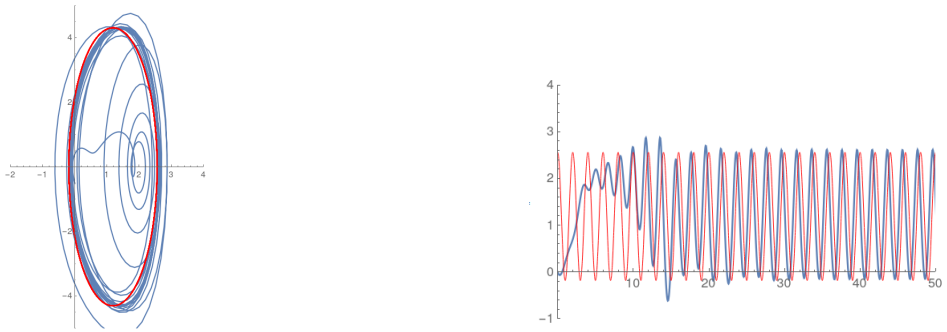

Figure 3.10: Left: $c = -20, T = 1$. The trajectory decays onto the fixed point at $\sqrt{-c}$. Right: $c = -19.5$ and $T = 1$. With the same initial conditions, the trajectory now spirals outward. A stability change has occurred at the fixed point.

CHAPTER 4

THE METHOD OF CHARACTERISTIC OSCILLATIONS

In this section, we put results of the previous sections into a more general context. The core idea is that in some contexts, periodic solutions to DDEs actually come from periodic solutions to related ODEs. We make this idea precise, and then leverage it to find a whole family of DDEs which contain infinitely many limit cycles, thereby offering some context for the remarkable singularity observed in the Delayed Duffing Equation. The idea of periodic ODEs yielding periodic solutions to DDEs dates back to [20]. In this section, we offer another result along the same lines, which is a generalization of techniques used in [12, 29] to find exact periodic solutions to the Delayed Duffing Equation.

We begin with an exact statement of the general idea, and then give an example.

Theorem 5. *Consider the ordinary differential equation*

$$\frac{dx}{dt} = f(x) \tag{4.1}$$

where $x(t) \in \mathbb{R}^m$ for some m. Suppose that $\hat{x}(t)$ is a periodic solution of (4.1) with period P. Now consider the delay differential equation with delay T

$$\frac{dx}{dt} = g(x, x(t-T)) \tag{4.2}$$

where $g(x, x(t-T))$ is just a delayed version of $f(x)$, i.e. $g(x, x) = f(x)$. If the period P of $\hat{x}$ divides T, i.e. if there exists an integer n such that $nP = T$, then $\hat{x}(t)$ is a solution to (4.2).

We can think of g as being the "same function" as f, where some occurrences of $x(t)$ are replaced by $x(t-T)$.

Proof. By assumption, $\hat{x}$ is periodic with period P, which means exactly that $\hat{x}(t - P) = \hat{x}(t)$ for all t. Inductively, this also implies that $\hat{x}(t - P) = \hat{x}(t - 2P) = \cdots = \hat{x}(t - nP)$ for all t. By assumption, $nP = T$, so

$$\hat{x}(t) = \hat{x}(t - T) \tag{4.3}$$

for all t. Hence,

$$\begin{aligned} \frac{d\hat{x}}{dt} &= f(\hat{x}) \\ &= g(\hat{x}, \hat{x}) \\ &= g(\hat{x}, \hat{x}(t - T)) \end{aligned}$$

so that $\hat{x}$ is a solution to the delay equation (4.2) as well.

This concludes the proof. □

We propose the name ***characteristic oscillations*** for the special solutions (4.3) described in Theorem 5. The term is borrowed from the method of characteristics in PDEs, in which one finds special solution curves along which the dynamics are governed by an ODE. The situation is quite analogous here: Theorem 5 finds particular oscillatory solutions along which the dynamics are governed by an ODE rather than a DDE.

Notice that Theorem 5 was stated in the context of an explicit ODE and has implications for explicit DDEs of retarded type. However, the result could easily be generalized to implicit ODEs $F\left(x'(t), x\right) = 0$ and implicit DDEs

$$G\left(x'(t), x'(t - T), x(t), x(t - T)\right)$$

with the same proof, and hence the result applies to equations in which the delay appears in the derivative terms as well. (Such equations are known as *delay equations of neutral type.* See [16], chapter 2.7 for details).

The implication of the above Theorem is that if we can find periodic solutions to an ODE with appropriate period, then we will have found periodic solutions to the corresponding DDE as well. The observation seems extremely simple, but we illustrate in the example of the delayed Duffing Equation how powerful it is. This example was first explored in [12, 29].

4.1 Example: The Delayed Duffing Oscillator

Consider again the Duffing ODE equation

$$\ddot{x} + x + x^3 = 0 \tag{4.4}$$

and its delayed counterpart

$$\ddot{x} + x(t - T) + x^3 = 0. \tag{4.5}$$

What we will find in this example is that the limit cycles discussed in Chapter 3 are in fact characteristic oscillations. If we write (4.4) as a first order system, it reads

$$\dot{x} = y$$
$$\dot{y} = -x - x^3$$

which is in the form of equation (4.1) with $f(x, y) = (y, -x - x^3)$. Similarly, equation (4.5) can be written as the first order system

$$\dot{x} = y$$
$$\dot{y} = -x(t - T) - x^3$$

which is of the form (4.2) with $g(x, y, x(t - T), y(t - T)) = (y, -x(t - T) - x^3)$. Note that $g(x, y, x, y) = f(x, y)$, as required by the assumptions of Theorem 5. Then Theorem 5 asserts that periodic solutions of the ODE Duffing equation (4.4) are exact solutions to the delayed Duffing equation (4.5), provided the period P of the solution satisfies $nP = T$ for some $n \in \mathbb{N}$. We show that such solutions with the appropriate period to equation (4.4) do exist — in fact, there are infinitely many such solutions.

To aid in illustrating the idea, we use the fact that the Duffing ODE is one of the rare examples of a nonlinear ODE for which we can find an exact solution, in the form of Jacobi elliptic functions:

$$x(t) = A \operatorname{cn}(b_1 t + b_2, m). \tag{4.6}$$

With appropriate choice of initial condition, we can take $b_2 = 0$. See [27] Chapter 2 for details on Jacobi elliptic functions, and their use in solving the Duffing ODE. For the purposes of our example, however, familiarity with Jacobi elliptic functions is not crucial. The most important relevant facts are that $A\operatorname{cn}(b_1 t, m)$ is periodic in t, and that the period is determined by the parameters b_1 and m. Moreover, the solutions given in (4.6) display a *period-amplitude relationship* — as the amplitude A increases, the period decreases. In effect, this implies that the parameters A, b_1,

and m are not independent. For those interested, the exact relationship between A, b_1, and m is derived in [27], equation (67). It turns out that

$$m = \frac{A^2}{2 + 2A^2} \text{ and } b_1 = \sqrt{1 + A^2}. \tag{4.7}$$

One feature of the Duffing ODE is that the period of a solution, which is given by

$$P(A) = \frac{4K\left(\frac{A^2}{2+2A^2}\right)}{\sqrt{1 + A^2}},$$

decreases to 0 as the amplitude A increases to infinity. Here, $K(\cdot)$ denotes the complete elliptic integral of the first kind. Again, see [27] for details on these functions. Figure 4.1 illustrates the period of oscillations in the Duffing ODE as a function of the amplitude A.

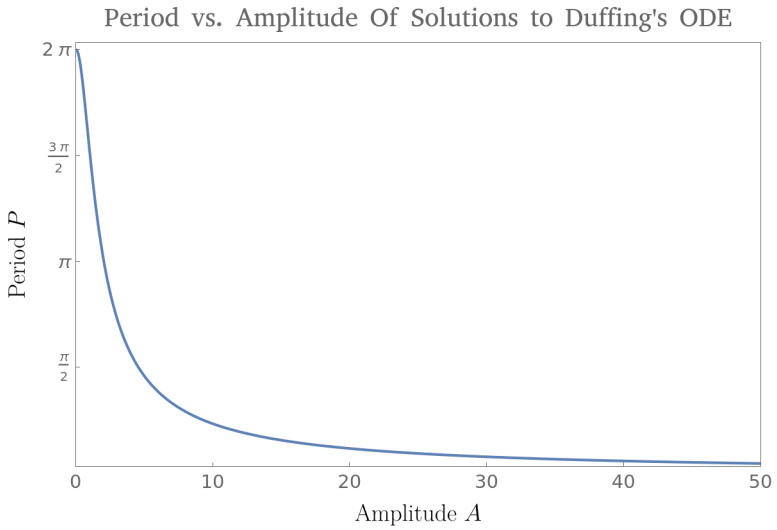

Figure 4.1: The period of solutions to Duffing's ODE $A\,\mathrm{cn}\,(b_1 t, m)$ decrease to 0 as $A \to \infty$.

Now, according to Theorem 5, a periodic solution $x(t)$ of the Duffing ODE is an

exact solution of the Duffing's DDE whenever the period of $x(t)$ satisfies $P = T/n$ for some $n \in \mathbb{N}$. To be concrete, let us set the delay $T = 2$. Therefore, if we can find solutions of Duffing's ODE with periods of $2/n$ for $n \in \mathbb{N}$, these will be exact solutions to Duffing's DDE. Because the period $P(A)$ goes to 0 as A goes to infinity, it is possible to find amplitudes A such that $P(A) = 2/n$ for each $n \geq 1$. Figure 4.2 illustrates that for each n, we can find a value of A so that the solution $A\,\mathrm{cn}\,(b_1 t, m)$ will have period $2/n$. This confirms the surprising result from Chapter 3 that the delayed Duffing equation has infinitely many limit cycles.

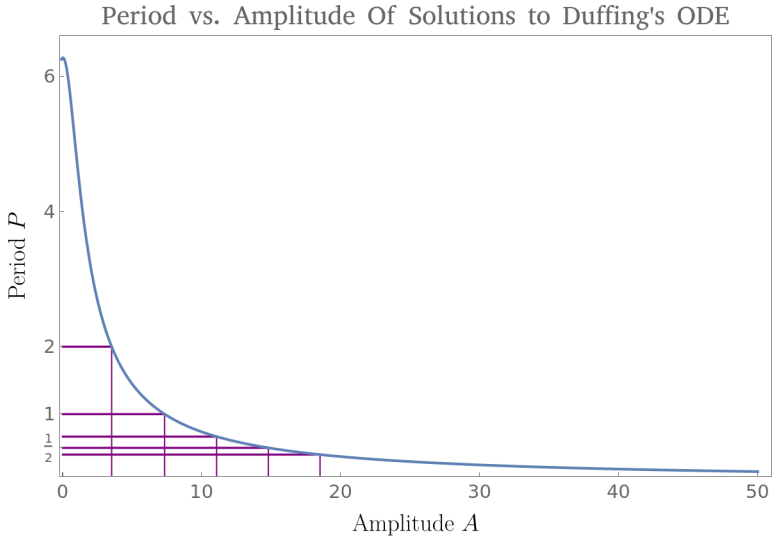

Figure 4.2: For value of the n, we can find an amplitude A such that the solution to Duffing's ODE has period $2/n$. The figure shows $n = 1, 2, 3, 4$, in purple. The corresponding values of A are 3.507, 7.317 and 11.059, and 14.783. Clearly this could be repeated for any value of n. This illustrates that we can find infinitely many solutions to the Duffing ODE that satisfy $x(t-2) = x(t)$, and hence, that there are infinitely many characteristic oscillations in the delayed Duffing equation with delay $T = 2$. A similar analysis applies to all values of the delay.

The solutions with the appropriate period also have amplitudes which can be

calculated using equation (67) of [27]. For this example, the first few have amplitudes given by $3.51, 7.32, 11.06$, and 14.78. We emphasize again that with these values of A and the corresponding values of b_1 and m given by equation (4.7), the solution $A\,\mathrm{cn}\,(b_1 t, m)$ are *exact solutions* to the delay equation $x''(t) + x(t-2) + x^3 = 0$.

How do we interpret these solutions to the delayed Duffing equation which are inherited from the solutions to the non-delayed Duffing equation? One way to think about it is that the period of the inherited solutions are timed in just such a way that the delayed state $x(t-T)$ and the current state $x(t)$ precisely coincide. In these special cases, the (normally dramatic) effects of delay are not present. This can be true for solutions which complete exactly 1 revolution in the amount of time T, as well as solutions that complete exactly 2, 3, or any integer number of revolutions in the time T. Because the oscillations of the Duffing equation get arbitrarily fast as the amplitude increases, it turns out that there are infinitely many such special orbits. The situation is illustrated in Figure 4.3.

It is worth comparing these values of A to the approximate values of the approximate limit cycle amplitudes found in Chapter 3, if we plug in $T = 2$ to Equation (3.3). This comparison is performed in Table 4.1.

We notice from Table 4.1 that the values of A coming from the harmonic balance method with even values of n agree very well with the exact values of A found by studying characteristic oscillations. However, the method of characteristic oscillations seems to be missing all of the limit cycles corresponding to odd values of n. Why is this happening? To answer this question, we must extend the method of characteristic oscillations, which we do in the next section.

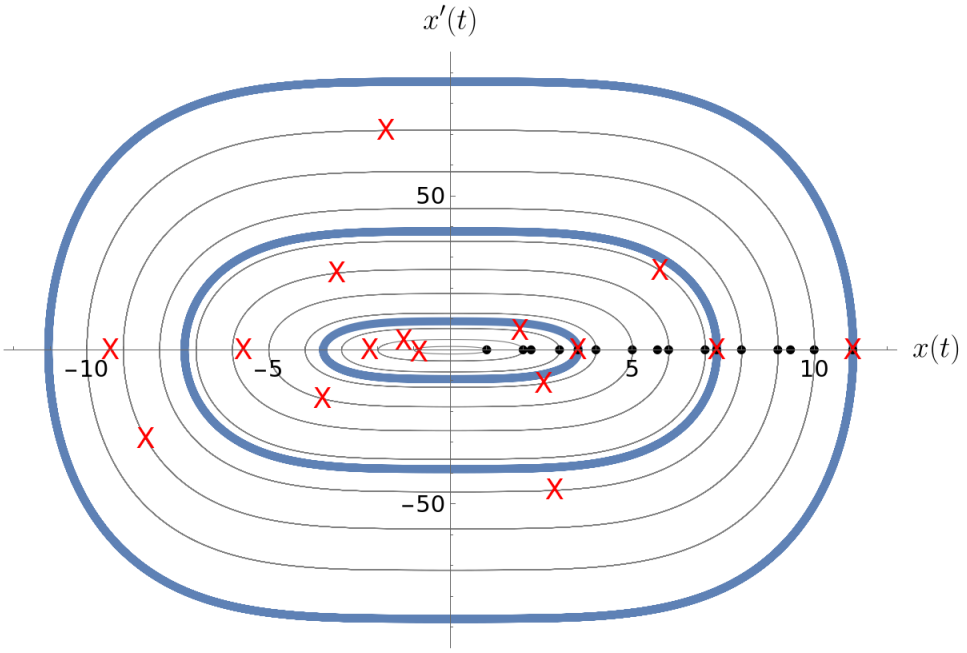

Figure 4.3: A picture of the phase portrait of the Duffing ODE, Equation (4.4). Trajectories of several initial conditions — shown by black dots — are shown. Trajectories start at their initial condition and then travel clockwise along the shown curves. After a predetermined amount of time — in this case, 2 time units — a red X is used to mark the solution's current position. Notice that on the curves shown in bold orange, the red X exactly aligns with the initial condition. These trajectories are exactly the characteristic oscillations for the delayed Duffing equation with $T = 2$ — on these trajectories, the delay has no effect.

4.2 More General Characteristics

In the previous section, we saw in Theorem 5 that in certain cases, oscillatory solutions to a DDE are governed by a corresponding ODE. We applied this idea to the Duffing DDE and were thereby able to confirm the surprising result that the Duffing DDE has infinitely many limit cycles. Having obtained an exact expression for these limit cycles, we checked the amplitudes of the limit cycles obtained by the

n	Characteristic Oscillation Amplitude with $P = 2/n$	Harmonic Balance Amplitude of L.C. $2n$	Percent Error in Harmonic Balance Approx
1	3.51	3.44	1.9%
2	7.32	7.16	2.1%
3	11.06	10.82	2.1%
4	14.78	14.46	2.2%

Table 4.1: A comparison of the limit cycle amplitudes found by the method of characteristic oscillations (exact) to the amplitudes found by harmonic balance (approximate, Equation 3.3) with $T = 2$. Notably, the method of characteristic oscillations seems to only find the *even numbered* limit cycles. Section 4.2 explains this discrepancy.

method of characteristics against the approximate amplitudes found by the method of harmonic balance. We found very good agreement, but noted that the method of characteristics did not seem to find all of the limit cycles predicted by the harmonic balance method (as well as numerical integration). In this section we address this discrepancy. The key comes from the fact that Theorem 5 can be generalized.

We have seen that certain solutions to DDEs come from solutions to corresponding ODEs. But the nature of this correspondence is more general than what is stated in Theorem 5. A more general result is

Theorem 6. *Consider the ordinary differential equation*

$$\frac{dx}{dt} = f(x) \tag{4.8}$$

where $x(t) \in \mathbb{R}^m$ for some m. Suppose that $\hat{x}(t)$ is a solution of (4.8) which satisfies the functional equation $\hat{x}(t-T) = h(\hat{x}(t))$ for some $h : \mathbb{R}^m \to \mathbb{R}^m$. Now consider the

delay differential equation with delay T

$$\frac{dx}{dt} = g(x, x(t-T)) \tag{4.9}$$

where g satisfies the relationship $g(x, h(x)) = f(x)$. Then $\hat{x}(t)$ is a solution to the delay equation (4.9).

Proof. The proof is straightforward:

$$\begin{aligned} g(\hat{x}(t), \hat{x}(t-T)) &= g(\hat{x}, h(\hat{x})) \\ &= f(\hat{x}) \\ &= \frac{d\hat{x}}{dt}. \end{aligned}$$

□

Again, we will call solutions $\hat{x}(t)$ to (4.9) which arise in this way ***characteristic solutions,*** in analogy to the method of characteristics from PDEs.

Corollary 7. *With the same hypotheses of Theorem 6, suppose that a solution $\hat{x}(t)$ of Equation (4.8) satisfies $\hat{x}(t-T) = -\hat{x}(t)$, and suppose $g(x, -x) = f(x)$. Then $\hat{x}(t)$ is a solution to*

$$\frac{dx}{dt} = g(x, -x(t-T)).$$

We illustrate the usefulness of Theorem (6) in the delay Duffing example. Consider the *negative damping* Duffing ODE

$$x'' - x + x^3 = 0. \tag{4.10}$$

According to Corollary 7, any solution to (4.10) that satisfies

$$x(t-T) = -x(t) \tag{4.11}$$

will be a solution to the delayed Duffing equation $x'' + x(t-T) + x^3 = 0$. Like the positive damping Duffing ODE, solutions to Equation (4.10) are periodic (with the exception of equilibrium solutions, and 2 heteroclinic orbits, see [29] for full details). Therefore, if we can find a periodic orbit $\hat{x}(t)$ centered on the origin whose period P satisfies $P = 2T$, then we after time T, $\hat{x}(t)$ will be halfway around its orbit. By the symmetries of Equation (4.10), this implies $\hat{x}(t-T) = -\hat{x}(t)$. Similarly, *any* solution to (4.10)which is halfway around its orbit after time T will satisfy Equation (4.11). The constraint that this imposes on the period P is that $P = \frac{2T}{2n-1}$ for $n = 1, 2, \cdots$. Just as in the case of the positive damping Duffing ODE, solutions to the negative damping Duffing ODE have periods that go to 0 as the amplitude of the solution goes to infinity. Hence, by Corollary 7, we conclude

Claim 8. The delayed Duffing Equation has infinitely many characteristic oscillations which come from solutions to the negative damping Duffing ODE, Equation (4.10).

The amplitude of these additional characteristic solutions can be computed in the same way as what was shown in Figure 4.2, using appropriate Period-Amplitude relationship for the negative stiffness Duffing equation,

$$P(A) = \frac{4K\left(\frac{A^2}{2-2A^2}\right)}{\sqrt{1-A^2}}.$$

The result is pictured in Figure 4.4.

Finally, we note that by considering these more generalized characteristics, we have found the missing limit cycles, predicted by the harmonic balance method in Equation (3.3).

To summarize where we are so far, we have used the method of harmonic balance

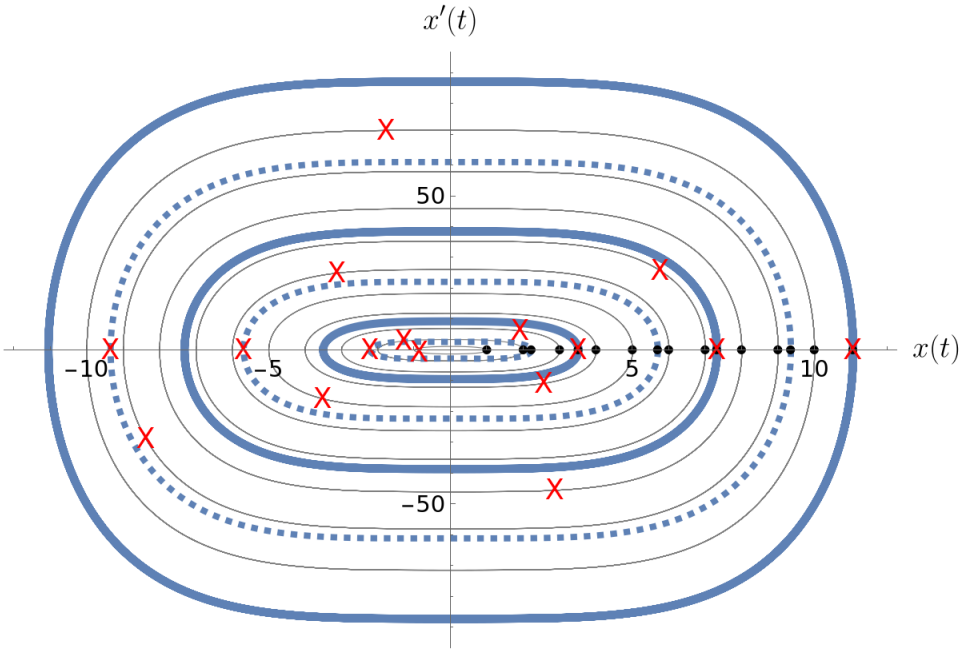

Figure 4.4: As in Figure 4.3, phase portrait of the Duffing ODE. Initial conditions are shown by black dots. After 2 time units, a solution's position is marked with a red X. The characteristic solutions of the negative stiffness Duffing equation are shown in dark blue: for these solutions, after 2 seconds, the particle is exactly half way around its orbit. These solutions satisfy $x(t-2) = -x(t)$, and by Corollary 7, are therefore solutions to the delayed Duffing equation.

on the delay Duffing equation and obtained the surprising prediction that the equation contained infinitely many limit cycles. We then extended the method and were able to calculate approximate values for various interesting bifurcations of those limit cycles. We have now provided a theoretical framework, which we have called the method of characteristic oscillations, to see how these limit cycles arise as exact solutions to certain related ordinary differential equations. In the next section, we further use the method of characteristics to prove a result which generalizes the Duffing example, to find a whole large class of delay equations which contain infinitely many limit cycles.

4.3 Delay Systems with Infinitely Many Limit Cycles

In this section, we step away from the specifics of the delayed Duffing oscillator and try to generalize what we have learned to obtain a more general result about delay systems which contain infinitely many limit cycles. We have seen from Theorem 5 that a sufficient condition for a DDE with delay T to have infinitely many limit cycles is for the corresponding non-delayed ODE to have oscillatory solutions with period $P = T/n$ for infinitely many values of n. This condition was met by the Duffing ODE because it

1. Has periodic orbits for all (non-equilibrium) initial conditions, and
2. Has orbits whose periods go continuously to zero as their amplitudes go to ∞.

It is easy to see that these two conditions (together with the intermediate value theorem) are sufficient to guarantee that an ODE will have infinitely many periodic solutions with period $P = T/n$ for some value of n. This observation enables us to state the following theorem, which suggests that the phenomenon of infinitely many limit cycles is common to many delay systems.

Theorem 9. *Let $H(x, y) : \mathbb{R}^2 \to \mathbb{R}$ be a Hamiltonian function such that $H(0, 0) = 0$ and H is strictly convex (i.e. the Hessian matrix is everywhere positive definite). Let γ_c be the curve defined by the equation $H(x, y) = c$. Further suppose that $\lim_{c\to\infty} \int_{\gamma_c} (\partial_y H(x, y))^{-1}\, dx = 0$. Then the corresponding Hamiltonian vector field*

$$\begin{aligned} x'(t) &= \frac{\partial H}{\partial y} \\ y'(t) &= -\frac{\partial H}{\partial x} \end{aligned} \tag{4.12}$$

will have periodic solutions for all non-equilibrium initial conditions, and the period of those motions goes to 0 *as the amplitude goes to infinity. Therefore, if we let* $\eta_1(x,y,z,w)$ *and* $\eta_2(x,y,z,w)$ *be functions satisfying* $\eta_1(x,y,x,y) = \frac{\partial H}{\partial y}(x,y)$ *and* $\eta_2(x,y,x,y) = -\frac{\partial H}{\partial x}(x,y)$*, then the delayed dynamical system given by*

$$x'(t) = \eta_1(x, y, x(t-T), y(t-T))$$

$$y'(t) = \eta_2(x, y, x(t-T), y(t-T))$$

contains infinitely many limit cycles.

Proof. The condition that H is strictly convex ensures that $(0,0)$ is the only equilibrium point and all other solutions are periodic. This follows from the fact that level sets of convex functions are themselves convex, and strictly convex curves in $\mathbb{R}^2$ are homeomorphic to circles. Hence, the orbits of the system (9) are all periodic. The period of the solution on orbit γ_c can be found be separation of variables:

$$x' = \frac{\partial H}{\partial y} \implies T = \int_0^T dt = \int_{\gamma_c} \frac{dx}{\partial H/\partial y}.$$

Therefore, the condition that

$$\lim_{c\to\infty} \int_{\gamma_c} \left(\frac{\partial H}{\partial y}\right)^{-1} dx = 0$$

is exactly the condition that the period of the oscillations goes to 0 as the amplitude goes to infinity. As discussed above, these two conditions ensure that this vector field will have infinitely many periodic motions with period satisfying $P = T/n$. Finally, Theorem 5 now guarantees that these solutions will be limit cycles of the system

$$x'(t) = \eta_1(x, y, x(t-T), y(t-T))$$

$$y'(t) = \eta_2(x, y, x(t-T), y(t-T))$$

□

Example 10. Consider the Hamiltonian $H(x,y) = ax^{2m} + by^{2n}$ where $a, b \in \mathbb{R}_{>0}$ and $m, n \in \mathbb{Z}_{>0}$ with at least one of m or n greater than 1. We will show that this Hamiltonian satisfies the conditions of Theorem 9, and hence, any delay equation made from such a Hamiltonian will contain infinitely many limit cycles with amplitudes going to infinity.

Verifying that H is strictly convex is trivial. Hence, the Hamiltonian vector field

$$\begin{aligned} x' &= 2nby^{2n-1} \\ y' &= -2max^{2m-1} \end{aligned} \tag{4.13}$$

will have closed orbits for all initial conditions other than $x(0) = y(0) = 0$. We now verify that the period of these oscillations goes to 0 as the amplitude goes to infinity. Consider the initial conditions $x(0) = \mu$ and $y(0) = 0$. Now, let $h = H(\mu, 0)$, so that h represents the energy level of the trajectory with initial condition $(\mu, 0)$, and let γ_h denote the corresponding level curve of H. Then the period $P(\mu)$ of this solution is given by

$$\begin{aligned} P(\mu) &= \int_{\gamma_h} \frac{dx}{\dot{x}} \\ \frac{P(\mu)}{4} &= \int_{\mu}^{0} \frac{dx}{\partial H/\partial y} \end{aligned}$$

We take $P(\mu)/4$ because the integral from $x = 0$ to $x = \mu$ represents a quarter of the oscillation. Now, on γ_h, we have that

$$y = \pm \left(\frac{h - ax^{2m}}{b} \right)^{\frac{1}{2n}}.$$

Therefore, on γ_h, we have

$$\begin{aligned}\frac{\partial H}{\partial y} &= 2nby^{2n-1} \\ &= \pm 2nb\left(\frac{h-ax^{2m}}{b}\right)^{\frac{2n-1}{2n}}.\end{aligned}$$

Because the direction of motion of system (10) is clockwise, a trajectory beginning at $(\mu, 0)$ will have a negative y value for the first quarter period, hence we take the negative value for $\frac{\partial H}{\partial y}$. We therefore have that

$$\begin{aligned}\frac{P(\mu)}{4} &= \int_{\mu}^{0} \frac{1}{-2nb\left(\frac{h-ax^{2m}}{b}\right)^{\frac{2n-1}{2n}}} dx \\ &= \frac{1}{2nb}\int_{0}^{\mu}\left(\frac{h-ax^{2m}}{b}\right)^{\frac{1-2n}{2n}} dx \\ &= \frac{1}{2nb^{1/(2n)}}\int_{0}^{\mu} h^{\frac{1-2n}{2n}}\left(1-\frac{a}{h}x^{2m}\right)^{\frac{1-2n}{2n}} dx.\end{aligned}$$

Our goal is to show that this integral goes to 0 as $\mu \to \infty$. Here it is worth noting that h can be expressed in terms of μ by $h = a\mu^{2m}$. We will use this fact later.

Now it turns out that the integral defining $P(\mu)/4$ can be evaluated exactly in terms of the beta function

$$B(z_1, z_2) = \int_0^1 t^{z_1-1}(1-t)^{z_2-1}dt$$

where z_1, z_2 are in general complex numbers with real part greater than 0. The beta function is closely related to the gamma function: [6]

$$B(z_1, z_2) = \frac{\Gamma(z_1)\Gamma(z_2)}{\Gamma(z_1+z_2)}.$$

To transform our integral for $P(\mu)$ into the form of the beta function, we make the

substitution $v = x/\mu$, so that $dx = \mu dv$. Our integral becomes

$$\begin{aligned}\frac{P(\mu)}{4} &= \frac{1}{2nb^{1/(2n)}} \int_0^{\mu} h^{\frac{1-2n}{2n}} \left(1 - \frac{a}{h}x^{2m}\right)^{\frac{1-2n}{2n}} dx \\ &= \frac{h^{\frac{1-2n}{2n}}}{2nb^{1/(2n)}} \int_0^{\mu} \left(1 - \left(\frac{x}{\mu}\right)^{2m}\right)^{\frac{1-2n}{2n}} dx \\ &= \frac{h^{\frac{1-2n}{2n}}}{2nb^{1/(2n)}} \mu \int_0^1 \left(1 - v^{2m}\right)^{\frac{1-2n}{2n}} dv.\end{aligned}$$

Next, let $t = v^{2m}$. Then $dt = 2mv^{2m-1}dv$, or equivalently, $dv = \frac{1}{2m}t^{\frac{1-2m}{2m}}dt$. Then we have

$$\frac{P(\mu)}{4} = \frac{h^{\frac{1-2n}{2n}}}{2nb^{1/(2n)}} \mu \frac{1}{2m} \int_0^1 (1-t)^{\frac{1-2n}{2n}} t^{\frac{1-2m}{2m}} dt.$$

The integral is now exactly $B\left(\frac{1}{2n}, \frac{1}{2m}\right)$, which has no dependence on the initial condition μ. With some algebra, we can now simplify and find that the period $P(\mu)$ is given by

$$P(\mu) = \frac{1}{nm} b^{-\frac{1}{2n}} a^{-\frac{1}{2m}} B\left(\frac{1}{2n}, \frac{1}{2m}\right) \cdot h^{\frac{1}{2n}+\frac{1}{2m}-1} \tag{4.14}$$

where, again, $h = a\mu^{2m}$. Importantly, the only term in Equation (4.14) that depends on the initial condition μ is the factor $h^{\frac{1}{2n}+\frac{1}{2m}-1}$. Note that if $m > 1$ or $n > 1$, the exponent $\frac{1}{2n} + \frac{1}{2m} - 1$ is less than 0. Hence, as $\mu \to \infty$, $h^{\frac{1}{2n}+\frac{1}{2m}-1} \to 0$, and so $P(\mu) \to 0$.

This shows that the system (10) satisfies the conditions of Theorem 9. Hence any delayed version of system (10) will have infinitely many limit cycles, with amplitudes going to infinity and periods going to 0. For example, the system

$$\begin{aligned}x' &= y \\ y' &= -x(t-T)^{2m-1} - x^{2m-1}\end{aligned}$$

will have infinitely many limit cycles for $m > 1$.

Corollary 11. *There is no upper bound to the number of limit cycles for polynomial delay systems when the polynomial is odd and of degree greater than* 1.

The above theorem and corollary give an emphatic answer to Hilbert's 16th problem in the delay case — not only is there no upper bound to the number of limit cycles that can exist in a polynomial delay equation, it is in fact quite easy to cook up examples which have infinitely many limit cycles! This illustrates one of the many ways in which the dynamics of delay equations contain greater complexities than anything which can arise in the ODE case.

For the sake of completeness, we note that significant weakening of the hypotheses of Theorem 9 is possible, and that the conclusion of infinitely many limit cycles will still hold. In Theorem 9, the hypothesis that H is strictly convex guarantees that the solution curves will all be closed and convex. The convexity assumption is convenient for being easy to check, but is not strictly necessary. Additionally, Theorem 9 assumes that the period on closed orbit goes to 0 as the amplitude goes to infinity — analogous to what happens in the Duffing oscillator — but for the result to hold, it doesn't matter which solutions have period going to 0. The fully general statement is below — in practice, the weaker assumptions can be difficult to check.

Theorem 12. *Let $x' = F(x, y)$, $y' = G(x,y)$ be a smooth dynamical system in the plane. Suppose there exists an open set $\Omega \subset \mathbb{R}^2$ such that Ω is foliated by closed solution curves of this system. Further, let $T(\gamma)$ denote the period of the solution on*

the curve γ, and suppose that $\sup_\gamma\{T(\gamma) \mid \gamma \subset \Omega\} > 0$ *and* $\inf_\gamma\{T(\gamma) \mid \gamma \subset \Omega\} = 0$. *Then any delayed version of this system will contain infinitely many limit cycles.*

(By a delayed version, we mean a system of the form $x' = \eta_1(x, y, x(t-T), y(t, -T))$, $y' = \eta_2(x, y, x(t-T), y(t-T))$, *where* $\eta_1(x, y, x, y) = F(x, y)$ *and* $\eta_2(x, y, x, y) = G(x, y)$.*)*

Proof. The proof is essentially identical to the proof of Theorem 9, and is therefore left as an exercise to the reader. □

Finally, we note that Theorem 12 can also be generalized to systems of degree greater than 2, in much the same way.

4.4 Conclusion

In this chapter, we introduced a novel technique for studying oscillations in delay differential equations: the method of characteristic oscillations. Inspired by the method of characteristics from the theory of partial differential equations, this approach identifies special solutions of DDEs on which the dynamics reduce to those of an associated ODE. We believe this technique is a powerful tool for analyzing delay-induced oscillations, as it allows one to leverage the full analytical machinery developed for ODEs.

Using this method, we confirmed that the infinitely many limit cycles predicted via harmonic balance in the delayed Duffing equation do indeed exist, and that their explicit forms are given by Jacobi elliptic functions. We also derived sufficient

conditions for the existence of such families of periodic solutions and exhibited an infinite class of DDEs that possess this structure.

This analysis underscores a fundamental distinction between delay and ordinary differential equations. Whereas it is known that no polynomial planar ODE can admit infinitely many limit cycles, such behavior appears to be common—perhaps even generic—in the delay setting.

We close with a few open questions for further investigation. While the method of characteristic oscillations provides detailed information about the amplitude, period, and sometimes even closed-form expressions of periodic solutions, it currently offers no insight into their stability. In the case of the delayed Duffing equation, it is known (see Fiedler et al. [12]) that characteristic oscillations arising from $x'' - x + x^3 = 0$ are all stable, whereas those from $x'' + x + x^3 = 0$ are all unstable. Whether a general method exists for determining the stability of characteristic oscillations remains an open problem. The author has attempted to apply the infinite-dimensional generalization of Floquet theory (see Hale and Lunel [16], Chapter 9), but without success so far.

A final, tantalizing question arises: Is every periodic solution of a DDE a characteristic oscillation? That is, given a known periodic solution to a delay equation, can one always find an associated ODE that governs the dynamics on that orbit? We leave this intriguing possibility to future researchers.

CHAPTER 5

CHAOS IN DDES

In this chapter, we further investigate the chaotic behavior observed in the delayed Duffing equation at the end of Chapter 3. Recall that for delay values greater than approximately $T \approx 3.61$, the solutions appear aperiodic and erratic. Figure 5.1 shows a phase plot and frequency spectrum—obtained via the Fast Fourier Transform—of a trajectory at this delay. The frequency spectrum reveals a continuous band of frequencies, suggesting that the trajectory is composed of many incommensurate modes. This suggests the trajectory is aperiodic, one of the hallmarks of chaos.

We begin the chapter with a brief overview of chaotic dynamics in ODEs and highlight some of the challenges in extending this theory to delay systems. We then introduce the concept of the dimension of a limit set, with particular attention to chaotic attractors and their fractional (non-integer) dimensions. In particular, we explore the concept of correlation dimension and compute it for the attractor arising in the delayed Duffing equation.

Finally, motivated by an intriguing conjecture due to Kaplan and Yorke [13], we introduce the Lyapunov spectrum associated with chaotic attractors. This spectrum is conjectured to relate directly to the attractor's dimension. We compute the Lyapunov spectrum for several values of the delay T in the Duffing equation, and present numerical evidence supporting the idea that the Kaplan–Yorke conjecture

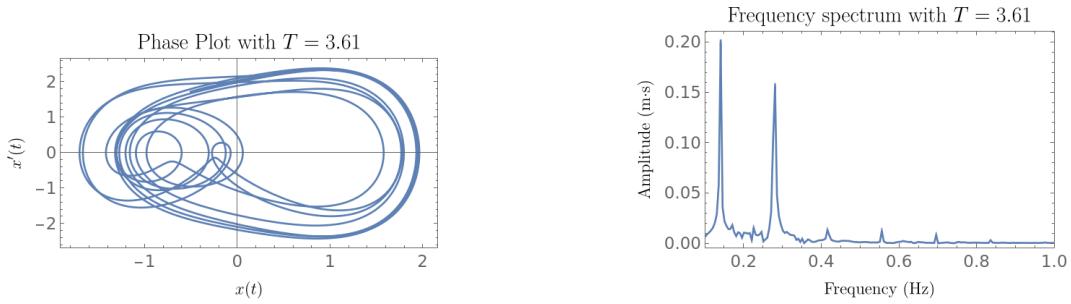

Figure 5.1: Erratic behavior of $x'' + x(t - T) + x^3 = 0$ for $T = 3.61$. Above this value of T, the Fourier frequency spectrum reveals that the trajectory is composed of continuous band of frequencies.

may extend to delay systems.

5.0.1 Background

The phenomenon of chaos occupies a central and still-evolving place in the study of nonlinear dynamical systems. First brought to widespread attention by Edward Lorenz in the 1960s during his attempts to model atmospheric convection, chaotic dynamics emerged as a striking feature of deterministic systems with only a few degrees of freedom. The now-famous Lorenz system—three coupled nonlinear differential equations—exhibited solutions that, though governed by simple rules, displayed long-term behavior that appeared random and unpredictable. This sensitivity to initial conditions, famously paraphrased as the "butterfly effect," meant that even minute differences in starting points could lead to drastically different outcomes over time.

Since Lorenz's discovery, chaotic behavior has been found throughout the sciences—in the irregular oscillations of certain chemical reactions, in turbulence and convective flows in fluid dynamics, in the boom-and-bust cycles of population models in ecology, and even in models of cardiac rhythms and the dynamics of the brain. The recognition

that deterministic systems can behave unpredictably has reshaped how scientists think about prediction, control, and the limits of modeling in complex systems. See [35] Chapter 9 for a nice history of the Lorenz system and the subsequent developments in the theory of chaos.

Despite decades of research, there remains no single, universally accepted mathematical definition of chaos. However, several related features often appear together and help to characterize chaotic behavior. One influential perspective, due to Devaney, applies to dynamics induced by a continuous map on a metric space, $f : X \to X$. Devaney defines a chaotic system in this setting as one that is topologically transitive, has dense periodic orbits, and exhibits sensitive dependence on initial conditions [7]. These three properties together imply a form of deterministic unpredictability: the system evolves according to fixed laws, but its long-term behavior resists concise prediction. (As an interesting aside: in 1992, it was shown by Banks et al. that for discrete dynamical systems, only the first two criteria are necessary — if a system is topologically transitive and has dense periodic orbits, it necessarily has sensitive dependence on initial conditions! See [1] for a remarkably short proof).

Beyond the topological view, there is also an ergodic-theoretic perspective on chaos, in the context of *measure preserving dynamical systems*. At a high level, this approach studies the statistical properties of trajectories in phase space, focusing on how typical orbits distribute over time. In this setting, chaos is often associated with the presence of positive Lyapunov exponents—indicating exponential divergence of nearby trajectories—and with invariant measures that describe the long-term statistical behavior of the system. See [23] for a very readable introduction to this

perspective. While this probabilistic lens requires a different set of tools than the topological one, the two perspectives both recognize chaos as a deeply intricate and pervasive phenomenon in nonlinear systems.

We will not delve into the formal machinery of either the topological or measure-theoretic perspectives on chaos, as these frameworks are not directly relevant to the work presented in this thesis. We mention them here primarily to highlight the contrast with the state of chaos theory in delay systems. Due to the infinite-dimensional nature of the phase space in delay differential equations (DDEs), there has been little success in extending the ergodic theory approach to this setting. The fundamental technical hurdle lies in defining a suitable notion of "measure" on a space of functions. For example, in the Banach space $C([-T, 0], R)$ the unit ball is non-compact, which complicates the construction of invariant measures and the application of classical ergodic tools. The challenge of adapting the measure-theoretic perspective to DDEs remains a major open problem. See [22] for a discussion of existing approaches and unresolved difficulties.

As a result, many of the powerful techniques from ergodic theory are unavailable to us in the delay setting. In studying the chaotic behavior of the delayed Duffing equation, we must therefore rely primarily on tools from the topological viewpoint. The next two sections develop this perspective in detail.

5.1 Sensitive Dependence: The Lyapunov Spectrum

A fundamental tool in the study of chaotic dynamical systems is the *Lyapunov Spectrum,* which provides a measurement of the sensitivity of long term outcomes to small perturbations of initial conditions. The idea is the following. Given an initial condition x_0, take an infinitesimally small ball of points centered on x_0. As you continue along the trajectory of x_0, the ball will be distorted. Because the initial ball is of infinitesimal radius, the dynamics will be governed by a linearized system, so the resulting shape will stretch into an ellipsoid. If the axes of this ellipsoid are sorted in descending order of length, then the i^{th} Lyapunov exponent λ_i quantifies the average exponential rate of expansion or contraction of the i^{th} axis. The collection of all Lyapunov exponents is called the Lyapunov Spectrum. Figure 5.2 illustrates this idea.

To formalize this intuitive geometric picture somewhat, consider a perturbation that is aligned with the i^{th} axis of the stretched ellipsoid. Call the length of this axis at time t $\epsilon_i(t)$. Because the dynamics of $\epsilon_i(t)$ will be governed by a linear system, we should have that $\epsilon_i(t) = \epsilon_i(0)e^{\lambda_i t}$. Solving for λ_i, and then imposing that we want the initial perturbation to be infinitesimal and that we are interested in the asymptotic expansion rate, we obtain the following formal definition for the i^{th} Lyapunov exponent:

$$\lambda_i = \lim_{t\to\infty} \lim_{\epsilon_i(0)\to 0} \frac{1}{t} \log \frac{\epsilon_i(t)}{\epsilon_i(0)}.$$

From this intuitive geometric definition, it is clear that if the phase space of a system is n-dimensional, then the system will have n different Lyapunov exponents. Now, the condition that a system has sensitive dependence on initial conditions

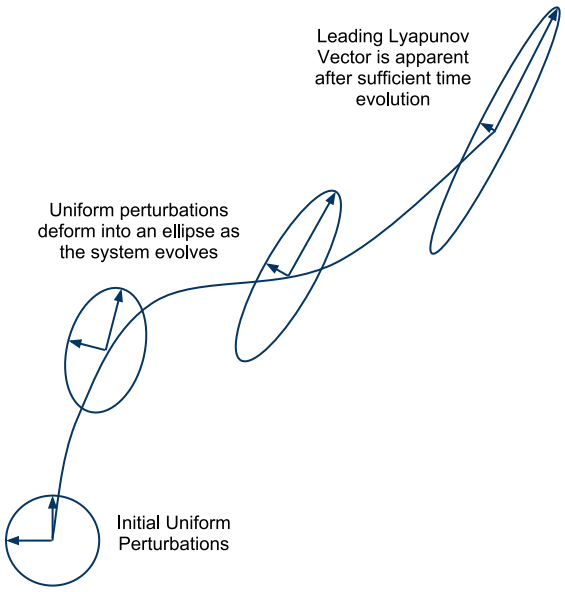

Figure 5.2: A diagrammatic depiction of the stretching of a ball of initial perturbations into an ellipse. The average exponential rate of stretching of the i^{th} axis corresponds to the i^{th} Lyapunov exponent λ_i. This diagram adapted from "Leading Lyapunov Exponent" by Mrocklin, licensed under CC BY-SA 3.0 via Wikimedia Commons: https://commons.wikimedia.org/w/index.php?curid=11995643.

means that nearby trajectories tend to diverge from each other in the long run. The divergence of nearby trajectories implies that at least one of the Lyapunov exponents is greater than 0, which is a common hallmark of chaos. This is the most common usage of Lyapunov exponents, and for many purposes just computing the largest Lyapunov exponent suffices. However, we can say more. For a trajectory to converge to a chaotic attractor, which occupies a bounded region of phase space, the system must be dissipative (at least locally), which implies that at least one of the Lyapunov exponents must be negative. Finally, it can be shown that for trajectories which remain bounded, one Lyapunov exponent must always be 0, corresponding to the flow direction. See for example [30, 33] for details. We therefore expect a chaotic finite dimensional system to have a Lyapunov spectrum consisting of at least one positive Lyapunov exponent, one exponent equal to 0, and at least one negative exponent.

Much of this discussion carries over to the delay setting and function spaces. A theorem of Ruelle guarantees that a Lyapunov spectrum exists for infinite dimensional dynamical systems, and that the spectrum is discrete and contains at most finitely many positive exponents. The proof can be found in [28], and an accessible discussion can be found in [9].

5.1.1 Calculating Lyapunov Exponents in ODEs

Directly computing the Lyapunov exponents from their definition — as time-averaged growth rates of infinitesimal perturbations — quickly runs into numerical difficulty. In systems with both expanding and contracting directions, the components of a

tangent vector aligned with contracting directions shrink exponentially, often falling below machine precision within a short integration window. This means that after enough time, only the dominant expanding direction remains detectable in the numerical data, making it impossible to resolve the full Lyapunov spectrum.

To overcome this, one instead simulates the evolution of an entire basis of tangent vectors along with the trajectory itself. This is done using the *variational equation*, which is the linearization of the nonlinear system about a trajectory. If $\dot{x} = f(x)$ is the original system of order n, then the variational equation is given by

$$\dot{\delta x} = Df(x(t)) \cdot \delta x,$$

which describes how infinitesimal tangent vectors to an initial condition x_0 evolve in tangent space. Here, Df represents the Jacobian matrix of f. The variational equation allows for a local, linear description of divergence or convergence between nearby trajectories. This equation arises by considering two nearby trajectories$x(t)$ and $x(t) + u(t)$, expanding $f(x + u) \approx f(x) + Df(x) \cdot u$, and subtracting the equations of motion to find how the perturbation evolves. Therefore, an improved strategy for computing the Lyapunov spectrum would be to form a matrix of n initial perturbations M_0, and forming the augmented equations

$$\begin{aligned} \dot{x} &= f(x) \\ \dot{M} &= Df(x(t)) \cdot M \end{aligned} \tag{5.1}$$

with initial conditions $x(0) = x_0$ and $M(0) = M_0$. However, even when integrating the variational equation for a full basis of initial perturbation vectors, numerical issues persist. Over time, all vectors tend to align with the direction of maximal

expansion due to exponential divergence – once again overwhelming information about the contracting or neutral directions. To address this, one uses a technique of periodic re-orthonormalization, where after discrete time intervals the tangent vectors are orthonormalized, and the logarithms of their stretch factors are accumulated. This prevents the vectors from all aligning with the dominant expansion direction. The Oseledec's Multiplicative Ergodic Theorem [25] guarantees that this process converges, under mild conditions, to the true Lyapunov spectrum. In practice, instead of using the classical Gram–Schmidt procedure to re-orthonormalize – which is numerically unstable – one uses the QR factorization: at each re-orthonormalization step, the matrix of tangent vectors is factored as $M = QR$, where Q contains orthonormal columns and R is upper triangular and contains the singular values of M on its diagonal. Therefore, the diagonal entries of R represent the stretching factors along the orthogonal directions during that time interval. After each discrete time step, the logarithms of the diagonal of R are accumulated. The resulting values are then divided by the integration time. This method, known as the Benettin algorithm, provides a robust and efficient numerical approach for estimating the full Lyapunov spectrum of a dynamical system [3, 33, 30].

5.1.2 Calculating Lyapunov Exponents in DDEs.

The Benettin algorithm for calculating Lyapunov exponents is fairly standard in the world of ODEs, and is based entirely on rigorous theory. The situation is not so clear in the delay setting. The first and most obvious point is that in the delay setting, there are infinitely many Lyapunov exponents, making their calculation theoretically

impossible. Another complication is that the infinite dimensional analogue to the variational equation involves the Frèchet derivative, which is unpractical to implement directly for numerical methods.

However, all is not lost. There are methods for approximating a DDE by a high dimensional ODE. This approximation allows one to then use the ODE algorithms to compute the Lyapunov spectrum — or at least an initial segment of it — for DDEs. The most elementary of these ODE approximation methods goes by the name *the method of lines*, which we describe below. This approach is laid out in [34].

Consider a DDE

$$\frac{dx}{dt} = F(x(t), x(t-T)).$$

The idea is that instead of evolving a whole function over the interval $[-T, 0]$, we will discretize the function into N equally spaced sample points, and evolve each of them in time. This will reduce the problem from evolving infinitely many points to finitely many.

To do so, fix an integer N, and let $h = T/n$ denote the step size between grid points. Define $x_0, \ldots, x_N$ by $x_i(t+h) = x_{i-1}(t)$ for $i = 0, \ldots, N$. The index i indicates how many steps of size h the variable x_i is behind the "current time" variable, x_0. For example, $x_N(t) = x_0(t - Nh) = x_0(t - T)$. From this observation, it's clear that the differential equation governing x_0 is given by $x_0' = F(x_0, x_N)$. For $1 \leq i \leq N-1$, we can define $\frac{dx_i}{dt}$ simply by observing that the derivative of x_i is just the slope of the solution curve passing through x_i, which can be approximated using centered finite differences. For the last point x_N, we use just a right-handed finite difference

approximation. In all, we obtain the system

$$\begin{aligned} \frac{dx_0}{dt} &= F(x_0, x_N) \\ \frac{dx_i}{dt} &= \frac{x_{i-1} - x_{i+1}}{2h} \text{ for } 1 \leq i \leq N-1 \\ \frac{dx_N}{dt} &= \frac{x_{N-1} - x_N}{h}. \end{aligned} \tag{5.2}$$

This system of $N+1$ ODEs can then be solved using any standard ODE solver. Moreover, using this setup, one can form the variational equation and implement the Benettin algorithm to compute the Lyapunov spectrum, as described in Section 5.1.1.

Being an ODE system of order $N+1$, system (5.2) will have $N+1$ Lyapunov exponents. How can we be sure that these Lyapunov exponents accurately reflect the Lyapunov exponents of the full, infinite dimensional DDE? To our knowledge, there are no tight theoretical guarantees of this. In practice, however, one can simply compute the Lyapunov spectrum for increasing values of N, and track how the largest (and therefore most relevant) exponents change as a function of N. As observed in [9], the leading exponents quickly stabilize as a function of N, giving the practitioner confidence that the exponents accurately represent the full, infinite dimensional system.

In general, this method of approximating our DDE by an ODE is quite good. Figure (5.3) compares the result of integrating the full delayed Duffing equation using the method of steps against the approximation using a finite order ODE. The agreement is quite good. This gives us further confidence in the method's ability to accurately represent the dynamics of the full DDE.

We therefore now have a scheme for computing arbitrarily many Lyapunov exponents

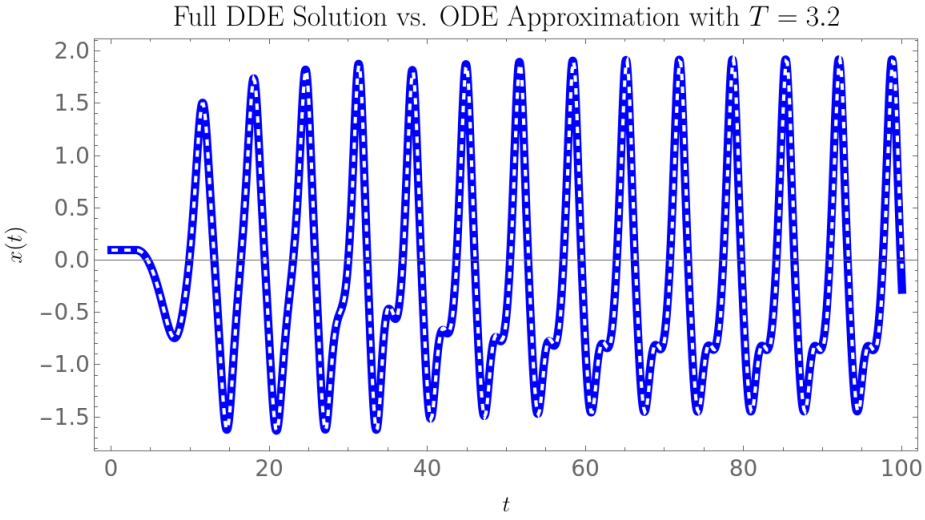

Figure 5.3: A comparison of integrating the full DDE $x'' + x(t - T) + x^3 = 0$ using the method of steps (thick blue curve) vs integrating the ODE approximation using Equation (5.2) with $N = 50$ (dashed white curve). The agreement is so good as to make the curves indistinguishable.

for a DDE. First, we employ the method of lines given by Equations (5.2) to find a system of ODEs approximating the original system. We can then form the variational equation (5.1) of the resulting system, and then use the Benettin algorithm. For this work, we implement this algorithm in the language Julia. All of the code is provided in the Appendix (7).

Below in Figure 5.4 is a plot of just the largest Lyapunov exponent of the delay Duffing equation, as T ranges from 2 to 50.

Our results in Figure 5.4 verify that the delay Duffing equation does indeed have sensitive dependence on initial conditions. Unfortunately, the Benettin algorithm for delay equations scales quadratically in complexity as a function of T, making fine-grained resolution of λ_1 over large scales infeasible on my little laptop, despite all

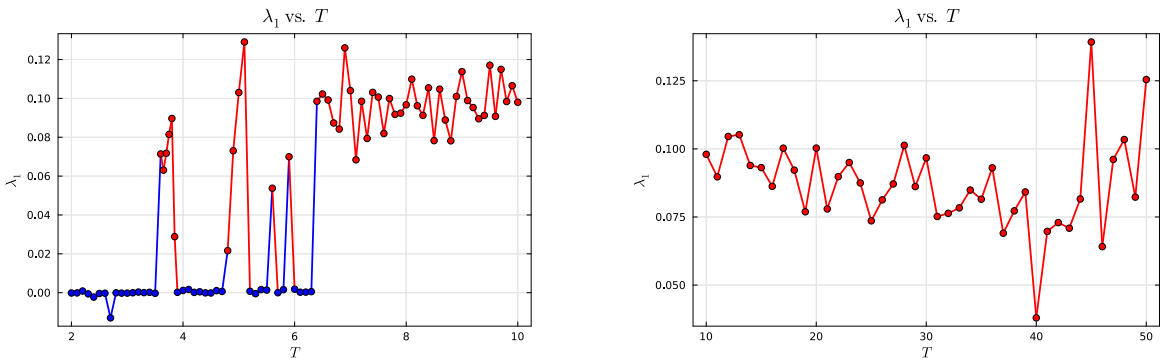

Figure 5.4: Computation of the leading Lyapunov exponent of the delay Duffing equation as a function of the delay T. Computations employed the Benettin algorithm in Julia, listed in Appendix 7.1, using 30000 steps and $N = 20T$.
Left: λ_1 as T ranges from 2 to 10 in increments of .1. Points colored blue have $\lambda_1 < 0$, corresponding to periodic motions. Points colored red have $\lambda_1 > 0$, corresponding to chaos. There is an initial window of chaos between $T = 3.6$ and $T = 3.9$, followed by windows of stability and windows of chaos. Beyond $T = 6.3$, if there are windows of stability, they are too small to be detected here.
Right: λ_1 as T ranges from 10 to 50 with increments of 1. The entire range has positive Lyapunov exponents, with no detected windows of stability.

the optimizations. In fact, quadratic complexity in T is the best case scenario. The difficulty comes from the fact that to use the method of lines to resolve sufficient detail for the ODE approximation to remain faithful to the full DDE, one must use, as a rule of thumb, $N \approx 20T$ sample points. This means that the order of the simulated ODE increases linearly with T. The real slow down comes from the need to QR factorize the perturbation matrix every step, as this step is quadratic in the size of the perturbation matrix. Additional slow down is possible if the step size needed for the integrator to continue meeting tolerances gets very small, which is absolutely possible over a large range of T values. We are therefore somewhat limited by computational capacity.

5.2 Dimension of Chaotic Attractors

To better understand and quantify chaotic behavior, it is often useful to examine the dimension of a chaotic attractor. This idea, while perhaps sounding abstract and somewhat arcane at first, offers an intuitive handle on the complexity of the system. In a seminal study of the Mackey-Glass equation—a delay differential equation known to exhibit rich chaotic dynamics—Farmer [9] showed that the attractor's dimension increases as the delay parameter grows. We follow a similar line of inquiry in the present context.

Chaotic attractors typically possess non-integer (fractal) dimensions, a hallmark of their intricate internal structure. This fractional dimension reflects the fact that the system's dynamics, while deterministic, evolve on a set that is more complex

than a smooth surface but not as unconstrained as the full phase space. One way to interpret this is to view the attractor's dimension as an estimate of the number of independent degrees of freedom participating in the chaos. For example, in a simplified weather model, a low-dimensional chaotic attractor might suggest that only a handful of variables (like temperature and pressure) are behaving unpredictably. In contrast, higher-dimensional chaos would imply that many more variables are fluctuating in an interconnected and erratic fashion, making the system's long-term behavior vastly harder to forecast.

In this section, we will give a quick introduction to the idea of fractional dimension through a discussion of Hausdorff dimension. We will then introduce a computationally feasible method for computing the dimension of a chaotic attractor, called the correlation dimension. Finally, we will describe an algorithm for computing the correlation dimension of a chaotic attractor in a DDE, and implement the algorithm in the study of the chaotic delay Duffing equation.

5.2.1 Fractional Dimension: Introduction

To begin, we first introduce the notion of fractional dimension, which initially can be quite a bewildering idea. What does it mean for a shape to be 1.3-dimensional, for example?

To deal with this question, we first need to examine what it means for a shape to be n-dimensional for integer n. We are used to thinking of dimension as representing the number of linearly independent directions in a vector space, for example. In this setting, dimension is a sort of cardinality, and obviously cannot be fractional.

An extension of this idea comes in the form of *Hausdorff dimension.* The Hausdorff dimension offers a way of computing the dimension of a set which agrees with our normal intuition for simple sets — such as lines, squares, and cubes — but can take on fractional values for more complex sets, such as self-similar fractals.

The core idea comes from noticing how the volume of shapes change when they are scaled. When a line segment is scaled by a factor of 2, its "volume" (i.e. its length) doubles. When a square is scaled by a factor of 2, its "volume" (i.e. area) increases by a factor of 4. When a cube is scaled by a factor 2, its volume increases by a factor of 8. Clearly, when an n-dimensional "cube" (which is to say, the Cartesian product of n intervals $I \times \cdots \times I$) is scaled by a factor of 2, its volume will increase by a factor of 2^n. This observation gives us a way of generalizing the notion of dimension — when a shape is scaled by a factor of 2, suppose its resulting volume is V_{scaled}. Then dimension of the shape then is the value of d such that $2^d = V_{scaled}$. In other words,

$$d = \frac{\log(V_{scaled})}{\log 2}.$$

This gives us the notion of Hausdorff dimension.

Now, to see how this idea can give rise to fractional dimension for more complicated sets, consider the well known Sierpinski triangle. As a reminder, the Sierpinski triangle is obtained by iteratively removing the middle third triangle from an equilateral triangle. The resulting figure in the limit does somehow seem to be in between one and two dimensional — it appears to have two independent directions, and yet it appears to take up zero area, being composed only of line segments. Hence, it appears to be a kind of in between shape. We can use the notion of Hausdorff dimension to

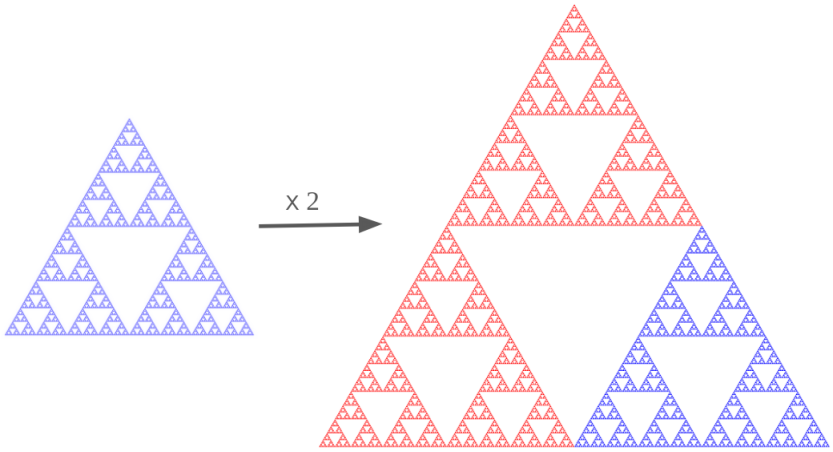

Figure 5.5: When the blue Sierpinski triangle has its side length scaled by a factor of 2, the overall volume of the triangle apparently triples. This implies the it has a Hausdorff dimension of $\log(3)/\log(2) \approx 1.58$.

calculate this.

As shown in Figure 5.5, scaling the Sierpinski triangle by a factor of 2 increases the "mass" of the triangle by a factor of 3. Hence, using the Hausdorff dimension, we find that $2^d = 3$, so that $d = \log(3)/\log(2) \approx 1.58$. It seems, then, that the Sierpinski triangle is closer to being 2 dimensional than 1 dimensional, but not by much.

Of course, formalizing all of this would require appropriately defining the "volume" of a fractal. Obviously the standard Lebesgue measure won't do the trick, since the Lebesgue measure of the Sierpinski triangle will always be 0. We skirt these details for the sake of brevity. In any case, hopefully this gives the reader an intuit idea of how it can be that a shape has a fractional dimension. This is the situation for the chaotic attractors which often appear in continuous dynamical systems. These objects have incredible intricate and complex structure, which often means their dimension is best described by a fractional value. For example, the well-known

Lorenz attractor is known to have Hausdorff dimension between 2 and 3, though the precise value is not known. [24]

We note that there is a way of extending Hausdorff dimension to sets which are not self similar, thereby extending the usefulness of the Hausdorff dimension. This extension goes by the name *box counting dimension.* We do not go into details here, but refer the reader to [35] Chapter 11 for details.

Finally, as a little exercise for the interested reader who wishes to test their understanding: Show that the Hausdorff dimension of the middle thirds Cantor set is $\log(2)/\log(3)$.

5.2.2 Correlation Dimension

In practice, the Hausdorff dimension — and even the more tractable box counting dimension — can be extremely difficult to compute. In this section, we describe another notion of fractional dimension, which is much more computationally feasible. This is the notion of *correlation dimension*, introduced in [14].

To build intuition for the correlation dimension, imagine randomly sampling a large number of points from a geometric shape and measuring how those points cluster at different scales. Specifically, for each point, consider counting how many other points lie within a ball of radius r centered at that point. By averaging this count over all points, we obtain a function $C(r)$ that describes the typical number of neighbors within radius r.

For example, picture sampling points uniformly inside a solid cube in $\mathbb{R}^3$. The average number of neighbors within radius r should scale roughly like the volume of

a ball in 3D space:

$$C(r) \sim r^3.$$

This cubic scaling reflects the three-dimensional nature of the solid cube from which the points were sampled. More generally, for a set with integer dimension d, $C(r)$ scales approximately as

$$C(r) \sim r^d$$

Therefore, the correlation dimension will be defined in terms of

$$d \sim \frac{C(r)}{\log(r)}$$

The correlation dimension successfully recovers the familiar integer dimensions for simple geometric objects, while providing a generalization to fractal and more complicated sets.

To be more precise, for a collection of N points, sampled from some kind of distribution on a set X. Then we first define the *correlation function* to be the average number of points within r of a given point, in the limit as N goes to infinity:

$$\begin{aligned} C(r) &:= \lim_{N\to\infty} \frac{1}{N^2} \left|\{(x_i, x_j) : |x_i - x_j| < r\}\right| \\ &= \lim_{N\to\infty} \frac{1}{N(N-1)} \sum_{i<j}^{N} H(r - |x_i - x_j|) \end{aligned}$$

where H denotes the Heaviside step function.

Based on the discussion above (and demonstrated in [14]), the correlation dimension

will scale as a power as the radius varies:

$$C(r) \sim r^{\nu}.$$

It is this ν that we're after. Now, we can't simply say

$$\nu = \frac{\log C(r)}{\log r}$$

because this quantity depends on r. A moment's thought reveals that the above quantity depends on r. For example, if X is compact, then when r is greater than the diameter of X, then all the points on X would be with r of each other, so $C(r)$ would become constant. Meanwhile, $\log r$ would go to ∞, so ν would always go to 0. This motivates the idea we should take r as small as possible, to just capture the local structure of X. Hence, we define

$$\nu = \lim_{r \to 0} \frac{\log C(r)}{\log r}$$

when the limit is defined.

In practice, of course, we are limited to sampling a set only finitely many times, and so we cannot compute the correlation dimension precisely. However, we can estimate it by examining how the correlation sum $C(r)$ scales with the distance parameter r. If the attractor has a well-defined correlation dimension ν, then for sufficiently small r we expect the relation

$$C(r) \sim r^{\nu}$$

to hold approximately. Taking logarithms of both sides yields the linear relationship

$$\log C(r) \approx \nu \log r + \text{const.}$$

Thus, one plots $\log C(r)$ versus $\log r$, and looks for a linear region in the plot. The slope of this linear segment gives an empirical estimate of the correlation dimension ν. For values of r that too are small, balls of radius r won't contain any other sample points, so the plot for small values of r will appear constant. For r to large, a sphere of radius r would contain the entire data set, so again, the log-log plot would become constant. It is therefore the intermediate ranges of r values where one gets a good estimate for the correlation dimension. Figure 5.6 illustrates the idea of correlation dimension.

In dynamical systems, the correlation dimension is typically computed from a trajectory by selecting points at different times along the orbit. Since the trajectory spends more time in some regions than others, the sampled points tend to cluster non-uniformly, with denser regions corresponding to frequently visited states. Thus, the correlation dimension captures the effective dimension in a dynamic sense.

One subtlety is that the correlation dimension is known to underestimate the more rigorous Hausdorff dimension, $D_{\text{correlation}} \leq D_{\text{Hausdorff}}$. However, the correlation dimension is often very close to the Hausdorff dimension [14].

We note that there are many other notions of fractional dimension that can be computed for chaotic attractors, most of which have advantages and drawback depending on the application. We refer the reader to [9] for a nice summary of all of the different measurements. We gloss over these, because for our purposes we are focused on the computationally feasible correlation dimension.

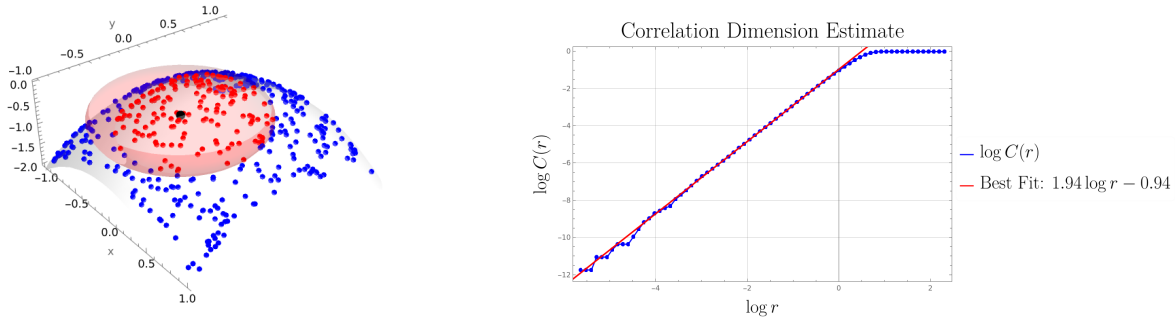

Figure 5.6: An example correlation dimension calculation, for the surface $z = -(x^2 + y^2)$.
Left: A depiction of the surface with a random distribution of points. The red sphere highlights all of the points within radius r of a given point x_0. As $r \to 0$, the number of points within r of x_0 will scale as r^2, reflecting the two dimensional nature of the surface.
Right: $C(r)$ is the average number of points within r of a given point, averaged over all points on the surface. We expect $C(r)$ to scale as r^2. The plot shows that $\log C(r) \sim 1.94 \log r$ during its most linear regime, which is very close to 2. Note that the linear approximation is only valid for intermediate values of r: when r is too small, there are not enough points that are within r of each other on average to get good scaling; when r is too large, the sphere encompasses the whole surface.

5.2.3 Computing Correlation Dimension in Delay Equations

In the preceding section we introduced the idea of correlation dimension as a computationally feasible method for computing the fractional dimension of chaotic attractors. In this section we describe the process by which we compute the correlation dimension for the delay Duffing equation.

To compute the correlation dimension of an attractor arising from a delay differential equation (DDE), we begin by numerically integrating the DDE over a long time interval—typically on the order of 2000 time units—to allow transients to decay and ensure that the system is evolves on its attractor for a sufficiently long time. From this trajectory, we extract a finite set of sample points from the tail end of the solution. Interpolation is used if necessary to ensure that samples are evenly spaced in time. This yields a dataset $\{x_1, x_2, \ldots, x_N\}$, which is treated as a representative sampling of the attractor in phase space. To be precise, if $x(t)$ represents the full time series obtained from numerical integration, then each x_i corresponds to evaluating x at some time, $x_i = x(t_i)$, where the time samples are equally distributed: $t_{i+1} - t_i = \Delta t$ for each i.

Next, we apply a *delay embedding* to these samples, forming reconstructed state vectors of the form

$$\boldsymbol{x}_i = (x(t_i), x(t_i - \tau), x(t_i - 2\tau), \ldots, x(t_i - m\tau))$$

for each index i. The goal of this step is to *reconstruct the geometry of the attractor* in a finite-dimensional Euclidean space. This procedure is based on the Takens' Embedding Theorem [37], which states that for generic smooth dynamical systems,

the geometry of an attractor can be faithfully reconstructed using time-delayed observations of a single scalar observable. Technically, Takens' original result applies only to attractors that are smooth manifolds. For more general attractors, possibly with fractal geometry, (such as chaotic attractors) the appropriate extension is due to Yorke et al. [31]. Their theorem asserts that if a limit set X of a smooth dynamical system has Hausdorff dimension d, then almost every delay embedding into $\mathbb{R}^m$ is an embedding, provided $m > 2 \cdot \dim_{\text{Hausdorff}}(X)$.

In practice, this means that one must choose the embedding dimension m large enough to capture the structure of the attractor, but the delay time τ itself need not be finely tuned—almost all choices will work, provided mm is sufficiently large. ("Almost every" is used here in the measure theoretic sense).

With the embedded dataset $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ in hand, we compute the correlation sum, $C(r)$, described in the previous section, which quantifies the average number of pairs of points that lie within a distance r of each other:

$$C(r) = \frac{1}{N(N-1)} \sum_{i<j} H\left(r - \|x_i - x_j\|\right),$$

where H is the Heaviside step function. We note that this computation is $\mathcal{O}(mN^2)$, which means that it is important to attempt to choose both N and m to be sufficiently large to get accuracy, but not so large as to make the computation infeasibly complex. To estimate the correlation dimension ν, we compute $\log(C(r))$ and $\log(r)$ across a range of r values. We then use a sliding window linear regression to locate the region where the relationship between $\log(C(r))$ and $\log(r)$ is most linear, and the slope of that linear segment gives the estimated correlation dimension ν.

To ensure that the embedding dimension m is sufficient to fully unfold the attractor, the procedure is repeated for increasing values of m. Once the estimated correlation dimension ν stabilizes (i.e., stops increasing significantly with increasing m), this saturation indicates that the embedding is adequate and the computed value of ν is a reliable estimate of the attractor's true correlation dimension.

Our code to implement this procedure is provided in the Appendix 7.

When we implement our algorithm for varying values of T on the delay Duffing equation, we see several striking features, presented in Figure 5.7. First, there is a very prominent and detectable jump in the correlation dimension at $T = 3.6$, which notably agrees with what was observed for the largest Lyapunov exponent, and indicates the onset of chaos. Additionally, we see that right around $T = 3.9$, there is a window of relative stability, in which the attractor dimension hovers around 1 again. During this regime, the system again displays periodic motion.

Interestingly, beyond $T \approx 5.5$ and up to $T \approx 65$, the correlation dimension hovers very close to 2. This is a surprising result! On one one hand, this would seem to suggest that in this parameter range, the attractor might be an invariant torus, and what we are seeing is actually quasi-periodic motion on a torus, rather than true chaos. However, what complicates this picture is that across much of this range, the first Lyapunov exponent was calculated to be positive, suggesting sensitive dependence of initial conditions. Therefore, the exact nature of the dynamics across this parameter regime remains somewhat mysterious. This would be a productive line of inquiry for future research.

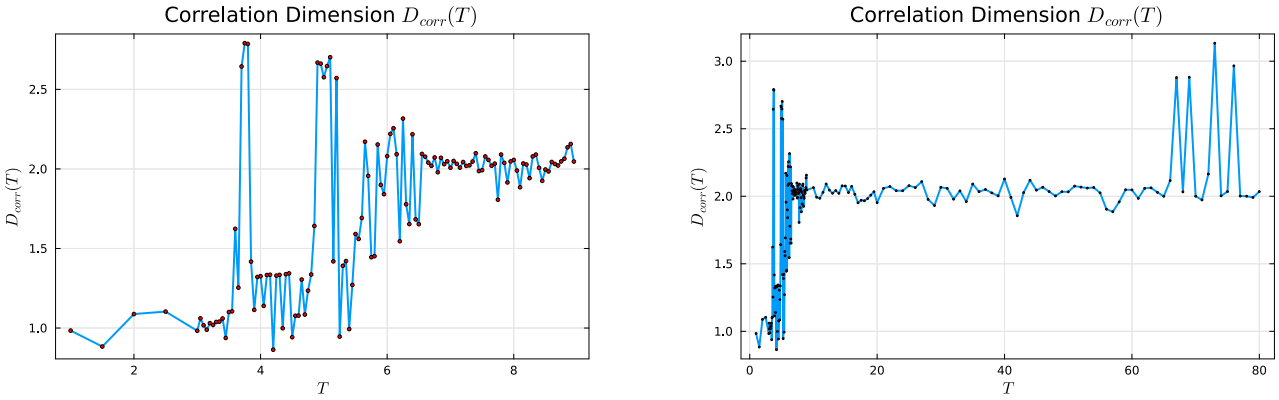

Figure 5.7: The computed correlation dimension D_{corr} of the attractor in the Delay Duffing Equation as T runs from 1 to 80. Correlation dimension was computed using the code presented in Appendix 7.2, with an embedding dimension of $m = 10$ and $15,000$ points sampled on the attractor.
Left: The correlation dimension for small values of T. For $1 \leq T \leq 3.6$, the attractor has dimension very close to 1, which is a good sanity check as we know the attractor is a limit cycle in these cases. At $T \approx 3.6$, the dimension jumps suddenly, suggesting the onset of chaos, which lasts until $T \approx 3.9$. Between $T \approx 4$ and $T \approx 5.5$, there are windows of chaos and windows of stability. After $T \approx 5.5$, the correlation dimension jumps and never comes down again, at least not in the range we see. The correlation dimension was computed at increments of $\Delta T = 0.1$.
Right: The correlation dimension for large T. Surprisingly, for T between roughly 10 and 65, the dimension hovers right around 2! This is unexpected. Between $T = 65$ and $T = 80$, we see several jumps in the attractor dimension, suggesting the onset of some kind of higher dimensional chaos. For this window, correlation dimension was computed with increment $\Delta T = 1$. It is possible there is finer level detail we were unable to resolve.

5.3 The Kaplan-Yorke Conjecture

Having now discussed both the Lyapunov spectrum and the dimension of strange attractors, we turn to an elegant and influential conjecture that links these two ideas: the Kaplan–Yorke conjecture (also sometimes referred to as the Lyapunov dimension

formula) [13]. This conjecture posits that the dimension of a chaotic attractor can be estimated directly from its Lyapunov exponents. Suppose the Lyapunov exponents are ordered from largest to smallest, $\lambda_1 \geq \cdots \geq \lambda_n$, and let j be the largest integer such that the sum $\sum_{i=1}^{j} \lambda_i \geq 0$. Then the conjectured dimension is given by the formula:

$$D_{KY} = j + \frac{\sum_{i=1}^{j} \lambda_i}{|\lambda_{j+1}|}. \tag{5.3}$$

The Kaplan-Yorke conjecture states that, for a broad class of dynamical systems, this value D_{KY} is equal to the information dimension of an attractor. While we do not define information dimension rigorously here, we refer the reader to [9]for an accessible discussion. In this work we use instead the correlation dimension D_{corr} described in the previous two sections because it has the advantage of being relatively computable. It is know that in general $D_{\text{corr}} \leq D_{\text{inf}}$, but the two are generally quite close when both can be computed.

The conjecture makes intuitive sense because it connects the stretching and contracting rates of nearby trajectories—captured by the Lyapunov exponents—to how volumes deform in phase space. The idea is that an attractor's dimension reflects how many directions exhibit sustained expansion, and the formula interpolates between dimensions based on when the accumulated stretching (the sum of exponents) transitions from growth to decay, capturing the "fractional" degree to which trajectories can spread before collapsing.

In a few special cases, the Kaplan–Yorke conjecture has been verified analytically and found to be remarkably consistent with numerical experiments in a wide range

of finite-dimensional chaotic systems. Intriguingly, there is evidence that the formula also remains valid in infinite-dimensional settings such as delay differential equations (DDEs), where rigorous extensions of standard dimension theory are more challenging. Farmer [9] for instance, found good agreement between the Kaplan–Yorke dimension and correlation dimension in the chaotic regime of the Mackey–Glass equation.

The difficulty in verifying the Kaplan-Yorke conjecture is that the it requires resolving some very small Lyapunov exponents. Resolving all Lyapunov exponents accurately is computationally difficult because smaller exponents converge much more slowly than the dominant ones, often requiring extremely long integration times. Additionally, numerical errors accumulate and amplify through repeated QR factorizations, making it challenging to maintain orthogonality and precision for subdominant directions over time. These problems are exacerbated by the nature of the Benettin algorithm for delay equations, which requires approximating the delay equation as a fairly high order ODE, making all of the associated computations much harder.

For these reasons, we have only been able to fully compute the Kaplan-Yorke dimension for a single value of the delay, and this was a herculean effort on my little laptop. Resolving Equation (5.3) to within 1×10^{-4} required $30,000,000$ steps of the Benettin algorithm with $N = 60$, corresponding to a 122 dimensional ODE.

We are, in this one case, able to confirm that the Kaplan-Yorke formula and the correlation dimension are very close to each other, with a relative error of $\approx 5.9\%$. This again confirms the feasibility of the Kaplan-Yorke conjecture in the context of delay equations, agreeing with the results of [9]. We note, however, that while the

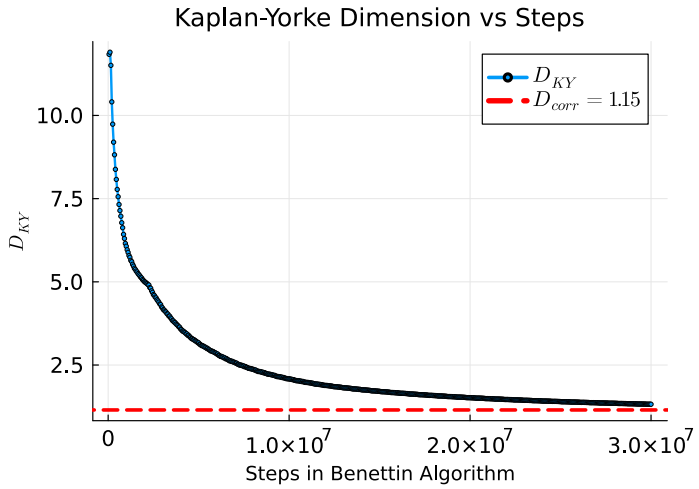

Figure 5.8: Convergence of the Kaplan-Yorke formula for the dimension of the attractor when $T = 3.65$, vs the computed correlation dimension. The Kaplan-Yorke yields a dimension of $D_{KY} = 1.221$, whereas the correlation dimension is $D_{corr} = 1.154$, a relative error of 5.89%.

truth of this conjecture is of theoretical interest, from a computational point of view, computing the Kaplan-Yorke dimension for delay equations is incomparably more expensive than computing the correlation dimension.

CHAPTER 6

CONCLUSIONS

We have explored the wild and intricate world of delay differential equations (DDEs), using the delayed Duffing equation as our guiding example. Far from a simple generalization of its undelayed counterpart, this system reveals a rich tapestry of dynamical behaviors that defy naive intuition. Our journey began with the foundations of linear DDE theory, which provided the groundwork for understanding stability and spectral properties. From there, we ventured into the realm of nonlinear oscillations, demonstrating how the method of Harmonic Balance can be a powerful and insightful approach for analyzing periodic solutions in delay systems.

A major highlight of our exploration was a delayed analogue of Hilbert's 16th problem. We showed that even simple-looking second-order polynomial DDEs can possess infinitely many limit cycles—a striking departure from the finite-limit-cycle behavior of ordinary differential equations. Through the method of characteristic oscillations, we constructed a family of such systems, offering both theoretical insight and constructive examples.

In the final leg of our tour, we turned to the fascinating terrain of chaotic dynamics in DDEs. By implementing numerical algorithms for computing Lyapunov exponents and attractor dimensions, we observed a general trend: as the delay parameter increases, so too does the effective dimensionality of the chaotic attractor. Our numerical investigations lent further support to the Kaplan–Yorke conjecture, which proposes a deep and elegant connection between the Lyapunov spectrum and the fractal dimension of chaotic sets.

We hope this tour through delay equations not only illuminates the rich behavior these systems can exhibit, but also equips the reader with a suite of analytical and numerical tools to investigate the delay systems that arise in their own research. Whether one encounters delay in physics, biology, control theory, or beyond, the insights and methods presented here underscore just how much structure—and surprise—can lie hidden in the past.

CHAPTER 7

APPENDIX: CODE

7.1 The Benettin Algorithm for DDEs

Below is our code for implementing the Benettin algorithm to compute the Lyapunov Spectrum for the delayed Duffing equation, $x''+x(t-T)+x^3=0$. Code is implemented in Julia version 1.11.5.

```
using DifferentialEquations
using LinearAlgebra
using Plots
using CSV
using DataFrames
using Base.Threads



function F_i(i, X, Y,N,T)
#Computes dxi/dt using method of lines.
    if i==1
        return Y[1]
    elseif 2<=i && i<=N
        return N*(X[i-1]-X[i+1])/(2*T)
    else
        return N*(X[N]-X[N+1])/T
    end
```

```
end


function G_i(i, X, Y,N,T)
#Computes dyi/dt using method of lines.
     if i==1
        return -X[1]^3-X[N+1]
    elseif 2<=i && i<=N
        return N*(Y[i-1]-Y[i+1])/(2*T)
    else
        return N*(Y[N]-Y[N+1])/T
    end
end



function discrete_DDE!(du, u, p, t)

#Defines system of 2N+2 ODEs for Delay Duffing using method of
   lines.

    N,T = p.N, p.T

    X = @view u[1:N+1]
    Y = @view u[N+2:end]

    dX = @view du[1:N+1]
```

```
    dY = @view du[N+2:end]

    for i in 1:N+1
        dX[i] = F_i(i, X, Y,N,T)
        dY[i] = G_i(i, X, Y,N,T)
    end
end



function Jac_template(N,T)

#Function for computing Jacobian of discrete_DDE, for use in
   variational equation. Most entries in Jacobian
#are constant, so precompute them for efficiency.

    block = zeros(N,N+1)
    for i in 1:N-1
        block[i,i]=N/(2*T)
        block[i,i+2]=-N/(2*T)
    end
    block[N,N]=N/T
    block[N,N+1]=-N/T

    J = [zeros(1,N+1) 1 zeros(1,N)
```

```
        block zeros(N,N+1)
        0 zeros(1,N-1) -1 zeros(1,N+1)
        zeros(N,N+1) block]
end



function update_J!(J,X)
#Function for updating Jacobian to J(X). Only entry N+2,1
   depends on X,Y, and in fact, only on X.
    N = (size(J)[1]-2) \div 2
    J[N+2, 1] = -3 * X[1]^2
    nothing
end



struct BenettinParams
#Container for parameters to get passed into benettin_system
   and benettin_loop
    N::Int
    T::Float64
    m::Int
    J::Matrix{Float64}
end
```

```
function benettin_system!(du_all, u_all, p, t)

#Forms augmented ODE using variational equation

    # Decompose combined state vector
    N, T, m, J = p.N, p.T, p.m, p.J
    u = @view u_all[1:2N+2]
    X = @view u[1:N+1]
    Y = @view u[N+2:end]
    Z = reshape(view(u_all,2N+3:length(u_all)), 2N+2, m)

    du = @view du_all[1:2N+2]
    dZ = reshape(view(du_all,2N+3:length(du_all)), 2N+2, m)

    # Compute du = f(u)
    discrete_DDE!(du, u, p, t)

    # Get Jacobian at current state
    update_J!(J,X)
    @assert J[N+2,1] == -3*X[1]^2 "J did not update!"

    # Variational equation: dZ/dt = J * Z
    dZ .= J * Z
```

```
end




function benettin_loop(N, T, m, Dt; steps=1000,burn = 200,
   L1listboole = false)

  #Performs Benettin algorithm on augmented system, using m
     perturbation vectors, step size of \Delta t.
  #burn is number of initial steps to ignore to improve
     convergence speed.
  #L1listboole - if true, plots value of largest Lyap exp vs
     number of steps



    # Initial condition: state vector
    u = [.2*ones(N+1); zeros(N+1)]

    # Initial condition: m orthonormal perturbations
    Z = Matrix{Float64}(I, 2N+2, m)

    # Accumulator for log-diagonal of R
    accum = zeros(m)
```

```
J = Jac_template(N,T)


p = BenettinParams(N, T, m, J)


start = time()


#Pre-integrate to get trajectory onto limit set
    tspan1 = (0.0, 400)
    prob = ODEProblem(discrete_DDE!, u, tspan1, p)
    # Integrate
    sol1 = solve(prob, Vern9(); reltol=1e-6, abstol=1e-6)
    # Unpack result
    u = sol1.u[end]


L1list=[]
for k in 1:steps


    # Pack full state vector
    uZ = [u; vec(Z)]
    tspan = (0.0, Dt)
    prob = ODEProblem(benettin_system!, uZ, tspan, p)


    # Integrate
    sol = solve(prob, Vern9(); reltol=1e-6, abstol=1e-6)
```

```
        # Unpack result

        uZ = sol.u[end]

    #Extract perturbation vectors
        Z .= reshape(view(uZ, 2N+3:(m+1)*(2N+2)), 2N+2, m)

    #Extract state vectors
        u = view(uZ,1:2N+2)

        # QR factorization and accumulation
        Q, R = qr(Z)

        if k > burn
            accum += log.(abs.(diag(R)))
        end
        Z .= Matrix(Q)  # reset perturbations to orthonormal
           directions

        @assert norm(Z'Z - I) < 1e-10 "Z is not orthonormal!"

        if L1listboole && k>burn
            if k % 20 ==1
                push!(L1list,accum[1]/((k-burn)*Dt))
            end
```

```
        end

    end
    fin = time()

    println("Loop with T=$T took $(fin-start) seconds")

    # Return Lyapunov exponents
    return [sort(accum,rev=true) ./ ((steps-burn) *Dt),sol1,
       L1list]

end



function build_data(filename)

#Takes input csv files with values of delay T in column 1.
   Computes Lyap spectrum for each value of T, writes out to
   same file.

    df = CSV.read(filename, DataFrame)
    Tvalues = Vector(df[2:end, 1])
    n = length(Tvalues)

  #New data structure to avoid data races
```

```
    results = Vector{Union{Nothing, Vector{Float64}}}(nothing,
        n)

  #Multi-thread because it's a big computation
    Threads.@threads for i in 1:n
        T = Tvalues[i]

    #Make N sufficiently large for each value of T
        N = max(Int(ceil(20*T)),40)

    #Write lyap spectrum for T to results
        results[i] = benettin_loop(N, T, 10, .01; steps=20000)
            [1]
    end

    for i in 1:n
        if results[i] !== nothing
            df[i+1, 2:length(results[i])+1] = results[i]
        end
    end

    CSV.write(filename, df)
end
```

7.2 Correlation Dimension Computation

This code is in two pieces. We perform the numerical integration in Mathematica, because Mathematica's integrator for DDEs and built in interpolation is convenient for our purposes. The delay embedding is computed in Mathematica to take advantage of the built in interpolation feature form NDSolve. These results are then exported to a CSV file, which we then import into Julia. The correlation dimension is computed in Julia, due to Julia's optimization of linear algebra functions. The computation is multi-threaded because it's such a big lift on my little laptop. Computations were performed in Mathematica 13 and Julia 1.11.5.

Mathematica:

```
delayEmbedding[f_, m_, \[Tau]_, \[CapitalDelta]t_, t0_] :=
  Module[{},
   (*f is interpolating funct from NDSolve*)
   embedding =
    Table[f[\[CapitalDelta]t*i + j*\[Tau] + t0], {i,
      0, (tmax - t0 - m*\[Tau])/\[CapitalDelta]t - 1}, {j, 0,
      m - 1}];
   Return[Transpose[embedding]]
   ];

Tmin = 1; Tmax =20; Tinc =.1;

AbsoluteTiming[
```

```
 For[T = Tmin, T <= Tmax, T += Tinc,

  (*Performs numerical integration and delay embedding for T
     values \
between Tmin and Tmax with step Tinc.
  Exports resulting embeddings to CSV*)

  T = 1. T;

    Clear[x, sol, xsol, embedding];
    tmax = 2000;
    sol =
     NDSolve[x''[t] + x[t - T] + x[t]^3 == 0 && x[t /; t <= 0]
        == .1,
      x, {t, 0, tmax}, MaxSteps -> Infinity];
    xsol = x /. First[sol];
    x[t_] := xsol[t];

    (*m=9 is sufficiently large embedding dimension. \[Tau]
       taken to be small \
irrational number*)

    embedding = delayEmbedding[x, 9, Sqrt[2]/15, .15, tmax -
       1202];
```

```
    (*Format for CSV*)
    Trow = {"#T",
    NumberForm[T, {5, 2}, NumberPadding -> {"0", "0"}]};
  combined = Join[{Trow}, embedding];
  Tstring =
   ToString[NumberForm[T, {3, 2}, NumberPadding -> {"0", "0"
      }]];
  filename = "T=" <> Tstring <> "_Embedding.csv";

  (*Progress Statement*)
  If[
   StringQ@Export[filename, combined, "CSV"],
   Print["Export successful: ", filename],
   Print["Export failed."]]
  ]
 ]
```

Julia:

```
using Base.Threads
using Statistics, LinearAlgebra
using Plots
using CSV, DataFrames
```

```
# Euclidean distance between two vectors
function euc(a, b)
    sqrt(sum((a .- b).^2))
end



function compute_pairwise_distances_parallel(X)
  #input: array X. Each column is distinct vector
  #threaded computation of all pairwise distances

    n = size(X, 2)  # number of points is the number of columns
    num_pairs = div(n * (n - 1), 2)

  #container for pairs of distances
    dists = Vector{Float64}(undef, num_pairs)

    # Mapping from (i, j) to index in dists:
    function index(i, j, n)
        return div((2*n - i)*(i - 1), 2) + (j - i)
    end



    @threads for i in 1:n-1
        for j in i+1:n
```

```
            idx = index(i, j,n)
            xi = @view X[:,i]
            xj = @view X[:,j]
            dists[idx] = euc(xi,xj)
        end
    end

    return dists
end

rs
function correlationSum(r, distances, len)
  #counts distances < r in the vector distances
  count(d -> d < r, distances) / binomial(len, 2)
end



function parallel_correlation_sums(rs, distances, len)
  #threaded function, takes vector of distances rs and vector
     of pairwise distances, computes C(r) for each r.

    Csums = Vector{Float64}(undef, length(rs))

    @threads for i in 1:length(rs)
        Csums[i] = correlationSum(rs[i], distances, len)
```

```
    end

    return Csums
end



function logspace(a, b, n)
  #Helper function, acts like Matlab's logspace
    exponents = range(a, stop=b, length=n)
    return 10 .^ exponents
end




function correlation_dimension(rs, Csums; windowSize=10,
   r2_threshold=0.8)
#Takes vector rs and vector Csums and performs sliding window
   linear regression to determine correlation dimension

    # Step 1: Clean and log-transform
    log_pairs = [(log10(r), log10(c)) for (r, c) in zip(rs,
       Csums)
                if isfinite(log10(r)) && isfinite(log10(c))]
```

```
    num_points = length(log_pairs)
    if num_points < windowSize
        return missing  # Not enough data
    end

    # Step 2: Sliding window regression
    slopes = Tuple{Float64, Float64, Float64, Float64}[]

    for i in 1:(length(log_pairs) - windowSize + 1)
        window = log_pairs[i:i+windowSize-1]
        x_vals = [p[1] for p in window]
        y_vals = [p[2] for p in window]

        X = hcat(ones(length(x_vals)), x_vals)
        B = X \ y_vals  # least squares fit
        slope = B[2]
        intercept = B[1]

        y_pred = X * B
        ss_res = sum((y_vals .- y_pred).^2)
        ss_tot = sum((y_vals .- mean(y_vals)).^2)
        r2 = 1 - ss_res / ss_tot

        residual_std = std(y_vals .- y_pred)
```

```
        push!(slopes, (mean(x_vals), slope, r2, residual_std))
    end

    # Then select:
    slope_threshold = 0.8

#Slope threshold prevents algorithm from select region with
    unreasonably low slope.

    filtered = filter(s -> s[3] > r2_threshold && s[2] >
        slope_threshold, slopes)

    if isempty(filtered)
        return missing  # or Missing(["NoStableSlope", Csums,
            rs])
    end

    min_std_dev = Inf
    best_fit = nothing

    for tup in filtered
        if tup[4] < min_std_dev
            min_std_dev = tup[4]
            best_fit = tup
```

```
        end
    end

    return best_fit
end









# Path to directory with embedding CSVs

dir_path = "/home/directory/path/goes/here"
results_file = dir_path*"/filename/goes/here"

# Get all CSV files in the directory
csv_files = filter(f -> endswith(f, ".csv"), readdir(dir_path;
   join=true))

for file in csv_files
    println("Processing: ", file)

    # Read CSV with no header
    lines = readlines(file)
    T = parse(Float64, split(lines[1], ",")[2])
```

```
    embedding = CSV.read(IOBuffer(join(lines[2:end], "\n")),
       DataFrame; header=false, types= Float64)

    println("T = ", T)



    # Remove the first row and convert the rest to a matrix
    embedding = Matrix(embedding)
    m, N = size(embedding)
    println("Embedding size: ", m, " x ", N)

    start=time()
    dists=compute_pairwise_distances_parallel(embedding);
    rs = logspace(-5, 1.6, 50)
    Csums = parallel_correlation_sums(rs,dists,length(dists))
    cdim=correlation_dimension(rs, Csums; windowSize=15,
       r2_threshold=.9)[2]
    done=time()
    print("Computation Took:", done-start," seconds.
       Correlation Dimension = ",cdim)
    row = DataFrame(T = [T], CDim = [cdim])
    CSV.write(results_file, row; append=true)
    println("Logged: T = ", T, ", CDim = ", cdim)

    println("Finished processing T = ", T)
```

```
    println()
end
```

BIBLIOGRAPHY

[1] J. Banks, J. Brooks, G. Cairns, G. Davis, and P. Stacey. On devaney's definition of chaos. *American Mathematical Monthly*, 99(4):332–334, 1992.

[2] Richard Bellman and Kenneth L. Cooke. Differential-difference equations. by richard bellman and kenneth l. cooke. pp. xvi, 465. 1963. (academic press). *The Mathematical Gazette*, 51(377):276–276, 1967.

[3] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. part ii: Numerical applications. *Meccanica*, 15:21–30, 1980.

[4] Robert M. Corless, Gaston H. Gonnet, David E. G. Hare, David J. Jeffrey, and Donald E. Knuth. On the lambert w function. *Advances in Computational Mathematics*, 5(1):329–359, 1996.

[5] Matthew Davidow, B. Shayak, and Richard H. Rand. Analysis of a remarkable singularity in a nonlinear dde. *Nonlinear Dynamics*, 90, 10 2017.

[6] Philip J. Davis. Gamma function and related functions. In Milton Abramowitz and Irene A. Stegun, editors, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, chapter 6, pages 253–293. Dover Publications, New York, 1972. See Section 6.2: Beta Function, p. 258.

[7] Robert L. Devaney. *An Introduction to Chaotic Dynamical Systems*. Addison-Wesley, 2nd edition, 1989. Introduces the classical Devaney definition of chaos.

[8] Gustav Doetsch. *Introduction to the Theory and Application of the Laplace Transformation*. Springer-Verlag, Berlin, 1974. Translated from the 5th German edition.

[9] J. Doyne Farmer. Chaotic attractors of an infinite-dimensional dynamical system. *Physica D: Nonlinear Phenomena*, 4(3):366–393, 1982.

[10] Thomas Erneux. *Applied Delay Differential Equations*, volume 200 of *Applied Mathematical Sciences*. Springer, New York, 2009.

[11] Mitchell J. Feigenbaum. Quantitative universality for a class of nonlinear transformations. *Journal of Statistical Physics*, 19:25–52, 1978.

[12] Bernold Fiedler, Alejandro Lopez-Nieto, Richard H. Rand, Si Mohamed Sah, Isabelle Schneider, and Babette deWolff. Coexistence of infinitely many large, stable, rapidly oscillating periodic solutions in time-delayed duffing oscillators. *Journal of Differential Equations*, 268(10):5969–5995, 2020.

[13] P. Frederickson, J. L. Kaplan, E. D. Yorke, and J. A. Yorke. The lyapunov dimension of strange attractors. *Journal of Differential Equations*, 49(2):185–207, 1983.

[14] Peter Grassberger and Itamar Procaccia. Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9(1-2):189–208, 1983.

[15] John Guckenheimer and Philip Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, volume 42 of *Applied Mathematical Sciences*. Springer, 2nd edition, 2002.

[16] Jack K. Hale and Sjoerd M. Verduyn Lunel. *Introduction to Functional Differential Equations*, volume 99 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1993.

[17] David Hilbert. Mathematical problems. *Bulletin of the American Mathematical Society*, 8:437–479, 1902.

[18] Keum-Shik Hong and Hahn Park. Boundary control of container cranes as an axially moving string system. In *IFAC Proceedings Volumes*, volume 38, pages 132–137. Elsevier, 2005.

[19] Yu. Ilyashenko. *Finiteness Theorems for Limit Cycles*, volume 94 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991. MR 92k:58221.

[20] James L. Kaplan and James A. Yorke. Ordinary differential equations which yield periodic solutions of differential delay equations. *Journal of Mathematical Analysis and Applications*, 48(2):317–324, 1974.

[21] Faouzi Lakrad, Jamol Pender, and Richard Rand. Queues with delayed information: A dynamical systems perspective. *SIAM Journal on Applied Dynamical Systems*, 21(1):676–713, 2022.

[22] Jérôme Losson, Michael C. Mackey, Richard Taylor, and Marta Tyran-Kamińska. *Density Evolution Under Delayed Dynamics: An Open Problem*, volume 38 of *Fields Institute Monographs*. Springer, New York, NY, 1st edition, 2020.

[23] James D. Mackey and Andrzej Lasota. *Chaos, Fractals and Noise: Stochastic Aspects of Dynamics*, volume 97 of *Applied Mathematical Sciences*. Springer, 1989.

[24] Carlos Gustavo T. Moreira, Maria Josè Pacifico, and Sergio Romana Ibarra. Hausdorff dimension, lagrange and markov dynamical spectra for geometric lorenz attractors. *Bulletin of the American Mathematical Society*, 57(2):271–292, 2020.

[25] V. I. Oseledec. A multiplicative ergodic theorem. lyapunov characteristic numbers for dynamical systems. *Trudy Moskovskogo Matematicheskogo Obscestva*, 19:179–210, 1968. English transl. in *Transactions of the Moscow Mathematical Society*, vol. 19, pp. 197–231.

[26] R.H. Plaut and J.-C. Hsieh. Non-linear structural vibrations involving a time delay in damping. *Journal of Sound and Vibration*, 117(3):497–510, 1987.

[27] Richard H. Rand. *Lecture Notes on Nonlinear Vibrations*. Cornell University, 2005. Available online at https://dspace.library.cornell.edu/handle/1813/28989.

[28] David Ruelle. Characteristic exponents and invariant manifolds in hilbert space. *Annals of Mathematics*, 115(2):243–290, 1982.

[29] Si Mohamed Sah, Bernold Fiedler, B. Shayak, and Richard H. Rand. Unbounded sequences of stable limit cycles in the delayed duffing equation: an exact analysis. *Nonlinear Dynamics*, 103, 01 2021.

[30] Marco Sandri. Numerical calculation of lyapunov exponents. *The Mathematica Journal*, 6(3):78–84, summer 1996. Marco Sandri, University of Verona.

[31] Tim Sauer, James A. Yorke, and Martin Casdagli. Embedology. *Journal of Statistical Physics*, 65(3-4):579–616, 1991.

[32] Larry F. Shampine and Sylvester Thompson. Numerical solution of delay differential equations. In Balakumar Balachandran, Tamás Kalmár-Nagy, and David E. Gilsinn, editors, *Delay Differential Equations: Recent Advances and New Directions*, pages 245–270. Springer, 2009.

[33] Charalampos Skokos. The lyapunov characteristic exponents and their computation. *The European Physical Journal Special Topics*, 164:1–16, 2008.

[34] J. C. Sprott. Lyapunov exponents for delay differential equations. Technical report, Department of Physics, University of Wisconsin, Madison, WI, USA, oct 2006. Technical note, dated October 4, 2006.

[35] Steven H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering.* Chapman and Hall/CRC, New York, 3rd edition, 2024.

[36] M. K. Suchorsky, S. M. Sah, and R. H. Rand. Using delay to quench undesirable vibrations. *Nonlinear Dynamics*, 62(1-2):407–416, 2010.

[37] Floris Takens. Detecting strange attractors in turbulence. In David A. Rand and Lai-Sang Young, editors, *Dynamical Systems and Turbulence, Warwick 1980*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381. Springer, 1981.

[38] Huailei Wang, Haiyan Hu, and Zaihua Wang. Global dynamics of a duffing oscillator with delayed displacement feedback. *International Journal of Bifurcation and Chaos*, 14(08):2753–2775, 2004.

[39] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications I: Fixed-Point Theorems*, volume 66 of *Applied Mathematical Sciences*. Springer, 1986.